Collaboration to Clarify the Cost of Curation

D5.2—Roadmap report

Deliverable Lead:	Jisc
Related Work package:	WP5
Author(s):	Paul Stokes—Jisc Neil Grindley—Jisc
Dissemination level:	Public
Submission date:	28 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	15 Jan 2015	First draft	PLSS
0.02	12 Feb 2015	Expanded draft	PLSS
0.03	24 Feb 2015	First review	PLSS
0.04	28 Feb 2015	Second Review	NG
1.00	05 Mar 2015	Release 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.

{Third party acknowledgements}

D5.2—Roadmap 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):	Paul Stokes (Jisc), Neil Grindley (Jisc)

Table of Contents

Acknowledgements 3
Table of Contents 4
Figures 6
Tables 8
Executive Summary 9
1 Introduction 10
2 Methodology 11
2.1 Desk research, literature review and critique 11
2.2 Development process 11
2.3 Draft publication 12
2.4 Community review 12
2.5 Final publication 13
2.6 Post project monitoring 13
3 Draft format 14
3.1 Early thoughts 14
3.2 Why we went for the format we did 17
3.3 How we honed it down 18
3.4 The vision and messages 22
3.4.1 THE MESSAGE BEHIND THE MESSAGES 23
3.5 The final draft version 24
3.5.1 POSTCARDS 25
4 Synthesis of responses 26
4.1 Response framework 26
4.2 Response to explicit questions 27
4.2.1 VISION 27
4.2.2 MESSAGE 1 28
4.2.3 MESSAGE 2 29
4.2.4 MESSAGE 3 30
4.2.5 MESSAGE 4 31
4.2.6 MESSAGE 5 32
4.2.7 MESSAGE 6 33
4.2.8 GENERAL THEMES 34
4.3 Key points addressed as a result 38
5 Final version 39
5.1 Summary of changes 39
5.1.1 SPECIFIC ISSUES 39
5.1.2 GENERAL ISSUES 39
5.2 Final Roadmap 40
5.3 Final Roadmap—condensed version 53
5.4 Updates and activities still to come 61

6 The way ahead 62
7 Conclusion 63
References 64
Appendices 65
Appendix A Literature Review 66
RESEARCH ROADMAPS 66
INFRASTRUCTURE ROADMAPS 69
DIGITAL PRESERVATION ROADMAPS 75
ACTION PLANS RELATED TO DIGITAL PRESERVATION 77
OTHER TEXTS REVIEWED 83
Appendix B Draft Roadmap 85
ROADMAP 85
POSTCARDS 98
Appendix C Structured Interview Template 133
Appendix D Webinar Agenda 135
WEBEX AGENDA 135
Appendix E On-line questionnaire 136
Appendix F iPRES 2014 Melbourne Roadmap Workshop abstract 144
Appendix G iPRES 2014 Melbourne Roadmap Workshop report 146
Appendix H iPRES Workshop blog 152
DEFINING A ROADMAP FOR ECONOMICALLY EFFICIENT DIGITAL CURATION 152
Appendix I 4C Conference Roadmap presentation—Nov 2014 155

Figures

Figure 1—Early roadmap mind map—overall concept 15
Figure 2—Early roadmap mind map—content 15
Figure 3—Early roadmap mind map—The end game 16
Figure 4—Early roadmap mind map—The principles underpinning the roadmap 16
Figure 5—Early roadmap mind map—document structure 17
Figure 6—Roadmap structure, messages and action plan 22
Figure 7—Final roadmap 52
Figure 8—Final roadmap—condensed version 61
Figure 9—Roadmap booklet 97
Figure 10—Actions for Curation Practitioners—English 98
Figure 11—Actions for Curation Researchers—English 99
Figure 12—Actions for Data Users—English 100
Figure 13—Actions for Managers—English 101
Figure 14—Actions for Member Organisations—English 102
Figure 15—Actions for Policy Makers—English 103
Figure 16—Actions for Solution Providers—English 104
Figure 17—Actions for Curation Practitioners—German 105
Figure 18—Actions for Curation Researchers—German 106
Figure 19—Actions for Data Users—German 107
Figure 20—Actions for Managers—German 108
Figure 21—Actions for Member Organisations—German 109
Figure 22—Actions for Policy Makers—German 110
Figure 23—Actions for Solution Providers—German 111
Figure 24—Actions for Curation Practitioners—French 112
Figure 25—Actions for Curation Researchers—French 113
Figure 26—Actions for Data Users—French 114
Figure 27—Actions for Managers—French 115
Figure 28—Actions for Member Organisations—French 116
Figure 29—Actions for Policy Makers—French 117
Figure 30—Actions for Solution Providers—English 118
Figure 31—Actions for Curation Practitioners—Portuguese 119
Figure 32—Actions for Curation Researchers—Portuguese 120

Figure 33—Actions for Data Users—Portuguese 121
Figure 34—Actions for Managers—Portuguese 122
Figure 35—Actions for Member Organisations—Portuguese 123
Figure 36—Actions for Policy Makers—Portuguese 124
Figure 37—Actions for Solution Providers—Portuguese 125
Figure 38—Actions for Curation Practitioners—Dutch 126
Figure 39—Actions for Curation Researchers—Dutch 127
Figure 40—Actions for Data Users—Dutch 128
Figure 41—Actions for Managers—Dutch 129
Figure 42—Actions for Member Organisations—Dutch 130
Figure 43—Actions for Policy Makers—Dutch 131
Figure 44—Actions for Solution Providers—Dutch 132
Figure 45—Challenges identified at the Melbourne workshop breakouts 150
Figure 46—Roadmap issues identified at the Melbourne workshop breakouts 151
Figure 47—Roadmap presentation from the 4C Confrence November 2014 160

Tables

Table 1—Roadmap idea template 18
Table 2—DigitalPreservationEurope (DPE) Research Roadmap 67
Table 3—Blue Ribbon Task Force on Sustainable Digital Preservation and Access 68
Table 4—Open research challenges and research roadmap for SCAPE 68
Table 5—A Strategic Vision for UK e-Infrastructure 69
Table 6—E-Infrastructure Roadmap 70
Table 7—Science Data Infrastructure Roadmap 71
Table 8—2011 Strategic Roadmap for Australian Research Infrastructure 72
Table 9—e-IRG Roadmap 2012 73
Table 10—EGI Technology Roadmap 74
Table 11—Digital Cultural Heritage Roadmap for Preservation – Open Science Infrastructure for DCH in 2020 (draft version) 75
Table 12—APARSEN Roadmap 76
Table 13—LERU Roadmap for Research Data 77
Table 14—Digital Repositories Roadmap: looking forward 79
Table 15—The Resource Discovery Taskforce (RDTF) Vision 80
Table 16—Big Data Roadmap, Association of the British Pharmaceutical Industry (APBI) 81
Table 17—Roadmap for the European Platform on Life Cycle Assessment: facilitating data collection and sustainability assessments for policy and business 82

Executive Summary

The original proposal for the 4C project placed great deal of emphasis on the roadmap being a key output for the project. As a consequence of this, substantial effort was dedicated to the development of the roadmap and the early buy-in for its content, aims and objectives (a task described as “crucial” in the DoW).

The project has from day one been guided by the principle of being “open and social”. In essence this means we publish our outputs early in draft form with a view to gathering input from the community and letting the community guide the agenda. The roadmap development process adhered to that principle and, after a period of internal brainstorming and external consultation, the draft roadmap was released to the community in August 2014.

Very early on in the process the project as a whole developed a series of governing principles that shaped the roadmap authoring process and the final output. These were:

- to produce an accessible document
- to produce a meaningful document
- to produce a SMART document
- to produce a validated document

Following these principles resulted in the development of a short (26 page) roadmap document with a vision and six key messages, each with a brief narrative, a description of the associated benefits, and an action table with actions for each of the seven key stakeholder groups previously identified in other project outputs.

The community were both invited to respond using on-line tools and actively pursued through a series of group and individual dissemination and engagement activities.

The community were on the whole very supportive of the draft albeit concerned about the future of the roadmap. A synthesis of their responses was used to generate a short list of changes to the original draft focusing upon:

- Solution provider time scales
- Solution provider support and involvement
- Actions for data producers
- The rationale behind the roadmap format
- Further signposting to 4C resources

The final version of the roadmap was released in February 2015. Its uptake and the progress by the community towards achieving the vison outlined in it will be monitored by members of the post project consortium.

1 Introduction

The Roadmap for digital curation that accompanies this report is arguably one of the most important outputs from the 4C project. It was assigned significant resources—all partners had assigned effort—and a relatively long development period.

Very early on in the process the project as a whole developed a series of governing principles that shaped the authoring process and the final output. In short these principles were:

- to produce an accessible document that would be read and acted upon by decision makers
- to produce a document that was meaningful for all the previously identified stakeholders
- to produce a SMART[1] document
- to produce a document that was validated by the community

These principles were acted upon throughout the course of the development of the roadmap document which was published in draft form in August2014 and in its final form in February 2015.

This report provides an overview of the processes undertaken to shepherd the roadmap from its early outline through to the finished document. In particular it covers the open and social relationship between the project and the digital curation community and a synthesis of the results of interactions and reactions to the roadmap.

DoW Roadmap*

"The purpose of this activity is to arrive at coherent and evidence-based recommendations for future action and strategy in relation to the economic aspects of digital curation. The focus will be on measures that will assist diverse types of organisations to better understand and take control of the cost of managing digital assets over varied timescales, including the provision of cost-effective solutions and services to others. This roadmap report will synthesise and exploit the valuable intelligence that emerges from the other work packages and will also ensure that the content and conclusions are complementary and non-duplicative of work being taken forward by others. The output will be a report and an indicator of its success will be the level of input and validation it receives from the broad community as it is being drafted and finalised. An internal progress indicator for the roadmap will be the amount of early effort that starts to logically accrue to building an evidence-base to underpin the final report. A substantial amount of effort is dedicated to the Roadmap and early buy-in for its content, aims and objectives is crucial

Final Roadmap Report: Final report containing community validated analysis and recommendations for future action and investment strategies for the promotion of affordable digital curation solutions and services

[1] SMART—an acronym generally associated with objectives that are Specific, Measurable, Achievable, Realistic and Time bound (they have a specific target date for completion)

2 Methodology

The methodology we employed broke down into the following stages:

- Desk research, literature review and critique of associated research and publications
- Brainstorming
- Draft publication
- Community review
- Final publication
- Post project monitoring

2.1 Desk research, literature review and critique

Desk research was undertaken In order to establish how our roadmap might mesh with other related publications and initiatives. On top of this requirement, we wanted to establish the core content and stylistic requirements for our document. We wished to ascertain what the 'crowd' thought a roadmap should be and to see what worked for other roadmaps (and—perhaps more importantly—what didn't work). We also wished to see how effective past roadmaps had been over time (the thinking being that there would be little point in emulating a document if it had proven to be ineffectual in the long run). Anecdotally it seemed that critical acclaim on release was by no means a guarantee of long term success. Although we didn't have a large enough sample to make a definitive judgement (and most of those roadmaps were too fresh to draw conclusions) it would appear that the most critical factors in the success of a roadmap are: a community to drive it forward; and getting it in front of the right people, namely, those with the power to initiate change. The latter point was a major factor in our early decision to produce a short, easily assimilated roadmap aimed at decision makers.

The literature review can be found in Appendix A.

2.2 Development process

Following on from the literature review stage we undertook a series of iterative brainstorming sessions ranging from small internal project task groups through to sessions involving all project partners and the 4C Advisory Board. Early sessions were designed to establish a vision and time scale. They were attempting to answer the fundamental questions of what the roadmap was intended to achieve and where the community should be after a given "period"[2] had elapsed.

The needs and gap analysis work[3] and the study of stakeholders and stakeholder initiatives[4] undertaken earlier in the project informed these discussions. Having established the core vision of the roadmap there followed a second round of brainstorming to hone a set of principles that we felt would be prerequisites for achieving the vision.

Although the group had many domain experts we also widened the consultation to incorporate views from experts outside the project using a combination of targeted structured interviews (the template for

[2] At this stage the period in question was not yet decided upon.

[3] D3.1—Evaluation of Cost Models and Needs & Gaps Analysis [http://4cproject.eu/d3-1]

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

which can be found in Appendix C), focus groups and workshops. In addition most public facing events had a Roadmap agenda item where current issues identified by the roadmap task group were addressed.

2.3 Draft publication

The draft road map—a copy of which can be found in Appendix B, Roadmap—was published on the 4C website[5] and circulated in PDF format in early August 2014. Later in the year it was published in a number of alternative formats including an on-line interactive version[6], a print version and a series of print and PDF postcards[7] encapsulating the key actions for each of the stakeholder groups in five European languages[8]—see Appendix B, Postcards.

It is perhaps worth pointing out that the project chose to publish the draft versions in these multiple formats rather than wait until the end of the project and publish the final version in order to maximise the dissemination of the message and to gain the widest ranging and best possible interaction with the digital curation community. Had we waited to the end we would have had neither the time nor the resource to achieve the same impact.

Our decision to publish early and extensively was also influenced by the advice that we had received as part of our period 1 review to "*Take* [a] *position (this would help to elicit feedback) and make recommendations"*.

2.4 Community review

Community validation was approached through the use of a range of channels. These included:

- Focus groups
- Workshops (in particular at iPRES in Melbourne in October 2014—see Appendix F for a copy of the abstract and Appendix G for a report about the outcomes)
- Directed requests for responses from influential, knowledgeable members of the community
- An on-line questionnaire (Appendix D shows the question set)
- A dedicated session at the 4C Conference in November[9] (The presentation can be found in Appendix H)
- Web page with commenting enabled

One measure of the impact of the draft road map can be seen in the web statistics for the roadmap page. Despite only being published in the last quarter of the project, it was the second most popular landing page on the site (the most popular being the home page) over the lifetime of the project. It also had the second highest number of "hits" (in other words it was the second most viewed page on the 4C website) and had an average page viewing time three times longer than the average viewing time for the site as a whole.

[5] http://4cproject.eu/d5-1-draft-roadmap

[6] http://4cproject.eu/int-roadmap

[7] http://4cproject.eu/rm-resources

[8] Dutch, English, French, German and Portuguese

[9] The webcast of the session can be seen at http://www.dpconline.org/events/webcast4canddpa2014/1324-4cwebcastroadmapdayone

2.5 Final publication

The responses to the roadmap were collected right up to the last possible phase of the project. They were then combined into a single synthesis (the core parts of which can be found in Section 4) and used to define a set of changes to be applied to the daft before it was published as a finalised document. It was very gratifying for the team, especially given the degree of consultation in the initial development phase, to find that the community had a very positive response overall. There were relatively few changes needed to bring it into line.

2.6 Post project monitoring

Strictly speaking the post project monitoring of the roadmap, published at the end of the funded period, is out of scope for the project. However, we have considered who will take ownership of the document and how its progress will be monitored. These issues are addressed in both Section 6 of this document and in the 4C Project Sustainability Plan.

3 Draft format

The preparation of the draft road map had a number of distinct phases.

- Desk research
- Phase one Internal brainstorming to establish the ultimate aim of the roadmap and to set some baseline parameters
- External consultation in the form of webinars, focus groups and structured interviews
- Phase two internal brainstorming to hone the central theme (which became known as the vision later in the process) and time scale
- Phase three internal brainstorming to establish the pathways that would ultimately result in the fulfilment of the vision and hence the actions that need to be encouraged
- Phase four internal brainstorming to establish the format (which ultimately became the 6 messages) specific target audience and actions
- Publication of the draft

The desk research (literature review) has already been referred to in the previous section. The results can be found in Appendix A

The roadmap was discussed as part of the agenda it a number of webinars, with two (held in June 2014) dedicated solely to the roadmap. In both case a standard agenda was used (see Appendix D). The webinar format proved to be especially fruitful, enabling the team to bring together domain experts from around the world into the same “room” at a mutually convenient time with minimum disruption to their schedules.

The rest of this section concentrates on the way the roadmap evolved into its current form.

3.1 Early thoughts

Informed by the previous outputs from the project and the literature review the core work package 5 team first concentrated on mapping out the parameters for the road map. Figure 1 through to Figure 5 show an early mind map visualisation of the key features of these parameters.

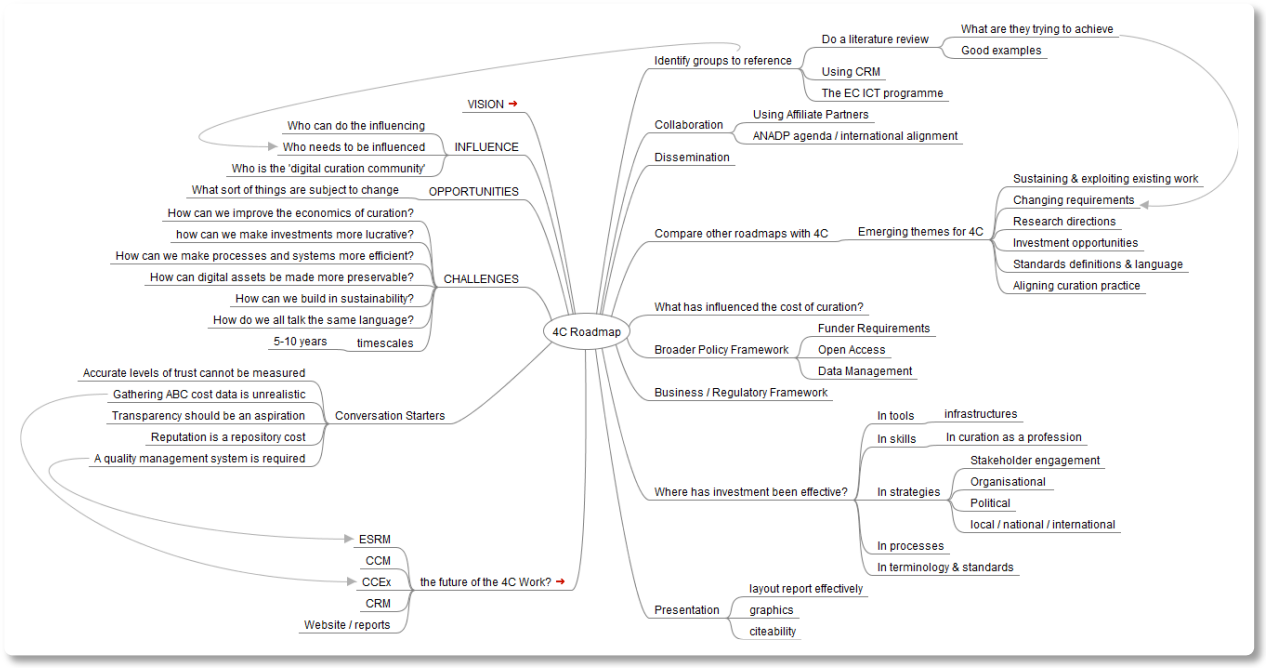


Figure 1—Early roadmap mind map—overall concept

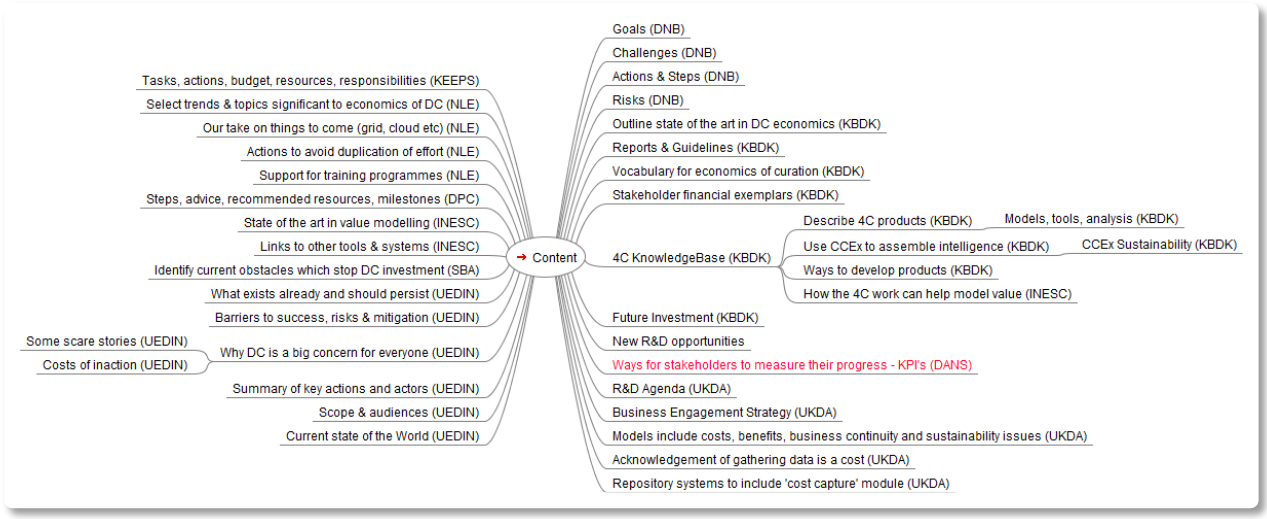


Figure 2—Early roadmap mind map—content

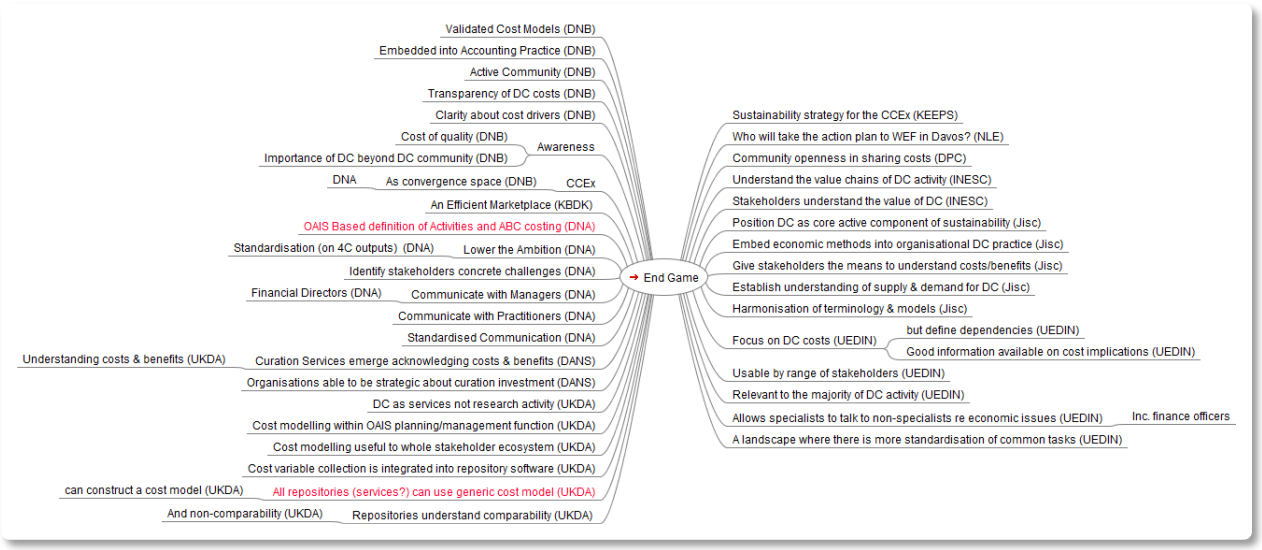


Figure 3—Early roadmap mind map—The end game

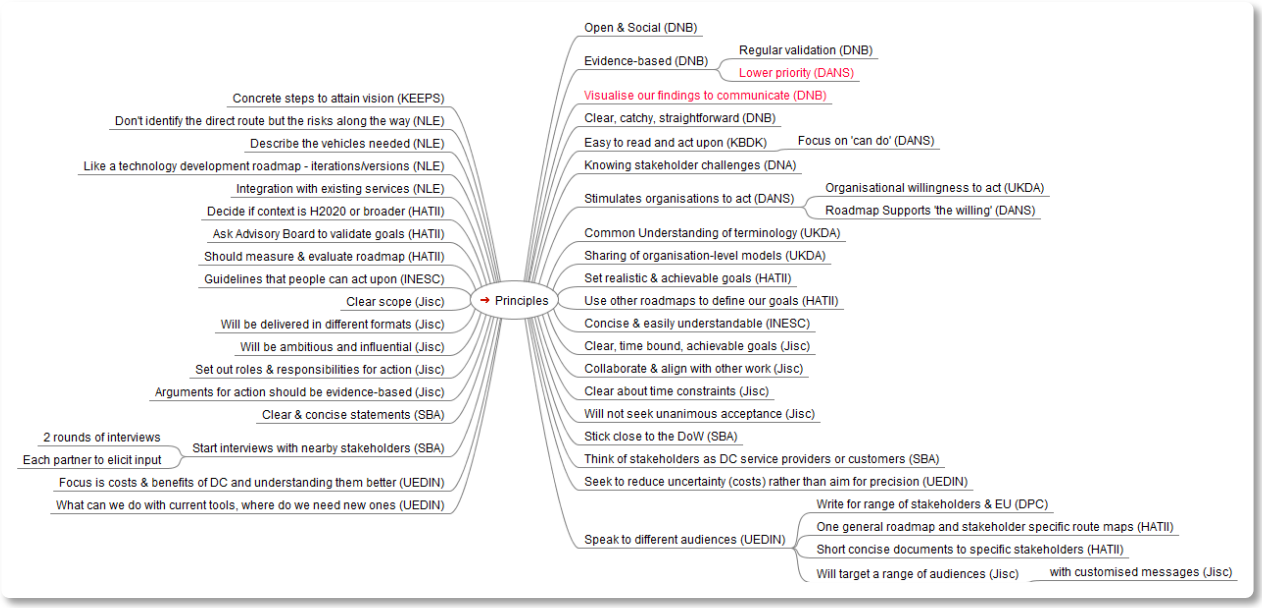


Figure 4—Early roadmap mind map—The principles underpinning the roadmap

Figure 5—Early roadmap mind map—document structure

3.2 Why we went for the format we did

Very early in the process we identified the key audience for the roadmap, namely decision makers and managers in the previously identified stakeholder groups. One of the defining characteristics of such people is that they rarely have time to read long and detailed documents about the "the way ahead". Indeed, the higher up the tree (in management terms) you go the less time there is likely to be for such activities. However, we wanted these key people to read our messages and act upon them. With this in mind we decided to take an executive summary approach. We condensed the key parameters of the road map down into a series of messages, (6 in all at the end of the process) and produced a short, punchy document with a series of clearly defined actions. Obviously the downside of such an approach is that those who do want to read the detail will find it lacking. We felt that the need to get the roadmap read by

decision makers outweighed the need for comprehensive detail. We also provided links to further information that could provide the specific detail if the reader wanted it.

After the initial draft we even went a stage further and produced a series of action postcards (see Appendix B, Postcards) each with just the messages and actions for a single stakeholder group.

3.3 How we honed it down

Identifying the vision was relatively easy compared to the later identification of the messages. Initially all project partners used an idea template to come up with the key parameters for the roadmap individually.

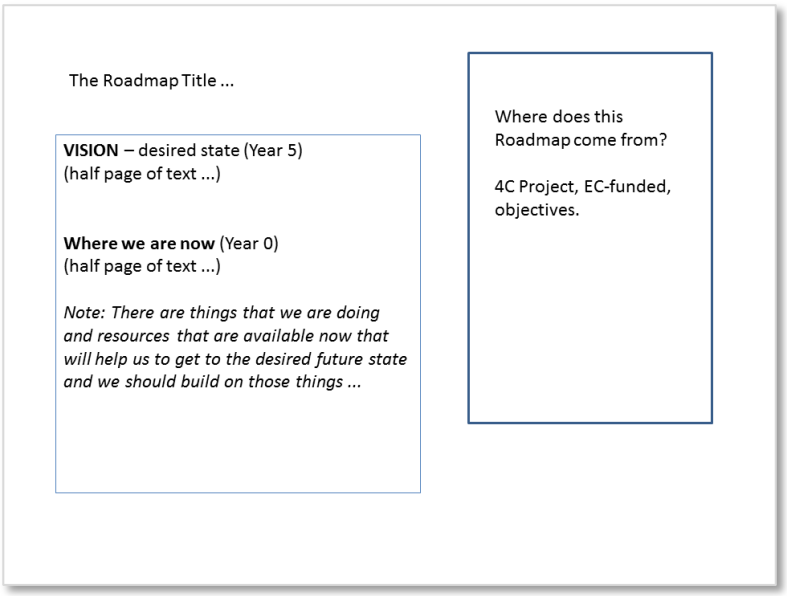
What do we want the Roadmap to Achieve?
Title: (snappy and memorable title for the document: e.g. 'Riding the Wave', 'Sustainable Economics for a Digital Planet')
End Game: (what is the desired outcome that the roadmap – if acted upon - will deliver)
Vision: (What will the roadmap be?)
Principles: (According to what principles and with what objectives in mind should we create the Roadmap)
Content: (What should/could be included in the narrative?)
Time period: (What period should the roadmap cover?)

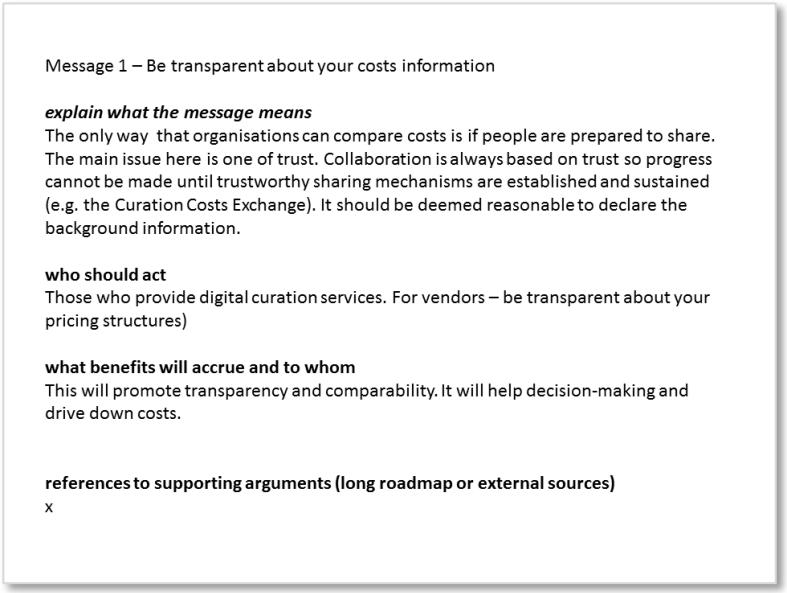
Table 1—Roadmap idea template

These separate ideas were combined to produce a series of mind maps (as shown earlier). These mind maps were used to place the ideas in front of various audiences and further refined.

There then followed the series of structured interviews (also mentioned earlier) and webinars with key (external to the project) stakeholder representatives. We also discussed the roadmap in weekly task group meeting, at every project face to face meeting and at every event attended by 4C where we had a presentation slot.

At this stage we were still contemplating a "traditional" roadmap format although alternative formats were already being discussed. After the June 2014 Steering group and Project meeting in Edinburgh we put aside an extra day dedicated in the most part to finalising the roadmap structure. By the end of that meeting we had the structure and messages organised as shown below.





Message 2 – Make smarter investments: through economies of scale

explain what the message means
It's best to assume that we're unlikely to have budgets raised for curation in line with enormous growth in volumes of content, so investment needs to be strategically / tactically targeted to the right places to create economies of scale and scope. (This will require decisions around appropriate infrastructure). It will also require a high level of commitment to collaboration and a realisation that if you want to remain in control, you need to adopt new thinking. People should be working smarter every year and this feeds into a wider issue around maturing strategy and practice right across the digital curation domain.

who should act

what benefits will accrue and to whom

references to supporting arguments (long roadmap or external sources)

Message 3 – Make Smarter Investments: demand and choose more efficient systems

explain what the message means
Efficient systems. In-house / outsource decisions, competitive tendering. Standardise on established methods and align practice. Increasing awareness of vendors and IT systems providers of curation needs. Make cold hard (80/20) decisions about what you want to achieve and design clear policies.

who should act
x

what benefits will accrue and to whom
x

references to supporting arguments (long roadmap or external sources)
x

Message 4 – Make Smarter Investments: Identify the value of digital assets and do selective digital curation

explain what the message means
Selection and appraisal of assets is needed because curating growing volumes of data is not sustainable. We can point to a handful of good examples where researchers (historians?) can inform selection decisions. This message will suit archivists, but we can't rely on old methods. The analysis has to be automated to be scalable. This a big topic for research. Given a clear policy, it should be possible to automate these tasks. If we try and save everything, we will start to randomly lose things. In response to people who worry that automation is a clumsy tool and things of value will get lost, we should argue that things of value have been getting lost for centuries by human decision-making. The crucial point here is that we longer have a choice because humans are unable to appraise gigabytes (let alone exabytes) of data.

who should act
X
what benefits will accrue and to whom
X
references to supporting arguments (long roadmap or external sources)
x

Message 5 – Funding should be dependent on the declaration of the whole lifecycle costs of sustaining digital assets

explain what the message means
This starts with funders not awarding grants for projects that can't demonstrate they know how much it will cost to sustain and make available the data they will be funded to create, either for specified amounts of time or in perpetuity. But it also needs to include non-project activity and funding transactions that are wider than Research Council-type grant awarding schemes. This message can be nuanced by the lifecycle of some data having a predicted life (7 years for legal data?). Some curation actions may not need to happen and it can be planned that little or no curation happens on data if this is the case.

who should act
X
what benefits will accrue and to whom
X
references to supporting arguments (long roadmap or external sources)
This would pick up on ADS 30 year retention assumptions. If you can keep it for 30 years, you can keep it forever.

Message 6 – Digital curation must be a sustainable structural service

explain what the message means
An example might a global and effective file format registry and this should be available in 5 years. Don't plan for one-off implementations or non-interoperable functions. Always assume scale and interoperability will be necessary. Don't think local, be joined-up. We believe in x and y will be ready and available in 5 years time. Give some concrete examples otherwise this is an empty statement.

who should act
x

what benefits will accrue and to whom
x

references to supporting arguments (long roadmap or external sources)
x

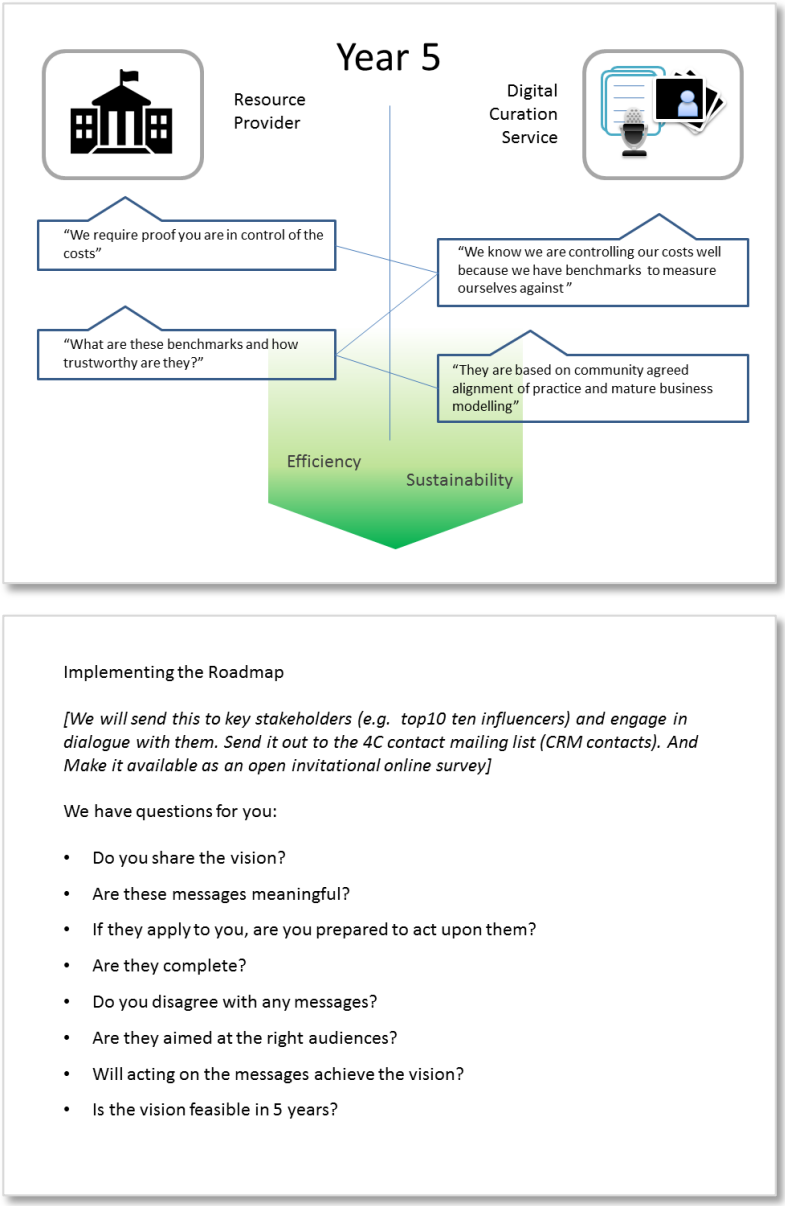


Figure 6—Roadmap structure, messages and action plan

3.4 The vision and messages

The vision and messages were condensed down into the following

- **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.
- **Message 1**—Identify the value of digital assets and make choices
- **Message 2**—Demand and choose more efficient systems
- **Message 3**—Develop scalable services and infrastructure
- **Message 4**—Design digital curation as a sustainable service
- **Message 5**—Make funding dependent on costing digital assets across the whole lifecycle
- **Message 6**—Be collaborative and transparent to drive down costs

3.4.1 The message behind the messages

"Identify the value of digital assets and make choices"

The concept of value has been an underlying theme throughout the project. The key point here is that you can't preserve everything, something that is becoming clearer to everyone as the volume of data being created rises exponentially. You have to choose what to keep and what to let go and that choice should be based upon some concept of value[10]. With that in mind it is essential that content owners have clear policies about their collections encompassing the scope, the type of assets, formats, etc. They also need criteria as to what they consider to have value.

Establishing value requires appraisal effort and this cost should be taken into account when costing curation. It is also important for organisations to consider some forms of (semi-)automatic appraisal to keep costs down.

"Demand and choose more efficient systems"

The digital curation market is immature and it can be difficult to source solutions appropriate to organisational requirements. This in part is due to difficulties in articulating that requirement. Adherence to widely accepted standards, both on the demand and supply side, will lead to better mutual understanding and encourage competitive tendering processes. In addition, it will help with the comparability of products.

"Develop scalable services and infrastructure"

The digital curation arena has a multitude of stakeholders, often with widely differing requirements. Services and infrastructure need to be appropriate for those requirements. It is not always necessary to develop and/or procure systems solely for in-house use. In many cases collaboration and /or outsourcing may provide a more efficient solution. Any solutions implemented need to be scalable as requirements change and additional budget needs to be found to ensure that suitable sustainability and re-evaluation planning is undertaken.

"Design digital curation as a sustainable service"

Effective digital curation requires active management. Active implies effort (and hence cost). Therefore it will always be necessary to fund curation and, given the inevitable increase is data stored, the funding will need to be reviewed regularly.

As ad hoc, reactive curation arrangements shift towards planned activities digital curation becomes more embedded in the normal activities of an organisation. It becomes a service for the organisation. As with any other business service, it is necessary to understand the business case for it. If there is a balance between the demand side and supply side the service is sustainable.

[10] It should be noted that value in tis context doesn't always relate to monetary value.

"Make funding dependent on costing digital assets across their whole lifecycle"

Digital curation requires a flow of resources to support it. Those resources are provided by funders (in the broadest sense of the word). Before funders supply those resources they need a plausible estimate of what the digital curation will cost over their lifetime (which obviously can be very different in different sectors). Digital curators need to be able to provide those costs (the implication being that those who can't won't be funded).

Note that in this context lifecycle may equate to funding cycle[11].

"Be collaborative and transparent to drive down costs."

Generally speaking stakeholders with assets to curate are looking for a "return" on their investment. The ability to compare your operation with that of a peer organisation provides you with the opportunity to identify potential areas of inefficiency and reduce your own costs.

3.5 The final draft version

The final version of the draft can be seen in Appendix B and can be downloaded from the 4C website[12].

It starts with a single page introduction followed by a description of the stakeholder groups.

- **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.

[11] See the discussion of the Digital curation Sustainability Model (DCSM) in the deliverable D4.2 - Assessment of Community Validation of the ESRM [http://4cproject.eu/d4-2-esrm-2]

[12] http://4cproject.eu/d5-1-draft-roadmap

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.

These groups are used in the action tables that follow.

Each message is presented full page followed by a single page of narrative about the message entitled "What this means and who should act".

The narrative section is in turn is followed by a section about the "Benefits and positive outcomes" and an Actions table (again, for each message). The target groups on the actions table are the same as those identified at the beginning of the roadmap.

The final pages contain:

- an envisioning of the roadmap outcomes in the form of a hypothetical conversation between a resource provider and a digital curation service
- A list of the 4C resource underpinning the concepts in the roadmap
- A request for feedback (with directions on how to provide it)

3.5.1 Postcards

As mentioned earlier we also provided an alternative format in the form of postcards directed at individual stakeholder groups—one card for each group—in five languages. See Appendix B, Postcards on page 98.

This condensed format consisted of the roadmap cover and vision on one side with an Actions table on the reverse showing the actions for that stakeholder in relation to the 6 messages.

4 Synthesis of responses

As outlined earlier, the members of the project took every opportunity to put the Roadmap in front of stakeholders and seek their feedback, both during the planning stages and after publication of the draft. As many of these opportunities were in the form of conference presentations and feedback was verbal it is difficult to come up with an accurate figure for the number of people who saw and commented upon the presentation. However, based upon the number of page views, downloads, questionnaire responses and attendance figures at events we can safely say that at least 250 engaged stakeholders participated in the development of the document. Based on this figure and given that all stakeholder groups previously identified by the projects were represented in the responses it is reasonable to claim that the community has validated the final document.

4.1 Response framework

In most cases, responses were requested (and received) using a standard framework. This frame work—structured around the vision, the messages and a section designed to capture "anything else"—was also used for the on-line questionnaire[13] (which has been included as Appendix B).

The frame work presented similar questions relating to each of the 6 messages, namely:

- Is this message meaningful to you?
  If a respondent agreed that any particular message was meaningful in their case they were asked a series of supplementary questions:
  - If this message applies to you, are you prepared to act on it?
  - Do you agree with this message?
  - Is this message aimed at the right audiences?
  - What are your reasons for the choices you made above?

Users were also asked specific questions relating to the vision and the roadmap.

The use of the framework allowed us to group, compare and contrast the responses from all sources. At the same time it also provided a way of capturing diverse responses. Respondents could "go off on a tangent" if they felt the need to do so.

[13] http://www.4cproject.eu/rmfeedback

4.2 Response to explicit questions

4.2.1 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 curation planning and management activities of all digital repositories."*

Do you share the Roadmap's vision?

Approximately 85% of respondents stated that they shared the vision. Reasons for doing so included:

> *"There are strong economic imperatives for this development."*
>
> *"…it is critical that the costs, benefits and the business cases are established and communicated."*
>
> *"I share this vision because I think it necessary in order to secure long-term funding for digital curation as opposed to 'just' project funding…"*
>
> *"1. Digital curation is not a matter of choice as the scale of data collection is constantly increasing. 2. Failure to manage costs will result in lost opportunities and perhaps more importantly if costs aren't strategically managed lost knowledge."*
>
> *"I completely agree"*

Given the strength of support it is perhaps more interesting to home in on the dissenting 15%. Closer examination of the responses shows that the main reasons for disagreement focused more on the achievability of the vision as opposed to the end goals. Over ambitious time scales and difficulties in convincing decision-makers and data producers of the need for digital curation were common themes.

> *"I share the vision but I am not quite sure that it is realistic. In my experience it is difficult to convince the record creators of the value of curation."*
>
> *"I think the timeframe is highly unrealistic given the amount of work ahead for the entire preservation community in order to achieve a better understanding of costs, benefits etc."*
>
> *"I share the vision, but I don't see it being universal practice ("all digital repositories") within 5 years"*
>
> *"Might require more time"*
>
> *"Things move very slowly here"*
>
> *"…progress is frustratingly slow"*

> *"Many of us are already aware of and working on these issues in isolation - strategic sector decision making could capitalise on this."*

> *"I believe that we will need at least 10 years to reach the vision. Issues include many different types of materials/formats, more and more complex formats and larger amounts of data. And funding is currently very limited in institutions all over Europe"*

Most of these responses originated from practitioners. Having said that, the vendor stakeholder group disagreed with the time scale for entirely the opposite reason. They felt it was an *under* ambitious timescale (shrinking it by at least two years was suggested), mainly because much of what was proposed was even now "in the works"[14]. Some even went so far as to suggest that within the timescale proposed the "business as usual" aspect of digital curation would mean that their current business models would need to be changed radically in order to avoid going out of business.

In purely quantitative terms, half the respondents thought the vision was feasible in 5 years and half didn't.

A final thought on the vision relates to the "service" concept.

> *"It may be hard to explain to colleagues that implementing more cost effective and efficient workflows is the same as implementing a 'service'."*

4.2.2 Message 1

"Identify the value of digital assets and make choices"

Almost all respondents found this message meaningful, agreed with it and were prepared to act upon it.

> *"As a curation practitioner I recognise the necessity in determining the value of digital objects placed within our care… …As such appraisal is an important aspect of the work we do and needs to be funded."*

> *"Selection is a prerequisite for preservation."*

> *"It is not possible to retain all data but equally it is difficult to determine now what might be important in the future. By considering this problem early we may come up with innovative solutions to the problem."*

However a small percentage felt it was aimed at the wrong audience.

> *"…the message must be sent to and received by the very highest levels, the rest of us can only tinker with the detail"*

Drilling down into the more detailed responses shows a slightly different story. A significant number of respondents, without disagreeing with the message, we're concerned with *how* the value of the assets might be ascertained. The (possible) root causes for these concerns were expressed as:

[14] It is interesting to note that these comments were made very close to the end of the project—at the 4C conference in November 2014 and at the vendor focus group in January 2015. Discussions nearly 2 years earlier with vendors in the opening stages of the project did not produce similar comments. This is perhaps a reflection of the rapid pace of change in the digital curation domain with vendors typically having quarterly product release cycles.

a) a lack of skills,
b) a lack of resources
c) lack of foresight (curators, especially in the cultural arena, have great difficulty in knowing precisely the future value of the assets in their care)

> *"[in the context of Research Data Management (RDM)] …researchers are not competent or experienced in planning activities on this item"*

> *"'Appraisal has to be (at least) semi-automated to be scalable and 'value' is an essential concept that will need to be algorithmically defined' – for email, that is going to be one complex algorithm"*

> *"…the value of digital objects are not always easy to assess in a straight forward way as cultural heritage often doesn't show its value until future users and their needs are known"*

On the other hand, most remained optimistic and were of the opinion that these were not insurmountable problems.

> *"By considering this problem early we may come up with innovative solutions to the problem."*

4.2.3 Message 2

"Demand and choose more efficient systems"

As with message 1, most respondents felt that this message was meaningful and, almost without exception, were prepared to act upon it.

> *"…need for plain language, simple templates or tools, and resources comparing options is critical to help people understand and implement complex standard."*

> *"If efficiencies are not achieved, digital preservation/curation is not sustainable."*

It's interesting to note that some appeared to see this message as an attack on proprietary software.

> *"…proprietary software dictating what we do in the future and in tangential areas is a problem".*

The context in which the point was raised—as part of the discussion about efficient systems—implies that they believe that proprietary systems are inherently inefficient and open source software is a better way forward. If that is the perception then commercial (proprietary) solution providers need to communicate more effectively to some stakeholders about the underlying technologies in their systems and provide evidence that refutes the assertion of inefficiency.

The point regarding communication and standards, and especially the bi-directional nature of any such communication was an underlying theme here (and in other places in the roadmap responses). Many stakeholders (from researchers through to solution providers) were of the opinion that all stakeholders needed to adopt the same standards and work together to incorporate those standards in systems and workflows.

> *"…standardisation is essential for vendors to understand the requirements of potential customers. Equally, those customers need to be on board with adopting the standards and enshrining that in policy"*

In contrast, some found the dependence on standards to be a step too far. They felt it would impose too great a burden on their already stretched resources.

> *"I find it too generic to say that everything should be built on ISO-standards and I doubt that we as members of the digital preservation community have the resources needed to not only specify our own requirements but also help create 'a common understanding and clear specifications'"*

Another felt that we as a community need to get our act together first before we would be in a position to 'demand more efficient systems'.

> *"Institutions involved in digital curation need to standardise and collaborate a lot more if we are to push the demand for more efficient systems"*

Finally, the legacy system long tail problem was raised. Some practitioners are constrained from adopting the latest systems and standards by the historic legacy systems they still have to work with.

4.2.4 Message 3

"Develop scalable services and infrastructure"

All respondents (who answered the direct questions relating to this message) agreed that this message was meaningful to them, agreed with the message and were prepared to act upon it. The only quantifiable dissent related to the audience that the message was aimed at. A small minority felt it was aimed at the wrong audience.

Endorsements included the following:

> *"…collaboration is important… …Very important to have strong leadership, buy-in, tools that will help less capable organizations reach the levels of those with more resources, succession plans, and strong oversight. Metadata is always a mess."*
>
> *"The actions are also nicely practical!"*
>
> *"…another excellent and meaningful message that I totally agree with. This must be the answer to the financial barriers that we face."*
>
> *"Outsourcing or sharing infrastructure is key to alleviating these costs while delivering the digital curation service expected by all stakeholders."*
>
> *"Scalability is necessary for digital preservation/curation activities to extend beyond large, well funded institutions."*

The latter endorsement is particularly interesting. Many practitioners tend to focus on scaling *up* their systems in order to cope with the ever increasing data deluge. It's as well to remember that the ability to scale *down* activity is also a valid and desirable property for a system, particularly in cases where new forms of storage technology can make previous procurement of 'capital' pieces of equipment redundant.

Given that developing scalable services and infrastructure seemed to be such an obvious thing to do, it's not surprising that some respondents went on to provide explanations as to why it's not already been done on a wide scale. The reasons essentially boiled down to trust and pride. Stakeholders are still relatively unwilling to trust other people and systems with their assets. They also feel that they should be able to cope on their own without seeking help. These are problems that are not unique to the curation community, but they are exacerbated by the irreplaceable nature of the assets in question. The adoption of standards and certification should go a long way towards dispelling these fears.

There were some who didn't wholeheartedly agree with the message.

> *"Collaboration is undoubtedly key to sustaining our digital preservation actions, but it's not exclusively about shared infrastructure"*

The point about infrastructure being only one facet of collaboration is one that the authors of the roadmap would agree with. It is in fact addressed in several places in the roadmap. Stakeholders are encouraged to collaborate at many different levels about many different aspects of digital curation.

Some put their dissent in stronger terms.

> *'I doubt the validity of many of the claims in the text, like "The switch to collaboration, sharing information and sharing resources to manage budgets for digital curation may be easily justified in financial terms."'*

It's difficult to counter statements of this type in the text of the roadmap. We simply don't have room. Nor do we have room to incorporate a reference section. However, we do provide links to the 4C outputs which *do* have appropriate counters and references. The updated version of the roadmap encourages users to consult these.

4.2.5 Message 4

"Design digital curation as a sustainable service"

This message didn't receive quite the overwhelming support as the previous one. However, the main reason for this was that some respondents didn't find it meaningful. Those that did find it meaningful were fully behind the concept.

> *"Yes this is absolutely right"*

> *"This is a natural consequence of a value proposition"*

The "yes, but…" points once again get to the heart of the matter. Firstly, although agreement was widespread many saw difficulties in implementation, particularly the need to gain acceptance for the idea from both the "powers that be" (by implication the funders) and practitioners on the ground.

> *"This will require a very big change in mindset for practitioners"*

> *"organizations store things and forget them or don't want to deal with them"*

Implicit in many responses was the assumption that the service would need to be in-house.

> *"we cannot change that overnight"*

Obviously this need not necessarily be the case. This mind set also appears to tie in with the 'Trust and Pride' issue mentioned earlier. The default assumption of many is that, if there is to be a service (whether it's sustainable or not), it's going to be an in-house service. There is perhaps a need to foster a new mind set for stakeholders across the board so the default becomes "We need a service that is sustainable and a best fit for our needs be it in-house or out-sourced".

4.2.6 Message 5

"Make funding dependent on costing digital assets across their whole lifecycle"

This message had one of the strongest approval ratings in that everyone who found it meaningful agreed with it and was prepared to act upon it.

> *"Critical"*
>
> *"Looking after research data is becoming increasingly important"*
>
> *"It is critical when beginning or planning a curation activity that the full costs are understood and committed to in order to ensure both success and sustainability"*

However, fewer people found it meaningful in the first place. It's difficult to speculate as to why because reasons for not finding messages meaningful were not explicitly captured within the frame work. On the other hand, analysis of verbal feedback from face to face discussions found that some didn't understand the concept. They weren't clear what the lifecycle was in this context, nor what costing across a lifecycle meant. This is perhaps an issue of standards and frameworks again. If all stakeholders subscribed to the same standards, spoke the same language so to speak, then the lifecycle and costing ideas might be clearer. A useful educational tool in this context would be the Digital Curation Sustainability Model (DCSM)[15]. It provides a systematic, standards based way of considering and discussing sustainability issues with senior managers and funders/investors as well as illustrating quite nicely the cyclical nature of the digital preservation lifecycle.

On the negative side, some respondents doubted the practicality of the exercise. Firstly there were doubts about the ability of practitioners to envisage the whole lifecycle in the first place, particularly when it comes to heritage preservation.

> *"Which means you have to predict the whole life cycle"*

They simply don't know what they're preserving could be worth to future researchers.

Secondly, there were doubts about the willingness and/or ability of stakeholders to share information. Fortunately the 4C project's Curation Cost Exchange (CCEx)[16] now provides a platform that will facilitate just such an exchange.

> *"Not sure that institutions will be able to communicate costs to Producers or that Producer will buy into, acknowledge, or care about costs."*

[15] See the appendix in D4.2—Community Validation of the ESRM for details of the DCSM

[16] http://curationexchange.org

The role of producers in the process of and costing of curation turned out to be an underlying theme in responses, one that was perhaps understated in the first draft of the roadmap.

Some stakeholders with specialist use cases, particularly those where the repository is closed and only ever used internally, found it hard to envisage just how the message could be applied to their situation.

> *"I agree in general, but again some areas would find difficult to apply to my organisation, where the repository is funded and used internally, by the business"*

The concept of value comes into play here and stakeholders should find the 4C reports from work package 4 of value here when making the case internally[17].

Lastly, the issue of copyright preventing stakeholders from realising the value of their assets was raised.

> *"exploitation of the value is hugely hindered by things like copyright!"*

Another "elephant in the room" issue that will require a great deal of discussion and perhaps even further legislation before it is resolved satisfactorily. Obviously the issue is of great relevance to the curation community, but it is difficult to see if it could (or even should) be incorporated into the roadmap.

4.2.7 Message 6

"Be collaborative and transparent to drive down costs."

This message generated the most controversy and conversation amongst stakeholders. Unlike the other messages, some respondents were not prepared to act upon it even though they found it meaningful and agreed with it.

On the positive side, many respondents were keen to find out what other organisations were spending and how effective their spending had been.

> *"We're very interested in the digital preservation choices made by other research libraries and how much those choices have cost them"*

> *"This is something that I would find extremely helpful and is really the practical base for achieving funding in the wider context."*

> *"…be not only clear about costs, but also about the related success and failures."*

Others, although keen, thought it would take too much effort and consume too much in the way of resources to make it worthwhile.

> *"I think that this message grossly underestimates the complexity and difficulties of sharing of cost information."*

> *"The message is very reasonable but very time consuming for many organisations."*

[17] D4.1—Indirect economic determinants [http://4cproject.eu/d4-1-ied],
D4.2 —Assessment of Community Validation of the ESRM [http://4cproject.eu/d4-2-esrm-2],
D4.3 —Quality and trustworthiness as economic determinants in digital curation [http://4cproject.eu/d4-3-quality-and-trustworthiness],
D4.4 —Report on Risk, Benefit, Impact and Value[http://4cproject.eu/d4-4-report-on-risk-benefit-impact-and-value],
D4.5—From Costs to Business Models[http://4cproject.eu/d4-5-from-costs-to-business-models]

Fortunately the CCEx platform referred to earlier should make this type of cooperation significantly easier. It is interesting to note that very few respondents raised the issue of commercial confidentiality as a barrier to collaboration. In fact this issue was mostly only raised internally by the project members and externally by solution providers. The latter group indicated that they would be more than willing to share "standard" information (standard scenario costings), but would have difficulty in sharing even anonymised customer costs.

Others thought the case for sharing wasn't yet obvious.

> *"…strong motivators for deploying such openness are difficult to imagine"*

One hopes that the other outputs from the project will help to reinforce the case.

An issue with the one-sidedness of the message, being about driving down costs only, was raised by one respondent.

> *"…collaboration is just as important to enhance value as drive down costs. They are two sides of the same coin"*

A point well made, and addressed in the work package 4 deliverables mentioned above.

Finally,

> *"…no-one is currently demanding that we curate/preserve their digital donations… … They automatically trust us to do what's best but don't actually check that we are doing that."*

An issue to be raised with producers, and funders perhaps.

4.2.8 General themes

Lifecycle

One general theme that emerged, albeit less forcefully than some of those above, stemmed from the fact that some respondents saw in the roadmap an implied message that curation would (and should) always be a process with a definite end. A process where a portion of the assets are not preserved. They saw in it an inference that there would be always be a point at which the life support machine was turned off and the assets were left to expire. Those with a cultural heritage remit in particular took issue with this.

> *"…I think it can be rather difficult to talk about value as an economic determinant when talking specifically about cultural heritage collections, as they are indispensable and there is often opinion expressed that 'everything should be preserved'."*

> *"…I think you overlooked the perspective of national libraries/archives, who… … need to think about the heritage for the longer term."*

> *"...in case of national libraries the time is not a valid issue—archiving means for eternity."*

Where the remit of an organisation requires practitioners to think in terms of preserving objects 'in perpetuity', the point was made by practitioners more than once that such a remit may be diluted (or even slightly threatened) by entertaining too much of a focus on a range of issues to do with cost, benefits and sustainability. The argument being that by allowing the discussion some room to take place, it was a tacit acknowledgement that the mission of the organisation and the resources required to support that mission could reasonably (and unhelpfully) be challenged.

Whilst this might (in some very particular contexts) be a form of denial, the approach of the 4C project is that—generally speaking—it is sensible to pro-actively engage with defining and promoting the value proposition of curated digital assets. Some organisations—particularly those with governmental mandates to preserve (such as national libraries and archives)—may reasonably claim that the resources required to sustain assets, infrastructure and curation activity are unlikely to become unavailable. But even in cases where the ongoing provision of infrastructure is (very nearly) assured and the incentives to sustain assets are enshrined into an organisation's mission, there must still be an acknowledgement that disruptive technologies and/or macro-economic forces can introduce threats that are entirely novel. It is not inconceivable, for instance, that the availability of global storage space will start to become a problem over time.[18]

For organisations such as those described above, the Roadmap's insistence on the need to regard all digital curation activity as a scalable and sustainable ongoing service that can demonstrate benefit to a designated community, should be understood as a reminder to exercise ongoing due diligence. The purpose of the DCSM (referenced earlier) is to provide all organisations with a framework for considering and monitoring their response to all forms of sustainability threat and opportunity.

Data producers

Another theme that emerged related to the role and responsibilities of the data producers.

> *"Producers need to take into consideration digital preservation needs before creation"*

They need to both be aware of the requirements of data preservation systems and the costs associated with meeting those requirements. In short, preservation responsibilities begin at the point of production, not at the point of ingest. Choices made by producers regarding the quality and certification of systems to be used can have a significant effect on the cost.

> *"[data] Providers need to think carefully about the level of certification they (a) need and (b) are willing to pay for."*

Similarly, the quality of the inputs provided to those systems can also radically affect costs.

> *"What are the [curation] staff doing? They're mostly making up for the poor quality of the inputs - mostly inadequate and poor quality metadata"*

There is a need for stakeholders at the very beginning of the supply chain to be better data producers. There is also a need for funders and policy makers to mandate this. It is not enough to mandate deposit. Reasonable metadata should be mandated as well.

> *"…policies that not merely mandated deposit but mandated reasonable metadata (and refused to confirm to funding agencies that data management plans had been followed unless adequate metadata had been supplied) would make a huge difference to total cost."*

Whilst on the subject of policy makers and funders, it was pointed out that policy maker and research funders may not be the same. Whilst this can be the case there is often significant overlap and we chose to include them in the same group of stakeholders.

It was also quite rightly pointed out that,

[18] See" Zettabyting off more than we can chew"—http://4cproject.eu/news-and-comment/4c-blog/157-zettabyting-off-more-than-we-can-chew

"The roles identified in the roadmap will only work if the people occupying those roles are on board."

In other words, we need to convince people at all levels in order to drive the roadmap forward.

Vendor support

Support for solution providers emerged across the board. Obviously purchasers want to achieve the best possible quality within budgetary constraints and that often puts pressure on vendors to drive their prices down. However, some of the purchasers recognised the fact that supporting the vendors to take risks with new products helps to drive innovation. In addition, articulating clearly the value propositions so that others can "come on board" helps to illustrate to the suppliers that there is sufficient demand for a product as illustrated but this quote from a user of curation services.

"There need to be a critical mass of demand before suppliers will invest in development"

The service providers endorsed this point of view as well.

"…we can find it hard to commit to building new solutions and services without clearer demand and support from the customer side."

One of the service providers articulated this as a set of principles which, although too long to include in the roadmap, are worth repeating here.

Purchasers of services should:

- *recognise that suppliers need to take risks when developing new products and services, understand these risks and work with suppliers to minimise and manage these risks.*
- *realise that bespoke solutions are expensive and often not sustainable from a provider perspective whereas if multiple organisations come together and articulate a need for a common and simple solution then this often a lot more attractive.*
- *take small steps with shared commitment between customer and supplier. Share the costs and risks through things like paid pilots, proofs of concepts, testbeds. Do this first rather than specify 'ideal world' systems that might never be cost effective for anyone to build.*
- *work with suppliers to promote successes, share good experience, help them to increase take-up through sales to others. This drives down costs, prices and risks for everyone.*

Completeness

Respondents were asked to comment upon the completeness of the message set. Nearly 90% thought the set was complete.

"Looks like a coherent set"

"They cover the main areas in which action needs to be taken."

"I think these messages are superb…"

But there were some dissenting points of view ranging from partial agreement

> *"…the topic is immature and so new ideas will emerge and the Roadmap should be under constant review and enhanced in the light of new knowledge and ideas."*

through to worries about the potential costs of implementation.

> *"…the roadmap is very ambitious… …a large part of your intended audience will immediately think of - and worry about - the costs of implementing the suggestions…"*

No respondents offered alternative, additional messages to be added to the roadmap. However, given the time and effort it took to come up with the six presented in the draft, this is perhaps not surprising. It would be interesting to ask respondents to revisit the question after some time has elapsed to see if further messages emerge.

General comments

On the whole the response has been very favourable.

> *"I very much welcome this initiative."*
>
> *"…to read the road map was a sheer joy."*
>
> *"I am impressed by how the draft has managed to address so many pertinent issues and also assigned tasks/considerations to so many diverse stakeholders."*
>
> *"…acting on these messages will at least provide a framework of best practice for all stakeholders to engage with"*
>
> *"In particular, the emphasis on defining format preservation policies and keeping them updated as well as the need to partially automate appraisal."*

But there were also some extremely negative reactions.

> *"The Roadmap presents a narrow and depressing view of the digital preservation landscape."*
>
> *"The needs of Curation Practitioners are overlooked in favour of high-level infrastructure investment messages"*
>
> *"I see a lot of forgone conclusions"*
>
> *"I haven't got time to read all of this"*

In addition some respondents thought that it wasn't always clear which audience the messages were intended for.

Some of the issues above arose because of the early decisions taken by the project to aim the roadmap at those most likely to be able to effect change, the decision makers and managers working within the key stakeholder groups. Detail that could clarify the rationale behind the "forgone conclusions" was perforce left out. Similarly, the specific needs of the practitioners were glossed over in favour of general needs for curation stakeholders as a whole.

Finally, it is perhaps worth noting that presenting digital preservation in a "narrow and depressing" way is more likely to lead to some affirmative action than would be the case if the message was only communicated in a relentlessly positive and broadly relevant manner.

4.3 Key points addressed as a result

Taking on-board the comments from respondents we undertook to adjust the final roadmap with additional material addressing the following:

- Solution provider time scales
- Solution provider support and involvement
- Actions for data producers
- The rationale behind the roadmap format
- Further signposting to 4C resources

5 Final version

5.1 Summary of changes

5.1.1 Specific issues

The changes made to the draft for the final version were as follows;

Solution provider time scales

Based upon the feedback we received form all the solution providers we interacted with—that many of the issues raised were already being addressed—the time scales for solution providers have all been shifted to be one year earlier.

Solution provider support and involvement

Extra actions and benefits have been added addressed at a number of stakeholders (including the solution providers themselves). These are designed to improve communication and relationships across the board.

Actions for data producers

Data producers have now been addressed explicitly as part of the Data Users stakeholder group. Further actions aimed at data producers and benefits have been added to the Actions table and benefits bullet points.

The Rationale behind the roadmap format

New text has been added to explain that the short text is intentional and signposting the underlying resources.

Further signposting to 4C resources

As mentioned above, the roadmap needs to be read in conjunction with ALL the 4C deliverables (including this one) to understand the rationale behind the "foregone conclusions" so extra signposting has been added.

5.1.2 General issues

A small number of other changes were made as well. Foremost amongst these were the following

- **Roadmap versus project plan**—text was added to clarify the status of the roadmap in the greater future digital curation vision. It is not a project plan and attempting to impose a project plan style critical path—a critical path that would be different for almost every stakeholder group—would make it impossibly complex and unusable. It is rather a meta plan outlining many parallel pathways to be undertaken by many different groups.
- **Condensed version**—an even shorter version was created to try and broaden the readership.

5.2 Final Roadmap

Introduction

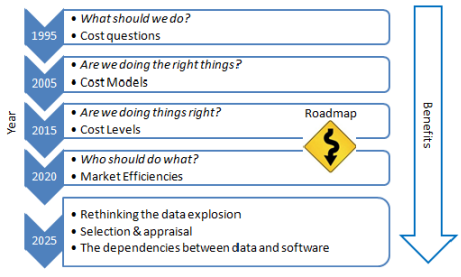
How can organisations working in a variety of different domains more cost-effectively look after and account for the digital assets in their care? This concise Roadmap sets out to address that question by outlining the steps that should be taken over the next five years in order to maximise the efficiency 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 institutions, funded research consortia, research arms of commercial entities (e.g. Microsoft, Google, IBM).

Data Producers/Users (and re-users)

Those who generate the data that will be curated. 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-producing or 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

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.

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.

Yes it is short…

The conciseness of this document is deliberate. We know that our key target readers—those who can make a difference when it comes to changing the face of digital curation—don't have time to read all the background materials and rationale.

Some won't even have time to read this document which is why we have produced other more condensed formats of the roadmap.

Those who *do* want the detail—the research and conclusions that have lead up to this point—will need to look for it in other published outputs from the 4C project, in particular the deliverable D5.2—Roadmap report. The resources in question are listed at the end of this document. They can also be downloaded from the 4C website.

http://4cproject.eu/community-resources/outputs-and-deliverables

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

This 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 Producers/ Users	Content experts to work with technologists to establish value criteria and represent 'designated communities'. Make choices about the level of certification you need and are willing to pay for	●	●			
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

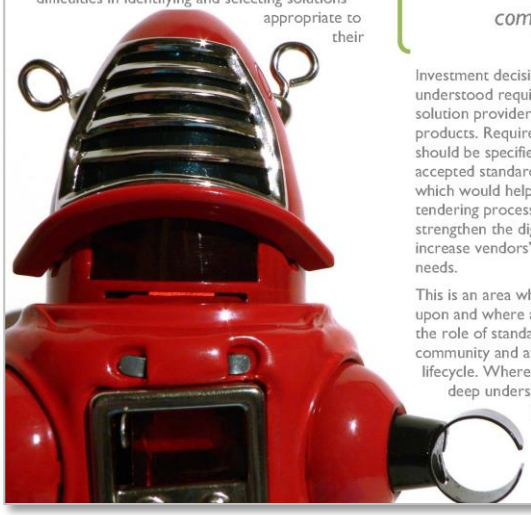
Demand and choose more efficient systems

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 (ISO 15489), 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.
- Utilising good practice make the curation dataflow more efficient and lowers costs
- 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 Producers/ Users	Demand better and more standardised interfaces to data and metadata making data more usable and thus demonstrating its value. Use good practices when working with data at all stages in the curation lifecycle.	●	●			
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 establishing 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 Producers/ 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. Ensure all providers and users of data utilise practices that can reduce cost.	●	●	●	●	●
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

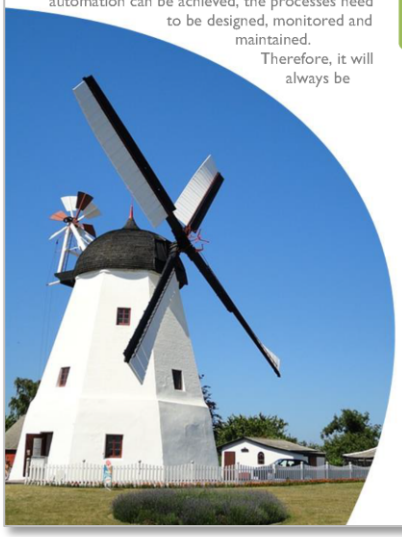
Design digital curation as a sustainable service

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.

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.
- Clear demand and support from data users allows solution providers to commit to building new solutions
- Promoting successes and sharing good experience helps increase take-up of services. This drives down costs, prices and risks for everyone

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 Producers/ Users	Methodically and empirically assert the value of digital assets and work with practitioners, managers and solution provides to undertake cost/benefit analyses and to promote successes		●	●	●	
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	●	●	●		

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 Producers/ 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

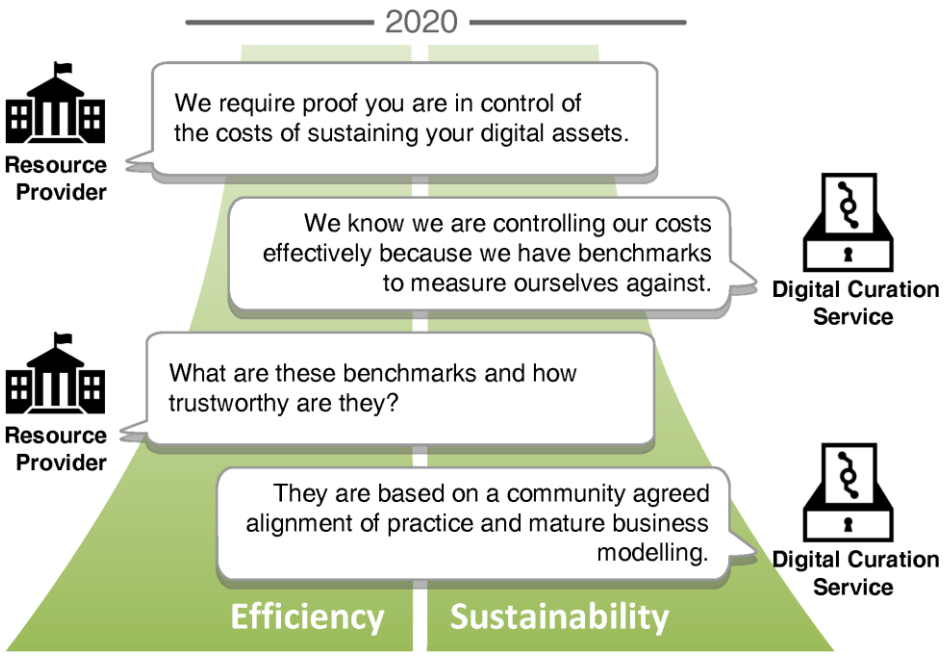
Benefits and positive outcomes

- 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 Producers/ Users	Understand the role and purpose of the 'designated community' for curation and ensure that managers & policy makers include users and solution providers 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 and deliverables of the 4C Project underpin much of what has been put forward in this roadmap. They are also 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

A Summary of Current Cost Models

A summary and description of 10 openly available cost models

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

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 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.

http://4cproject.eu/d4-4-report-on-risk-benefit-impact-and-value

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

An Economic Sustainability Reference Model, & Digital Curation Sustainability Model

Two strategic tools to facilitate discussion and to support planning of successful sustainability strategies for digital curation.

http://4cproject.eu/d4-2-esrm-2

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

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.

http://4cproject.eu/d4-5-from-costs-to-business-models

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

Roadmap report

An overview of the preparation of this roadmap and its validation by the digital preservation community.

http://4cproject.eu/d5-2-roadmap-report

24

So what do you think?

This version of the Roadmap (Feburary 2015) is based upon the draft orignally published in August 2014 and should be considered to be the Final version produced by the 4C project. The messages are based on wide-ranging research and engagement with the community.

From the time we published the initial draft we have sought further feedback from the stakeholder community. This version is the result of that extensive community validation exercise.

Having said that, we still want to know what you think. The digital curation landscape is changing rapidly so it is inevitable that some aspects of this roadmap will become outdated very quickly.

So…

What have we got right?

What have we got wrong?

What have we missed out?

What's changed?

The original on-line survey has now been replaced with a discussion forum at

http://4cproject.eu/roadmap-discussion

Eventually, discussion will be moved to the project's sustainable platform, the curation costs exchange, so please keep an eye on http://curationexchange.org for updates.

It's a Roadmap, not a Project Plan

Given the multiple stakeholder groups and large number of related, but at the same time semi-independent actions suggested in this roadmap, it would be incorrect to suggest that there is a single critical path from 2015 through to 2020. Obviously there is a degree of dependency, but even if we were able to propose such a path it would date very quickly. For this reason we have presented the actions for each stakeholder group as a series of independent parallel time frames.

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/roadmap—where they can download their own copy. Other formats—actions postcards, a condensed version of this document and an on-line interactive version—can also be found on the same web page.

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

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

4C

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 7—Final roadmap

5.3 Final Roadmap—condensed version

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.

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.

Data Producers/Users (and re-users)

Those who generate the data that will be curated. 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.

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.

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.

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.

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.

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

3

Identify the value of digital assets and make choices

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 Producers/ Users	Content experts to work with technologists to establish value criteria and represent 'designated communities'. Make choices about the level of certification you need and are willing to pay for	●	●			
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	●	●	●		

5

Demand and choose more efficient systems

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.
- Utilising good practice make the curation dataflow more efficient and lowers costs
- 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 Producers/ Users	Demand better and more standardised interfaces to data and metadata making data more usable and thus demonstrating its value. Use good practices when working with data at all stages in the curation lifecycle.	●	●			
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	●	●			

7

Develop scalable services and infrastructure

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 Producers/ 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. Ensure all providers and users of data utilise practices that can reduce cost.	●	●	●	●	●
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	●	●	●		

9

Design digital curation as a sustainable service

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.
- Clear demand and support from data users allows solution providers to commit to building new solutions
- Promoting successes and sharing good experience helps increase take-up of services. This drives down costs, prices and risks for everyone

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 Producers/ Users	Methodically and empirically assert the value of digital assets and work with practitioners, managers and solution provides to undertake cost/benefit analyses and to promote successes		●	●	●	
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	●	●	●		

11

Make funding dependent on costing digital assets across the whole lifecycle

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 Producers/ 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	●	●			

13

Be collaborative and transparent to drive down costs

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 Producers/ Users	Understand the role and purpose of the 'designated community' for curation and ensure that managers & policy makers include users and solution providers 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	●				

15

Resources

The Curation Costs Exchange
http://curationexchange.org

A Cost Concept Model and Gateway Specification
http://4cproject.eu/d3-2-ccm

A Summary of Current Cost Models
http://4cproject.eu/summary-of-cost-models

A Report on the Indirect Economic Determinants of Digital Curation
http://4cproject.eu/d4-1-ied

A report on Risk as an Indirect Economic Determinant
http://4cproject.eu/d4-4-report-on-risk-benefit-impact-and-value

An Evaluation of Costs Models and a Needs & Gap Analysis
http://4cproject.eu/d3-1

An Economic Sustainability Reference Model, & Digital Curation Sustainability Model
http://4cproject.eu/d4-2-esrm-2

From Costs to Business Models for Digital Curation
http://4cproject.eu/d4-5-from-costs-to-business-models

A Report on Quality & Trustworthiness as an Indirect Economic Determinant
http://4cproject.eu/d4-3-quality-and-trustworthiness

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

Roadmap report
http://4cproject.eu/d5-2-roadmap-report

Please comment

http://4cproject.eu/roadmap-discussion

Eventually, discussion will be moved to the projects sustainable platform, the curation costs exchange, so please keep an eye on http://curationexchange.org for updates.

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/roadmap—where they can download their own copy. Other formats—actions postcards, a longer version of this document and an on-line interactive version—can also be found there.

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

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
"Decaying Media" by Paul Stokes / CC BY-SA

16

Figure 8—Final roadmap—condensed version

5.4 Updates and activities still to come

Now that the final version of the roadmap has been released we have identified some additional actions to be undertaken. Members of the post project consortium have already undertaken to complete these.

- Post card update—The post cards need to be updates to reflect the changes made to the actions. They will be published on the 4C website when complete
- Web based interactive version—this needs to be updated to reflect the changes made.
- Short slide set—A slide deck is being produced to facilitate post project dissemination and uptake of the roadmap
- Dissemination—The roadmap has been presented in a number of fora since the end of the project (IDCC in March 2015 in London for instance). Consortium members have undertaken to continue these dissemination activities.

6 The way ahead

Although widely welcomed by the community there were some reservations as to the future of the roadmap.

> *"There is a need for a formalized project to achieve the vision."*
>
> *"Who will drive this work forward when 4c ends?"*

We have put in place a number of measures to take the plan forward now that the funded phase has ended. Firstly Jisc, the lead partner in the project, has undertaken to put together a workshop in one year's time to review/revisit the roadmap. The three major topics to be addressed will be:

- Is it still relevant?
- Has anyone been acting upon it?
- Does anything need to be updated?

We are also considering a pre–workshop information gathering survey that will ask key stakeholders where they are in relation to the timetable. The results of the survey and the workshop will be shared with the community.

Discussion forum

The on-line discussion forum[19] is currently located on the 4Cproject.eu website which will no longer be actively managed after the end of project review. The discussion board will be moved to an alternative platform (such as the Curation Costs Exchange) as soon as it is practical to do so.

Advocacy

Further future advocacy to bring on board the key decision maker roles identified in the roadmap is being undertaken by all partners, both internally and within their particular spheres of influence. Individual partners have produced localised dissemination and advocacy materials which are being shared within the 4C group. These along with the roadmap itself will be used to initiate further dialogue with the relevant people and organisations.

Partners have undertaken to feed back to the post project consortium on the results of their dissemination activities. This feedback will include sharing advocacy materials and some detail about who has been approached and outcomes.

[19] http://4cproject.eu/roadmap-discussion

7 Conclusion

The purpose of the Roadmap is to guide the community towards an agreed destination. That destination is a time and a place where it is easy for all organisations to setup and maintain affordable and efficient curation services.

Judging by the community's reaction, we believe that the 4C Roadmap easily satisfies the commitment we made to produce an effective and usable Roadmap. Our purpose from the outset was to provide the community with a resource that provides practical and easily understandable actions that will result in a broadly beneficial outcome. We have made a particular effort to provide this in a format that is easy for designated stakeholders to refer to and use for planning and advocacy purposes.

It is worth reiterating that the clarity of the vision and the generation of the messages was not the result of a single brainstorming session to write something appropriate for a roadmap-type text. We declared from the outset that we intended to be an open and social project and this philosophy was nowhere more evident than with the road map. The first draft to be released came about as a result of all of the project's activities and deliverables and was the consequence of a lot of engagement across all work packages. It encapsulates not only a lot of the expertise of the 4C consortium members but also the knowledge that was shared and developed during the course of the project. We went above and beyond in our efforts to be inclusive, even going to Melbourne (IPRES 14) to put the early draft in front of an audience of experts.

Putting together the 4C roadmap has been a large-scale exercise in collaboration and synthesis and although every effort has been made to keep the Roadmap concise and readable, it inevitably contains a rich amount of detail, particularly where it outlines actions to be taken by designated parties over specified timescales. There are, however, some important high-level principles that provide a foundation for the content and give the document its purpose. It is important that these points do not get lost amidst the detail. Conversely, its very conciseness means that some of the evidence base is external to the document. It is encapsulated in the other 4C activities and deliverables. An in-depth appreciation of roadmap should not be attempted in isolation from the rest of the activities of the 4C project.

The overall reaction to the Roadmap has been very positive, particularly with the written feedback where around 85% of respondents said that they shared the Roadmap's vision and nearly 90% thought that it represented a complete message set. Given the complexity and breadth of the subject area that it addresses—digital curation and sustainability—this is a very significant endorsement of the Roadmap's content.

Despite broad endorsement, there was still a lot of discussion about the messages, how feasible they were within the timescales and about the ownership of the challenges implicit in the messages. We listened to the views of the respondents and put in place a follow-up plan to safeguard the immediate future of the roadmap. We also made appropriate changes to the roadmap to address some other minor concerns raised by the community.

References

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

4C deliverable—D2.3—Final Stakeholder Report
[http://4cproject.eu/d2-3-final-stakeholder-report]

4C deliverable—D2.4—Final Report on Outreach Events
[http://4cproject.eu/d2-4-final-report-on-outreach-events]

4C deliverable—D3.1—Evaluation of Cost Models and Needs & Gaps Analysis
[http://4cproject.eu/d3-1]

4C deliverable—D4.1—Indirect Economic Determinants
[http://4cproject.eu/d4-1-ied]

4C deliverable—D4.2 —Assessment of Community Validation of the ESRM
[http://4cproject.eu/d4-2-esrm-2]

4C deliverable—D4.3 —Quality and trustworthiness as economic determinants in digital curation
[http://4cproject.eu/d4-3-quality-and-trustworthiness]

4C deliverable—D4.4 —Report on Risk, Benefit, Impact and Value
[http://4cproject.eu/d4-4-report-on-risk-benefit-impact-and-value]

4C deliverable—D4.5—From Costs to Business Models
[http://4cproject.eu/d4-5-from-costs-to-business-models]

Appendices

Appendix A Literature Review

WP5 ROADMAP LITERATURE REVIEW

Authors(s)	Jaan Krupp, Raivo Ruusalepp (NLE)
Version	0.3
Date	4/2/14

This survey is based on an initial list of roadmaps identified by the WP5 team. The aim of the survey is to identify roadmaps relevant to the 4C project topics. The list of roadmaps included is not final and the analysis can be extended.

The taxonomy of roadmaps includes the following topics:

- Research roadmaps
- Infrastructure roadmaps
- Digital preservation roadmaps
- Action plans connected to digital curation

Other, non-roadmap texts reviewed are listed at the end of the report.

Research roadmaps

Title	**DigitalPreservationEurope (DPE) Research Roadmap**
URL	http://www.digitalpreservationeurope.eu/publications/reports/dpe_research_roadmap_D72.pdf
Author(s)	Holger Brocks, Seamus Ross, Maurizio Lunghi, Hans Hofman, Colin Rosenthal, Robert Neumayer, Stefan Strathmann, Matthias Hemmje, Jurate Kupriene, Rimvydas Lauzikas, Raivo Ruusalepp, Chiara Cirinnà, Maurizio Messina, Zibute Petrauskiene, Jan Hutar, Emily Witham, Dominic Heutelbeck
Date published	31/10/2007
Period covered	-
Purpose	Despite the frequent discussions concerning the direction of research in digital preservation no concise and well-developed strategy representing a clear consensus has emerged. To address this need the DPE Research Roadmap's objective is to provide a concise overview of the core issues which have to be addressed in future digital preservation research based on an extensive crosswalk of existing preservation research agendas.

Title	DigitalPreservationEurope (DPE) Research Roadmap
Topics/themes	• Digital Object Level—Fundamental issues in digital preservation are raised at digital object level. There has been significant research done on file formats which has produced working solutions for handling the preservation of static digital resources as bit streams, but new types of digital resources emerge regularly and there has been little research on complex and dynamic objects to date. • Collection Level—The collection level is concerned with the longevity of digital collections. While most of the digital preservation research has, thus far, focused on individual types of digital objects, in practice digital repositories are dealing with collections of digital objects and their metadata. Interoperability between different archiving institutions is most often required at the collection level. • Repository Level—The management of digital repositories is based on the understanding of the organisational requirements for long-term preservation of digital material. While digital repositories are being set up, based on OAIS and other standards, research and best practice in managing a digital repository remains a growing area. • Process Level—Preservation processes are established in various contexts. Current preservation practices at most digital repositories still map traditional preservation techniques for physical materials onto digital content. The deluge of digital content that can be expected in the near future necessitates automation of ingest, preservation and delivery processes. • Organisational Environment—The organisational environment is characterised by the creation and use of information, including handling legal aspects such as policies and intellectual property rights. Collaboration between archiving institutions, but also with content creators and distributors are necessary to improve the efficiency and effectiveness of digital preservation. But collaboration might also be interdisciplinary, drawing upon research in related domains. DPE Recommended Research: • Restoration • Conservation • Management • Risk • Significant Properties of Digital Objects • Interoperability • Automation • Context • Storage • Experimentation

Table 2—DigitalPreservationEurope (DPE) Research Roadmap

Title	Blue Ribbon Task Force on Sustainable Digital Preservation and Access
URL	http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf
Author(s)	Francine Berman, Brian Lavoie, Paul Ayris, G. Sayeed Choudhury, Elizabeth Cohen, Paul Courant, Lee Dirks, Amy Friedlander, Vijay Gurbaxani, Anita Jones, Ann Kerr, Clifford Lynch, Daniel Rubinfeld, Chris Rusbridge, Roger Schonfeld, Abby Smith Rumsey, Anne Van Camp
Date published	February 2010
Period covered	-
Purpose	Aims to address the following questions: What digital information should be preserved? Who will preserve it? Who will pay for it?
Topics	• Economic perspective on digital preservation • Economic risks to sustainability • Sustainable preservation in scholarly discourse, research data, commercially owned cultural content and collectively produced web content • Recommendations for achieving sustainability

Table 3—Blue Ribbon Task Force on Sustainable Digital Preservation and Access

Title	Open research challenges and research roadmap for SCAPE
URL	http://www.scape-project.eu/wp-content/uploads/2013/08/SCAPE_D3.1_TUW_V1.0.pdf
Author(s)	Christoph Becker, Andreas Rauber, Norman Paton, Rainer Schmidt, Natasa Milic-Frayling, Brian Matthews
Date published	November 2012
Period covered	-
Purpose	Outlines the research roadmap of the SCAPE project, focused on the scalability of preservation systems in terms of storing and processing as well as decision making and control. It positions the research carried out in SCAPE within the European research landscape focused on digital preservation research. It further outlines the key goals of the R&D work packages in SCAPE, grouped according to sub-projects (preservation components, preservation platform, and preservation planning and watch).
Topics	• Research in digital preservation • Community involvement • Digital preservation challenges

Table 4—Open research challenges and research roadmap for SCAPE

Infrastructure roadmaps

Title	A Strategic Vision for UK e-Infrastructure
URL	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32499/12-517-strategic-vision-for-uk-e-infrastructure.pdf
Author(s)	-
Date published	04/01/2012
Period covered	10 years
Purpose	A roadmap for the development and use of advanced computing, data and networks
Topics/Themes	• Scientific and industrial growth • E-enabled Science and Innovation • The e-infrastructure ecosystem • An emerging business model • The E-infrastructure Leadership Council

Table 5—A Strategic Vision for UK e-Infrastructure

Title	E-Infrastructure Roadmap
URL	http://www.epsrc.ac.uk/SiteCollectionDocuments/ourportfolio/EInfrastructureRoadmap.pdf
Author(s)	Dr Louise Tillman, Dr Susan Morrell, Dr Tracy Hanlon, Dr Daniel Emmerson, Dr Michele Erat, Dr Edward Clark, Mr Timothy Erskine
Date published	21/01/2014
Period covered	-
Purpose	• Understand the whole UK e-infrastructure landscape, view it holistically and consider it within an international context. • Understand the requirements of the EPS research community that make use of e-infrastructure; ensuring there are no gaps or duplication. • Identify where EPSRC, and more specifically the EPSRC Research Infrastructure team can add the most value. • Provide a framework for spending reviews and business cases for funding opportunities from government. • Be used as a discussion tool with other stakeholders and Research Councils.

Title	E-Infrastructure Roadmap
Topics/Themes	• Capability—By Capability, we mean People Development, Computing and Data Skills and Research and Sector Domain Knowledge. These underpin the whole roadmap: having skilled people is both an input to and output of a healthy e-infrastructure ecosystem. Training in both tools (such as programming and software engineering, and basic data analysis) and research methods (applying computational techniques and data analytics as research tools) is required at all career stages. The key challenges are around ensuring that researchers in software/computational techniques get recognition, and have access to sustainable academic career paths. This is a particular problem for research software engineers and research technologists. There are also benefits to be gained from co-ordination and integration of training, such as sharing best practice and resources, and community building. • Connections—As the volume of data being generated through scientific research rises rapidly, and the scale of international collaboration increases it is essential that the research community has access to high speed, high capacity infrastructure that can be shared in an open and secure manner. • Software Development—The importance of software development at all levels of the software stack has been highlighted in a number of high-profile reports. Software is where much intellectual property, knowledge and understanding resides and this is why software has such longevity. Software and algorithm development also represents major investments in skilled scientists and engineers and the large suite of codes used in research therefore needs to be regarded as a research infrastructure in its own right, requiring support and maintenance along the innovation chain, and throughout its lifecycle. • Data Infrastructure—It is clear that data science is becoming more important. There is growing awareness amongst our research community that data is a valuable asset. To stay competitive, the UK needs to invest in the development of cutting-edge skills in data analytics and software development and provide a clear career path for data management professionals. In addition to traditional structured data, research will increasingly use unstructured data coming from non-traditional sources such as crowd sourcing initiatives and social media. We need to ensure that the skills and infrastructure are present in the UK to support these new challenges. • Hardware and Compute—Computing hardware used to carry out modelling, simulation, data analysis and visualisation ranges from desktop machines, to university and regional systems to the national HPC service and access to international machines and new services such as Cloud. Integration across the tiers of the eco-system is still immature, but increased co-ordination and integration of systems and services will allow the UK to maximise the impact of capital investment in this area plus provide users with easy access to the type of e-infrastructure they require.

Table 6—E-Infrastructure Roadmap

Title	Science Data Infrastructure Roadmap
URL	http://www.parse-insight.eu/downloads/PARSE-Insight_D2-2_Roadmap.pdf
Author(s)	PARSE.Insight consortium
Date published	05/06/2010
Period covered	-
Purpose	The purpose of this document is to provide an overview and initial details of a number of specific components, both technical and non-technical, which would be needed to supplement existing and already planned infrastructures for science data.
Topics	• Demand for a science data infrastructure • Requirements for a science data infrastructure • Possible financial infrastructure concepts and components • Possible organisational and social infrastructure concepts and components • Possible policy infrastructure concepts and components • Virtualisation of policies, resources and processes • Technical science data concepts and components

Table 7—Science Data Infrastructure Roadmap

Title	2011 Strategic Roadmap for Australian Research Infrastructure
URL	http://www.innovation.gov.au/science/Documents/2011StrategicRoadmapforAustralianResearchInfrastructure.pdf
Author(s)	-
Date published	September 2011
Period covered	-
Purpose	The 2011 Roadmap articulates the priority research infrastructure areas of a national scale (capability areas) to further develop Australia's research capacity and improve innovation and research outcomes over the next five to ten years. The capability areas have been identified through considered analysis of input provided by stakeholders, in conjunction with specialist advice from Expert Working Groups. The 2011 Roadmap highlights the need to sustain high performing facilities that remain a national priority, to ensure the availability of infrastructure services on which researchers and the sector can rely.
Topics/Themes	• Research outcome targeted capability areas—Marine Environment, Terrestrial Systems, Atmospheric Systems, Solid Earth, Urban Settlements, Sustainable Energy, Integrated Biosecurity, Cyber Security, Astronomy, Population Health Research Platforms, Translating Health Research, Cultures and Communities • Enabling Capability Areas—Integrated Biological Discovery, Biological Collections and Biobanks, Characterisation, Fabrication, Space Science, Digitisation Infrastructure, eResearch Infrastructure

Table 8—2011 Strategic Roadmap for Australian Research Infrastructure

Title	e-IRG Roadmap 2012
URL	http://www.e-irg.eu/images/stories/publ/e-irg_roadmap_2012-final.pdf
Author(s)	Gudmund Høst, Leif Laaksonen, Ivan Maric, Norbert Meyer, Kees Neggers, Dana Petcu, Bjørn Henrichsen, Sverker Holmgren, Frank van Iersel, Dieter Kranzlmüller, Rossend Llurba, Christian Straube, Anton Frank, Marie Sandberg
Date published	05/12/2012
Period covered	-
Purpose	The e-IRG Roadmap 2012 outlines a vision for the future of e-Infrastructures in Europe. To meet the challenges of implementing the EU's 2020 Strategy, this vision outlines Europe's need for a single "e-Infrastructure Commons" for knowledge, innovation and science as a living ecosystem that is open and accessible and continuously adapts to the changing requirements of research. This Roadmap presents the principles of the political, technological, and administrative framework needed for such an e-Infrastructure Commons.
Topics	Roadmap consists of three steps: • Describes the changing world: new social and technological developments, new research paradigms, such as those caused by the data deluge, and the resulting new user requirements; • Presents a vision on the required European e-Infrastructures Commons in 2020; • Addresses the consequences: Reorganise for 2020.

Table 9—e-IRG Roadmap 2012

Title	EGI Technology Roadmap
URL	https://documents.egi.eu/public/RetrieveFile?docid=612&version=18&filename=EGI-D5.4-v17.pdf
Author(s)	Steven Newhouse, Michel Drescher
Date published	07/09/2011
Period covered	-
Purpose	The purpose behind the EGI Technology Roadmap is to establish a multi-year view as to how EGI will evolve from a technology perspective from the previous project-based structures to a sustainable pan-European e-Infrastructure. Sustainability is a critical aspect of this vision, which is dealt with in EGI's Sustainability Plan; however, at the heart of any sustainability strategy there are three key points relevant for the Technology Roadmap: • Clearly defining services that are attractive, unique and needed by their consumers • Sourcing these services from the most effective technology solutions available • Delivering the defined services to a high-quality to the available resources
Topics	• Security • Information • Operations • Storage • Data • Compute • Virtualisation • Instrumentation & Clients

Table 10—EGI Technology Roadmap

Digital preservation roadmaps

Title	Digital Cultural Heritage Roadmap for Preservation – Open Science Infrastructure for DCH in 2020 (draft version)
URL	http://www.digitalmeetsculture.net/repository/dch-rp/public%20deliverables/DCH-RP_D3.4_Intermediate%20version%20of%20the%20Roadmap.pdf
Author(s)	Borje Justrell, Lajos Balint, Eva Toller
Date published	10/01/2014
Period covered	A short-term action plan (2014) is proposed by the DCH-RP project in order to initiate the development of a preservation services infrastructure on a level that will be self-sustainable and continue to progress on its own. This further progress is defined in terms of two further proposed time spans: • Medium-term (2016, i.e. two years after the end of DCH-RP), and • Long-term (2018 and beyond) for the logical continuation of the DCH-RP work.
Purpose	The overall vision for the DCH-RP roadmap is to implement a federated infrastructure, dedicated to support the application of open science in the arts and the humanities, which will make digital cultural heritage accessible and usable long term. The roadmap exercise as such is aiming to produce an instrument that will facilitate policy makers as well as management within cultural heritage institutions.
Topics/themes	• Harmonisation of data storage and preservation: would allow integrating in common environments the curation of research data with other digital objects—two domains which are currently addressed separately; • Improved interoperability: includes better integration of preservation within the overall workflows for digitisation and online access; in a way this is a set of measures to avoid building 'digital silos' within the organisation, for example when digitisation is carried out without taking into account needs for preservation, and/or accessibility online is disjointed from preservation; • Establishment of conditions for cross-sector integration: a key condition for maximising the efficiency of successful solutions, transferring knowledge and know-how; • Governance models for infrastructure integration: a necessary condition for successful institutional participation in larger e-Infrastructure initiatives, and aggregation and re-use of digital resources.

Table 11—Digital Cultural Heritage Roadmap for Preservation – Open Science Infrastructure for DCH in 2020 (draft version)

Title	APARSEN Roadmap
URL	http://www.alliancepermanentaccess.org/index.php/aparsen/aparsen-roadmap/
Author(s)	-
Date published	-
Period covered	-
Purpose	To identify research areas which can be integrated into broad topics. These broad topics are then merged successively into a common vision.
Topics/Themes	Common vision: • Trust— testing environment, authenticity & provenance, annotation, reputation & data quality, peer review and 3rd party certification • Sustainability—preservation services, storage solutions, cost/benefit data & modelling, brokerage, business cases • Usability—Common tools, interoperability & intelligibility, scalability • Access—identifiers & citability, data policies & governance, digital rights

Table 12—APARSEN Roadmap

Action plans related to digital preservation

Title	LERU Roadmap for Research Data
URL	http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf
Author(s)	Pablo Achard, Paul Ayris, Serge Fdida, Stefan Gradmann, Wolfram Horstmann, Ignasi Labastida, Liz Lyon, Katrien Maes, Susan Reilly, Anja Smit
Date published	December 2013
Period covered	-
Purpose	The Roadmap presents a series of blueprints which LERU members, indeed any European university, could use to begin to tackle the challenges which research data poses. It also has a series of messages for researchers, research institutions, support services and policymakers.
Topics/Themes	Roadmap looks at the challenges posed by research data management (RDM) from six viewpoints: • Policy and Leadership • Advocacy • Selection and Collection, Curation, Description, Citation, Legal Issues • Research data Infrastructure • Costs • Roles, Responsibilities and Skills Includes selected case studies and examples from LERU universities. The resulting Roadmap, like its predecessor on Open Access to research publications, presents a series of blueprints which LERU members, indeed any European university, could use to begin to tackle the challenges which research data poses. It also has a series of messages for researchers, research institutions, support services and policymakers.

Table 13—LERU Roadmap for Research Data

Title	Digital Repositories Roadmap: looking forward
URL	http://www.jisc.ac.uk/whatwedo/programmes/reppres/keydocs.aspx http://www.slideshare.net/eduservfoundation/the-repository-roadmap-are-we-heading-in-the-right-direction
Author(s)	Rachel Heery, Andy Powell
Date published	07/04/2006
Period covered	2006-2010
Purpose	This roadmap presents a vision for 2010 in which a high percentage of newly published UK scholarly output is made available on an open access basis and in which there is a growing recognition of the benefits of making research data, learning resources and other academic content freely available for sharing and re-use. Furthermore, geospatial information will be better integrated with other data through improved licensing agreements. This roadmap focuses on UK repositories for research outputs (text, data and other) and learning materials. Administrative records are out of scope. Furthermore, the roadmap is only concerned with objects created, owned and shared by members of the HE/FE community not those made available to HE/FE on a commercial basis. The roadmap will consider repository services associated with management and dissemination of research and learning outputs of UK institutions offered at institutional, national or subject-based disciplinary level. The roadmap will not include 'repositories' that manage and provide access to information about collections and services, ontologies and terminologies, nor analysis tools (often characterised as 'registry services').

Title	Digital Repositories Roadmap: looking forward
Topics/Themes	• Policy—Research councils and other funding bodies need to mandate that all scholarly publications generated by publicly-funded research are made available on an open access basis. The RAE needs to move significantly towards using open access copies of scholarly publications as a primary mechanism to support the assessment exercise. Motivated both by the open access agenda, and by the requirement to manage their digital assets effectively, institutions should build curation of scholarly publications, research data and learning objects into their information strategies. Although the long term preservation of all academic output is an important consideration, the aims and issues in this area need to be clearly articulated separately from (but in relation to) the aims of open access and asset management. • Cultural—The 'reward structures' and 'professional development' infrastructure within the academic community need to recognise open access as a valuable and important part of the profession. The community needs to find ways to encourage academics to share and re-use publications, research data and learning resources as openly as possible. • Technical—The technical infrastructure supporting open access needs to be based on a more thorough modelling of the materials being made available, the way such materials are described and identified and the mechanisms for automatically interlinking and manually citing scholarly output, research data and learning objects. There needs to be widespread agreement about the machine to machine interfaces (the services) that open access repositories should support in order to ingest and make available content and metadata. Finally, repositories should be well integrated into institutional and national access management approaches (such as Shibboleth). These activities will provide a solid environment within which a wide variety of software tools (open source and commercial) and added value services can be developed by both the public and private sectors. • Legal—The licensing of community-developed content needs to protect the intellectual property of institutions, individual academics and third-parties as necessary yet still be supportive of the open access approach. The community needs to find ways to avoid a situation where concerns about IPR are allowed to stifle the creative sharing and re-use of academic content.

Table 14—Digital Repositories Roadmap: looking forward

Title	The Resource Discovery Taskforce (RDTF) Vision
URL	http://discovery.ac.uk/vision/ http://discovery.ac.uk/files/pdf/jisc-rluk-vision-final-june2010.pdf
Author(s)	-
Date published	-
Period covered	2010-2012
Purpose	The Resource Discovery Taskforce (RDTF) Vision is to make resources more discoverable in order to add value for researchers, teachers, and managers of information assets within libraries archives and museums. Aims: • Clearly position and define the benefits of the RDTF Vision to research and education at the local and national level • Improve the discoverability of UK library, archives and museum content • Drive a shift in ethos to 'open' in institutions, services and funding bodies • Improve the quality and sustainability of new and existing resource discovery infrastructure • Be understood, endorsed and promoted by key stakeholders within the library, archives, and museums sector and beyond
Topics/Themes	Targets: • Progress embedding of the technical, licensing and metadata principles • Drive innovation and sustainable, benefits-led reuse of LAM open metadata • Identify and establish core efficiencies in dataflow and aggregation that can be achieved by key shared UK bibliographic data services • Establish open licenses for JISC library and archives service metadata and other key UK LAM aggregations • Develop demonstration exemplars of what is possible, strengthening the business case for open data and also identifying issues for sustainability • Open up and make discoverable important but hidden collections • Demonstrate and support approaches to inaccessible metadata and where no metadata exists • Persuade funding bodies and vendors to support the key principles • Engage with related initiatives to ensure that the approaches recommended in Discovery are compatible with relevant work occurring elsewhere. • Work with related JISC initiatives to explore how they can be integrated into the Discovery framework

Table 15—The Resource Discovery Taskforce (RDTF) Vision

Title	Big Data Roadmap, Association of the British Pharmaceutical Industry (APBI)
URL	http://www.abpi.org.uk/our-work/library/industry/Pages/big-data-road-map.aspx
Author(s)	-
Date published	21/11/2013
Period covered	4 years
Purpose	This road map sets out a four-point plan to direct progress in big data over four years. The aim is to outline the opportunities and challenges that big data presents, give an overview of the UK's big data assets and issues and recommend specific actions required to create the conditions for success.
Topics/Themes	• Increase awareness • Build capability and capacity • Create a sustainable data ecosystem • Accelerate high-value opportunities

Table 16—Big Data Roadmap, Association of the British Pharmaceutical Industry (APBI)

Title	Roadmap for the European Platform on Life Cycle Assessment: facilitating data collection and sustainability assessments for policy and business
URL	http://bookshop.europa.eu/en/roadmap-for-the-european-platform-on-life-cycle-assessment-pbLBNA26379/
Author(s)	Karen Allacker, Fulvio Ardente, Lorenzo Benini, Camillo De Camillis, Simone Fazio, Malgorzata Goralczyk, Lucia Mancini, Rana Pant, Marco Recchioni, Serenella Sala, Erwin M. Schau
Date published	2013
Period covered	-
Purpose	The EPCLA platform plays an increasingly visible and vital role in support of many policy needs, particularly in the context of life cycle data without which many recommendations would fail or be of questionable coherence and quality. The long term perspective is to make the platform the official (and most important) reference frame for LC-based activities in the European context.

Title	Roadmap for the European Platform on Life Cycle Assessment: facilitating data collection and sustainability assessments for policy and business
Topics	• Guidances – ILCD Guidance/Handbook, impact assessment • Data—EUROPEAN REFERENCE LIFE CYCLE DATABASE (ELCD): A reference database is essential for coherence and quality assurance. This is particularly the case for data used in a wide range of life cycle studies, such as for energy, and where these data are likely to have a significant influence on the outcomes of many assessments. Similarly, such a reference data source can be required in a policy support context, including in e.g. sector and product specific rules/applications. The need of new high quality datasets, especially for the so-called "secondary data", was hence pointed out several times. Thus new datasets on e.g. end of life scenarios, energy/transport, primary production, and raw materials may need to be developed and/or provided. • Governance—ADVISORY GROUPS: Formal agreements have been established with 3rd countries, with key data providers from business, tool and database providers such as consultants, as well as with developers of impact assessment methods needed to provide indicators of burdens in life cycle tools. The relations between EC and advisory groups may be further enhanced, by planning at least one meeting per year with all the AGs, discussing the state of the art and upcoming development opportunities of the EPLCA, plus additional specific meeting where needed. Some new advisory groups with e.g. member states or EC-DGs may be created. A new Advisory Group on the Format might also be useful. It would be also useful to create a transversal AG for the general discussion on e.g. the future data provision on the ELCD. • Knowledge sharing—RESOURCE DIRECTORY: A common repository of meta information on e.g. existing studies, tool and database, and other life cycle service providers remains vital. This includes identifying and promoting those that support compatibility with EC policy requirements. Some actions must be taken over a short-mid-term, to populate the resource directory to support policy needs for study results and for storage of EC financed studies. On this perspective should be enforced. • IT tools—FORMAT AND NOMENCLATURE HARMONIZATION: To use life cycle data from different sources, as well as to promote availability and quality assurance, a common format and nomenclature are essential. In a potential agreement with the Platform advisory groups, a common strategy may be adopted to solve the problems related to the remaining differences among ILCD requirements and other existing schemes adopted by database developers. Common nomenclature should be developed and adopted, slightly updating the ILCD data format.

Table 17—Roadmap for the European Platform on Life Cycle Assessment: facilitating data collection and sustainability assessments for policy and business

Other texts reviewed

PLAN FOR DISSEMINATING AND EXPLOITING KNOWLEDGE BEYOND SCIENTIFIC & ACADEMIC COMMUNITIES

http://shaman-ip.eu/sites/default/files/SHAMAN_D17.2_PlanDissExplKnowledge.pdf

COINVENT—Concept Invention Theory

http://www.iiia.csic.es/coinvent/goals

Concept

What: a computationally feasible, cognitively inspired formal model of concept invention grounded in a sound mathematical theory of concepts

Why: for engineering computer systems that effectively support humans in those genuinely creative tasks underlying abstract, serendipitous thinking both in the sciences and the arts

How: by (1) building upon Goguen's proposal of a Unified Concept Theory and drawing from cognitive and social theories of conceptual blending for concept invention, (2) implementing a proof of concept drawing from interdisciplinary research results, and (3) validating the model by deploying the proof of concept in two testbed scenarios: mathematical reasoning and melodic harmonization.

Goals

- to develop a novel, computationally feasible, formal model of conceptual blending that is sufficiently precise for capturing the fundamental insights of Fauconnier and Turner's theory, and is at the same time general enough to address the syntactic and semantic heterogeneity of knowledge representations;
- to gain a deeper understanding of conceptual blending and its role in computational creativity by linking this novel formal model to relevant, cognitively inspired computational models, such as analogical and case-based reasoning, induction, semantic alignment, and coherence-based reasoning;
- to design a generic, creative computational system based on this novel formal model capable of serendipitous invention and manipulation of novel abstract concepts, enhancing thus the creativity of humans when this system is instantiated to particular application domains for which conceptual blending is a core process of creative thinking.
- to validate our model and its computational realization in two representative working domains of creativity: mathematics and music.

CRe-AM: Creativity REsearch Adaptive roadMap

http://www.brunel.ac.uk/sed/ece/research/cmcr/cre-am-creativity-research-adaptive-roadmap

The CRe-AM project aims to bridge communities of creators with communities of technology providers and innovators, in a collective, strategic intelligence/roadmapping effort to streamline, coordinate and amplify collaborative work towards developing, enhancing, and mainstreaming new ICT technologies and tools by addressing the needs of different sectors of the creative industries (e.g. art/culture, crafts, publishing, design, games). Our premise is that ICT use could help make art more widely accessible, more inclusive, and generate significant awareness around it. The project will involve creators who currently use ICT tools in their everyday creative practices, and engage them in a collective dialogue with ICT researchers and developers, with a focus of empowering creators by giving them access to new forms of

facilitation, enhancement, and contextualization of the creative process and its product. The focus will be the future ICT R&D agenda, which will develop new tools for supporting the creative processes as well as enhancing and improving existing tools and platforms to be more adapted to, or to better care for, the needs of specific creators' groups. Thus, the project will also aim at forming a critical mass of ICT and creative communities working together. The main target users will be individual creators/workers and professionals, as well as SMEs, creative groups, communities, and organizations.

Appendix B Draft Roadmap

Roadmap

Full roadmap

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

Introduction

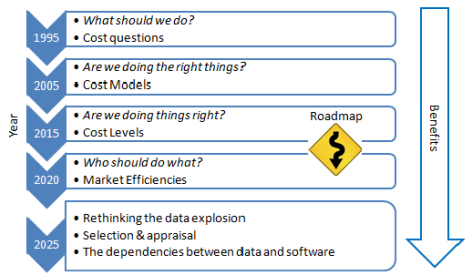
How can organisations working in a variety of different domains more cost-effectively look after and account for the digital assets in their care? This concise Roadmap sets out to address that question by outlining the steps that should be taken over the next five years in order to maximise the efficiency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

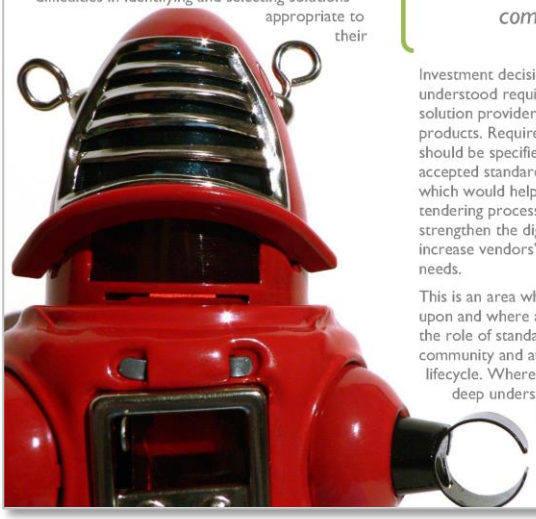
Demand and choose more efficient systems

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

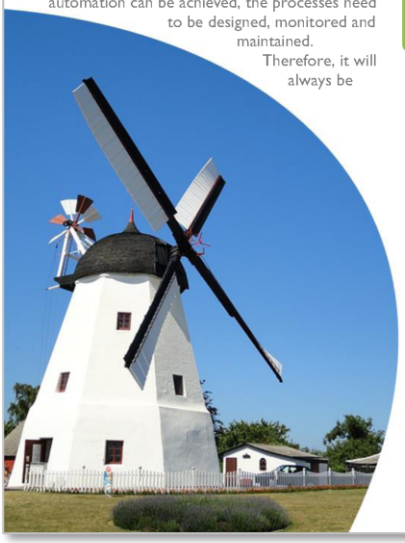
Design digital curation as a sustainable service

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.

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	●	●	●	●	

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

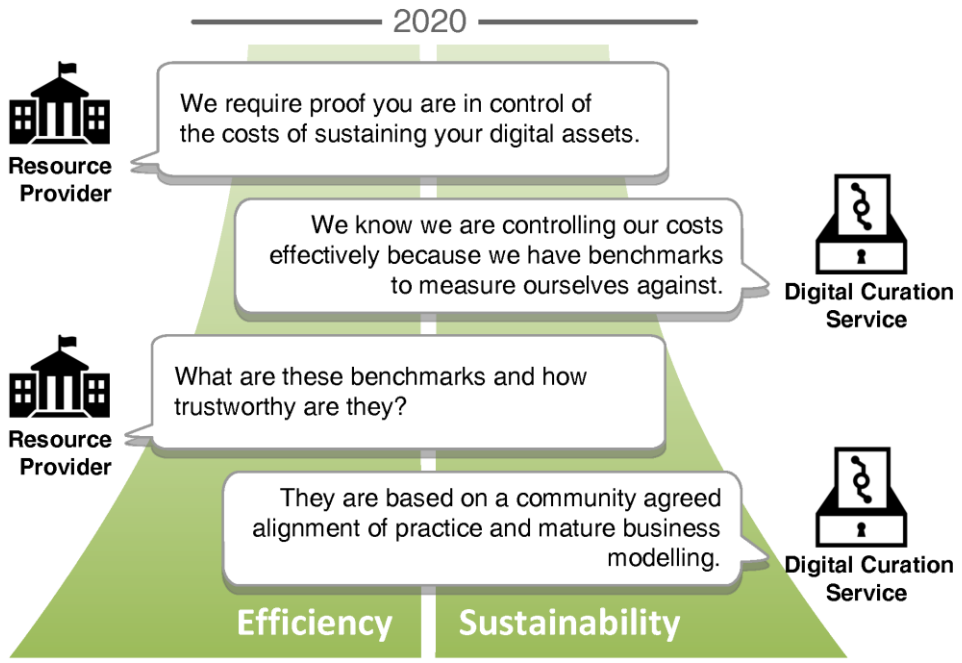
Benefits and positive outcomes

- 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

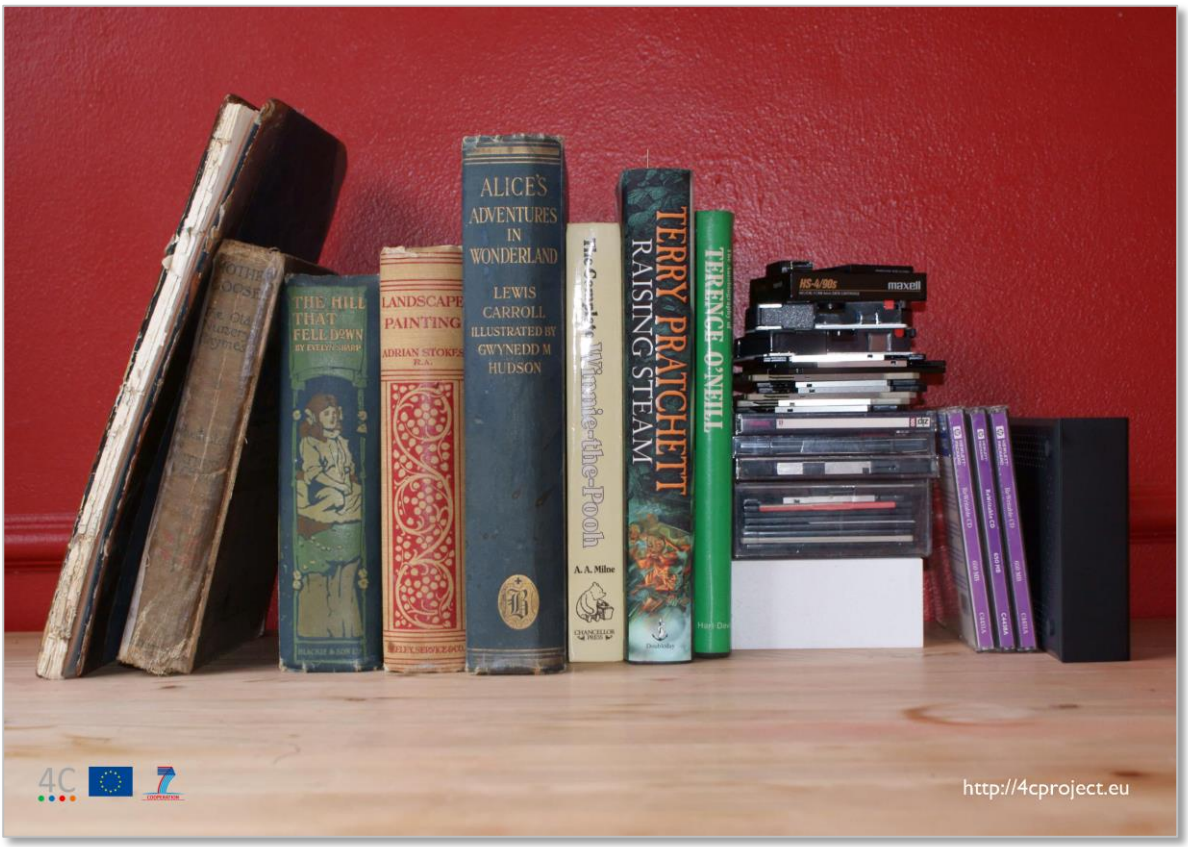


Figure 9—Roadmap booklet

Postcards

Stakeholder actions postcards—English

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

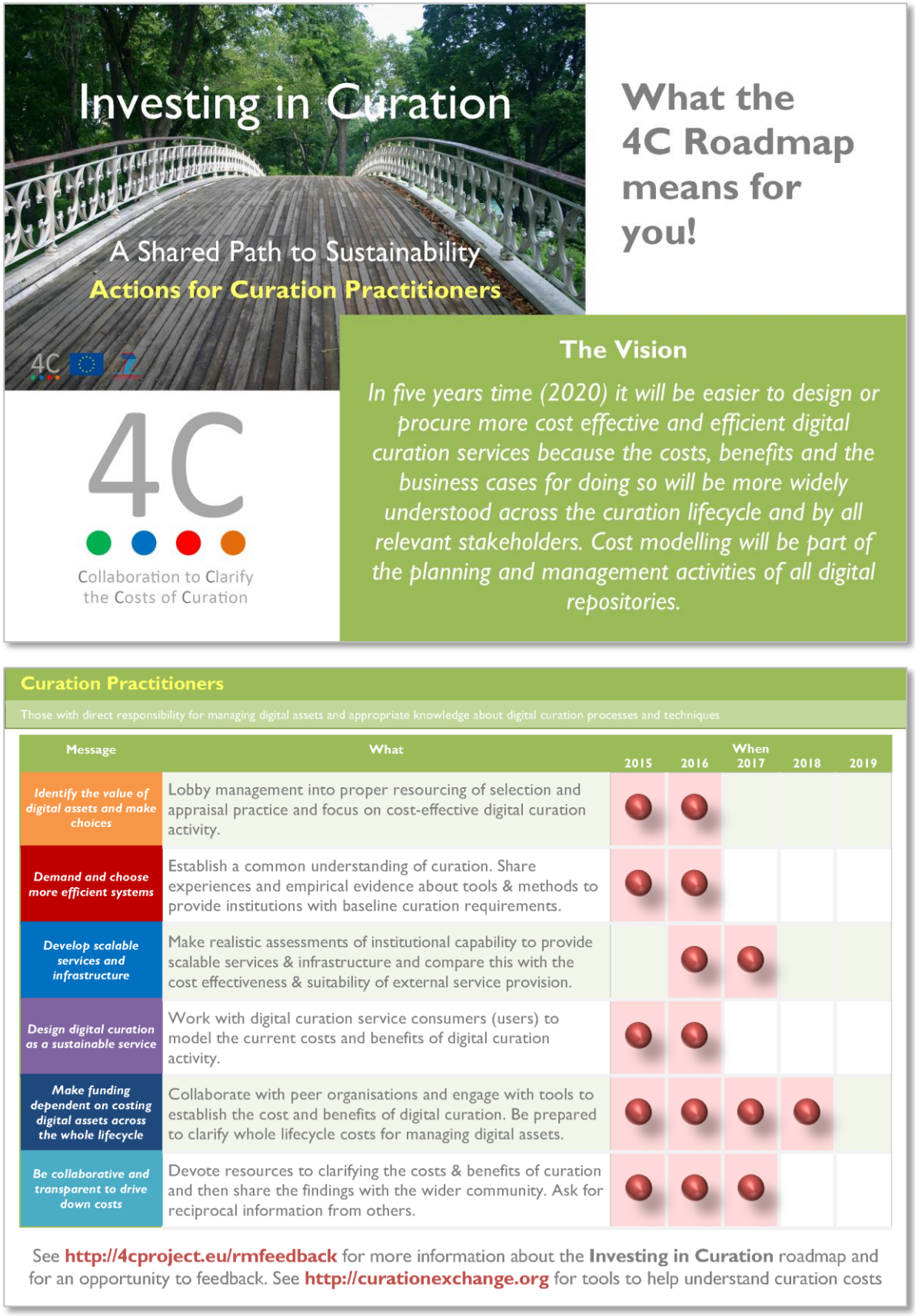


Figure 10—Actions for Curation Practitioners—English

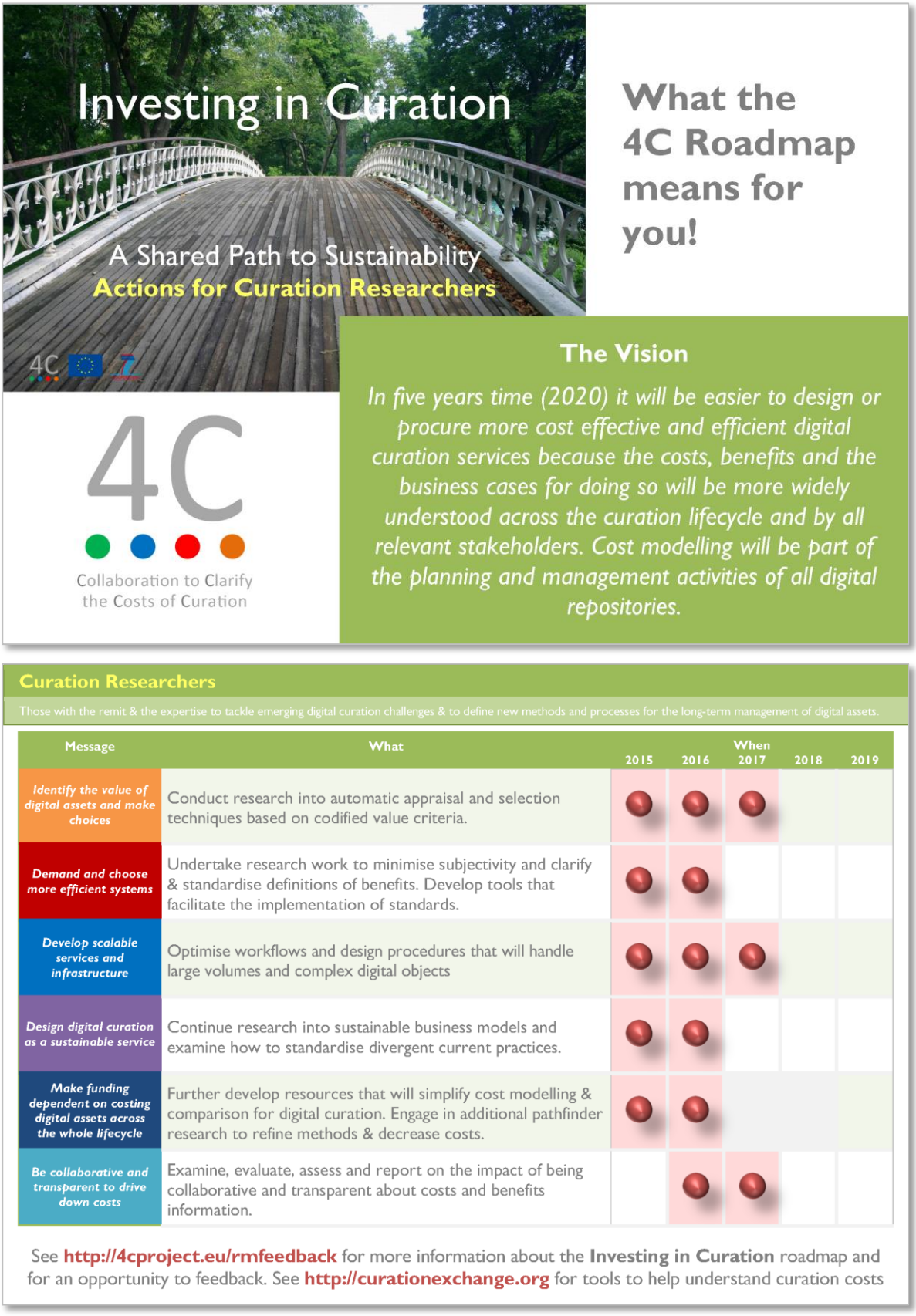


Figure 11—Actions for Curation Researchers—English

Figure 12—Actions for Data Users—English

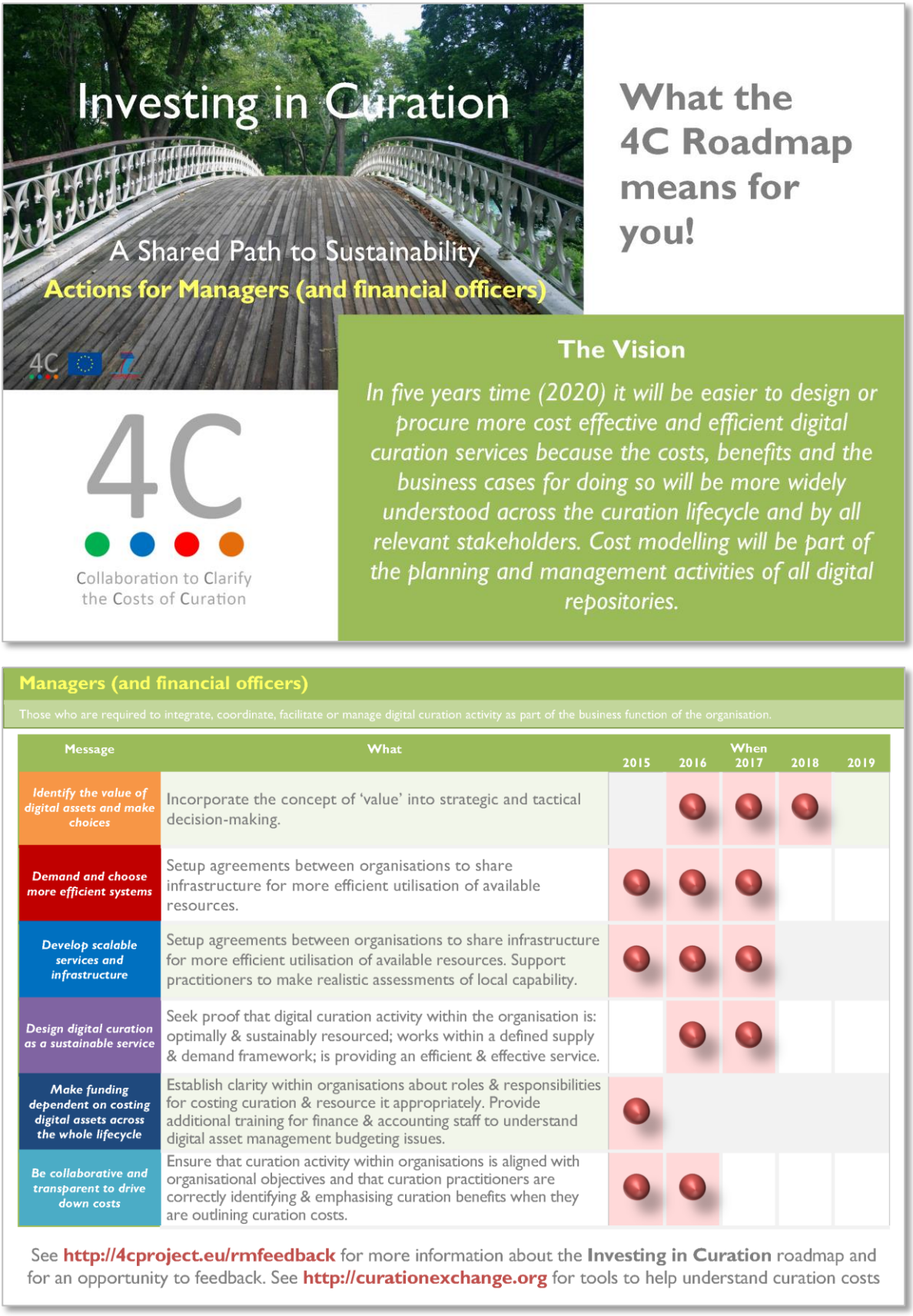


Figure 13—Actions for Managers—English

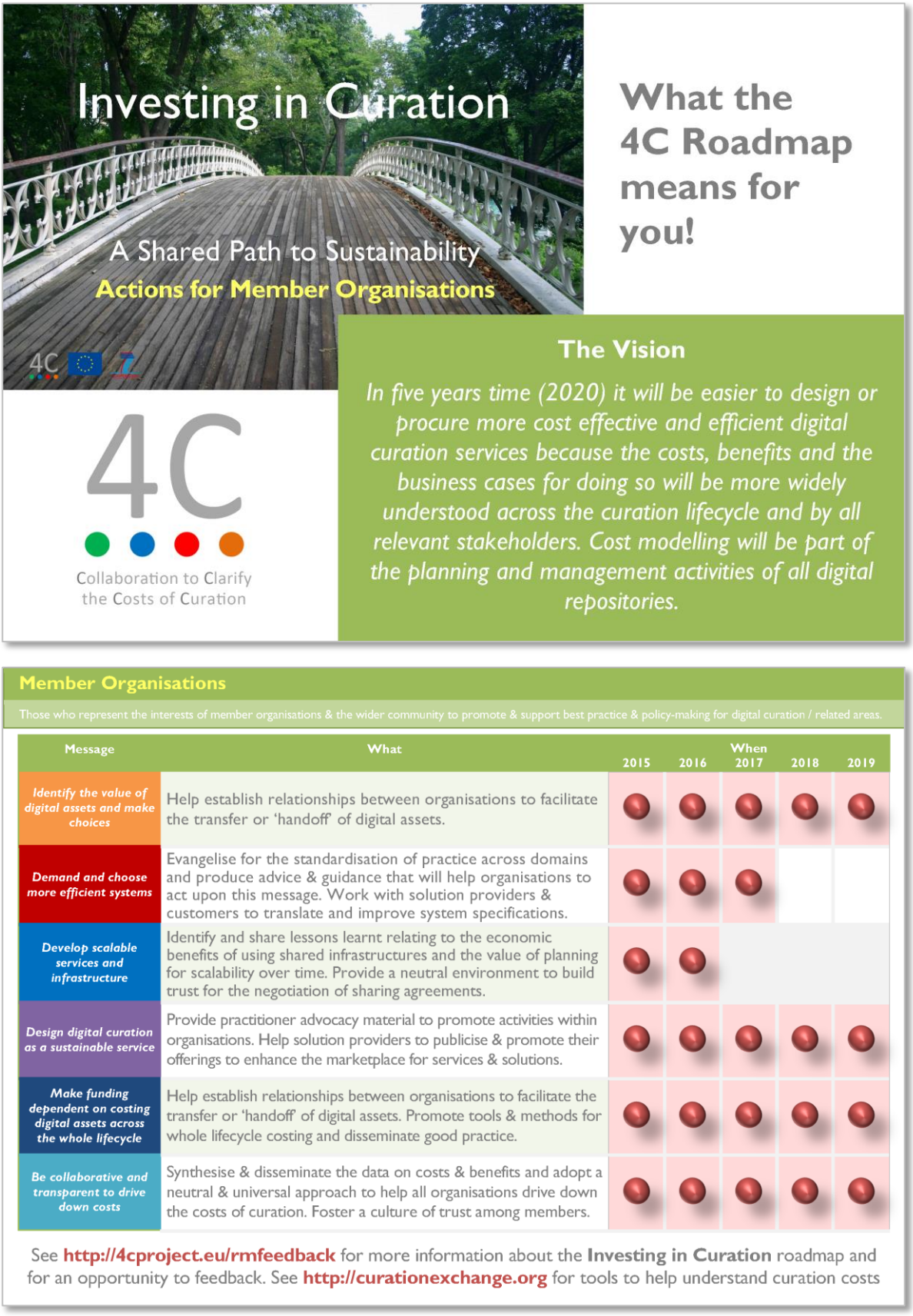


Figure 14—Actions for Member Organisations—English

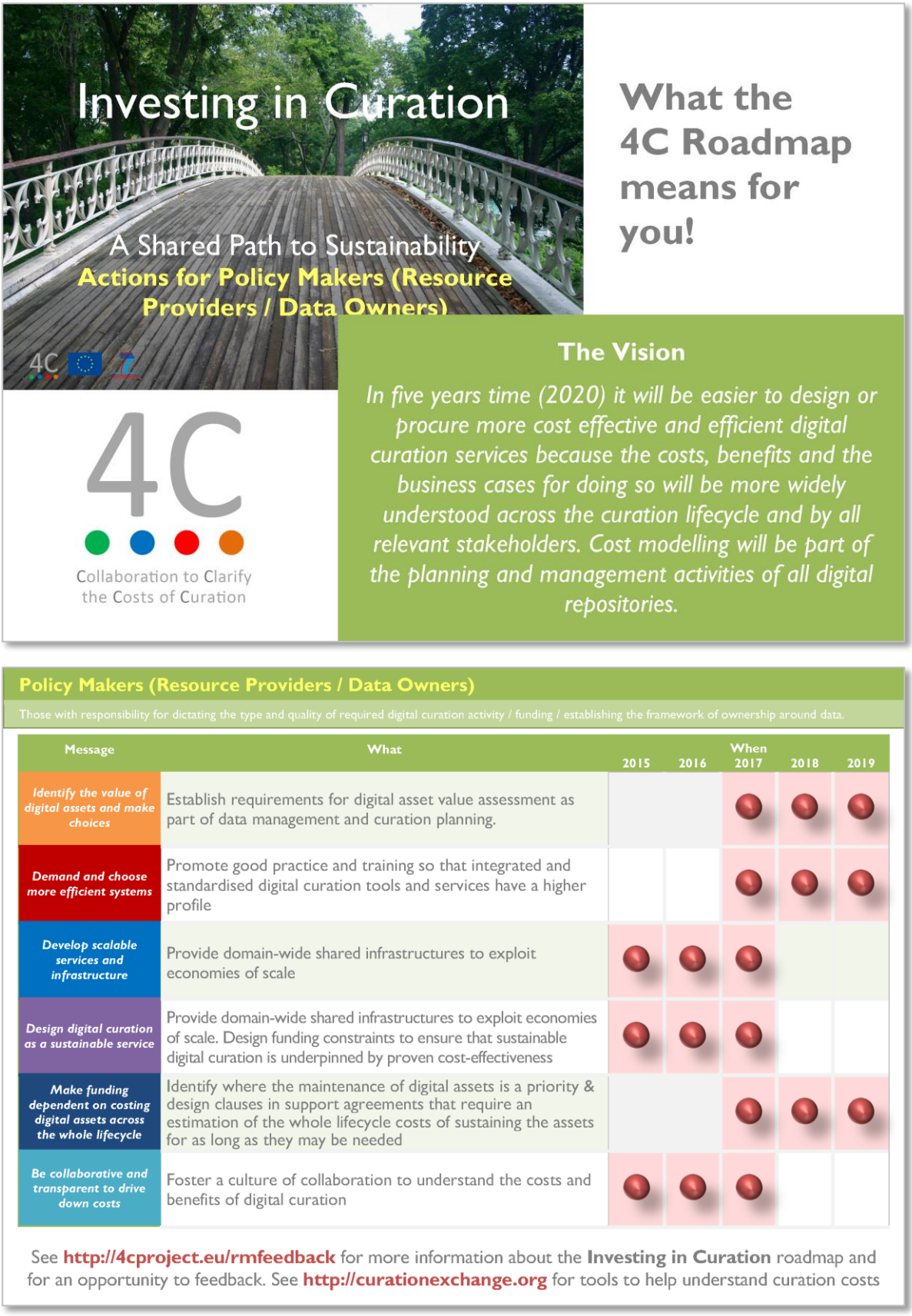


Figure 15—Actions for Policy Makers—English

Figure 16—Actions for Solution Providers—English

Stakeholder actions postcards—German

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

Figure 17—Actions for Curation Practitioners—German

Figure 18—Actions for Curation Researchers—German

Figure 19—Actions for Data Users—German

Figure 20—Actions for Managers—German

Figure 21—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 22—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 23—Actions for Solution Providers—German

Stakeholder actions postcards—French

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

Figure 24—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 25—Actions for Curation Researchers—French

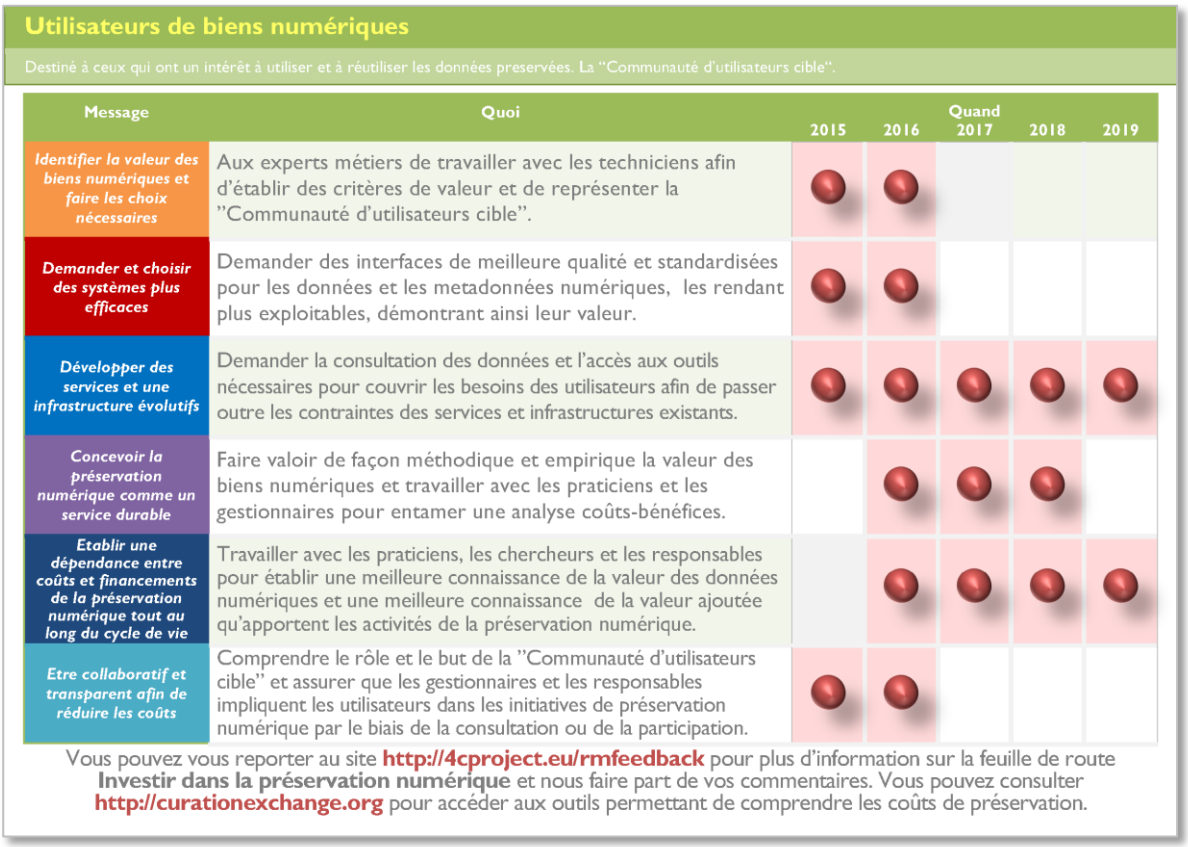


Figure 26—Actions for Data Users—French

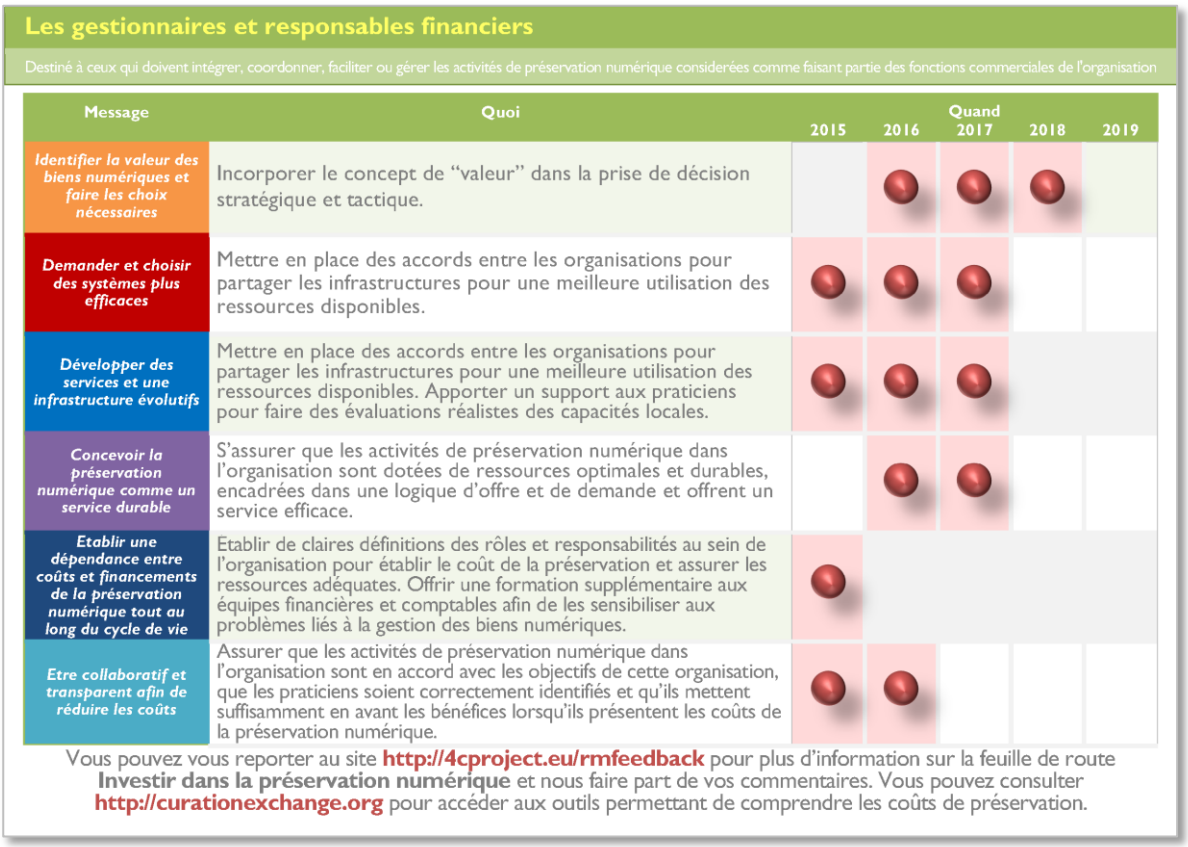


Figure 27—Actions for Managers—French

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.

Organisations membre

Pour ceux qui représentent les interêts des organisations membre ainsi que la communauté plus large pour promouvoir et soutenir de meilleures pratiques et la mise en place de règles pour la préservation numérique / domaines reliés

Message	Quoi	Quand 2015	2016	2017	2018	2019
Identifier la valeur des biens numériques et faire les choix nécessaires	Aider à l'établissement de relations entre les organisations pour faciliter le transfert ou le relais des biens numériques.	●	●	●	●	●
Demander et choisir des systèmes plus efficaces	Promouvoir la standardisation de pratiques dans différents modèles et produire des conseils et orientations pour aider les organisations à agir selon ce message. Travailler avec les fournisseurs de solutions et les clients pour traduire et améliorer les spécifications du système.	●	●	●		
Développer des services et une infrastructure évolutifs	Identifier et partager les enseignements concernant les bénéfices économiques à l'usage d'infrastructures partagées ainsi que la valeur de la planification de solutions évolutives. Offrir un environnement neutre pour construire la confiance dans la négotiation d'accords de partage.	●	●			
Concevoir la préservation numérique comme un service durable	Offrir du materiel pour promouvoir des activités dans les organisations. Aider les fournisseurs de solutions à publier et promouvoir leurs offres pour améliorer le marché de solutions et de services.	●	●	●	●	●
Etablir une dépendance entre coûts et financements de la préservation numérique tout au long du cycle de vie	Aider à l'établissement de relations entre les organisations pour faciliter le transfert ou le relais des biens numériques. Promouvoir des outils et méthodes pour établir les coûts tout au long du cycle complet et diffuser les bonnes pratiques.	●	●	●	●	●
Etre collaboratif et transparent afin de réduire les coûts	Synthétiser et diffuser les données sur les coûts et bénéfices ainsi qu'adopter une approche neutre et universelle pour aider toutes les organisations à baisser leurs coûts de préservation. Cultiver une climat de confiance entre les membres.	●	●	●	●	●

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 28—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 29—Actions for Policy Makers—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.

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 30—Actions for Solution Providers—English

Stakeholder actions postcards—Portuguese

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

Figure 31—Actions for Curation Practitioners—Portuguese

Figure 32—Actions for Curation Researchers—Portuguese

Figure 33—Actions for Data Users—Portuguese

O que *Roadmap* do 4C poderá fazer por si!

4C

Collaboration to Clarify the Costs of Curation

Visão

Dentro de 5 anos (2020) será mais fácil desenvolver ou adquirir serviços mais eficientes e económicos de gestão de informação digital, uma vez que os custos, os benefícios e os modelos de negócio subjacentes serão melhor compreendidos. A modelação de custos será uma atividade que fará parte do processo de planeamento e gestão de qualquer repositório digital.

Gestores (incluindo gestores financeiros)

Aqueles com responsabilidade por integrar, coordenar, facilitar ou gerir as atividades de gestão de informação digital nas suas organizações

Mensagem	O quê	Quando 2015	2016	2017	2018	2019
Identifique o valor da sua informação e tome decisões	Incorpore o conceito de "valor" da informação nas suas decisões técnicas e estratégicas.		●	●	●	
Exija e escolha sistemas mais eficientes	Estabeleça parcerias com outras organizações com vista à partilha de infraestruturas para uma utilização mais eficiente dos recursos disponíveis.	●	●	●		
Desenvolva infraestruturas e serviços escaláveis	Consulte os profissionais da informação que estão no terreno no sentido de determinar realisticamente a capacidade de resposta da sua organização.	●	●	●		
Considere a gestão de informação digital uma atividade sustentável	Procure evidências de que as atividades de gestão de informação são: geridas de forma eficiente e eficaz; funcionam segundo um modelo de procura e oferta; providenciam um serviço eficiente e ao encontro das necessidades dos seus consumidores.		●	●		
Torne o financiamento dependente do custo dos objetos digitais ao longo do seu ciclo de vida	Clarifique na sua organização quem são os responsáveis por determinar o custo das operações de curadoria digital e aloque recursos de forma apropriada. Dê formação adicional ao pessoal da parte financeira e contabilística para que sejam capazes de orçamentar melhor as atividades de gestão de informação.	●				
Seja transparente e colabore para reduzir custos	Certifique-se que as atividades de gestão de informação estão alinhadas com os objetivos estratégicos da organização e que os profissionais da informação são capazes de salientar os benefícios quando determinam os custos associados à gestão de informação.	●	●			

Consulte **http://4cproject.eu/rmfeedback** para mais informação sobre este roadmap. Em **http://curationexchange.org** poderá encontrar ferramentas de apoio à compreensão dos custos inerentes à gestão de informação digital.

Figure 34—Actions for Managers—Portuguese

Figure 35—Actions for Member Organisations—Portuguese

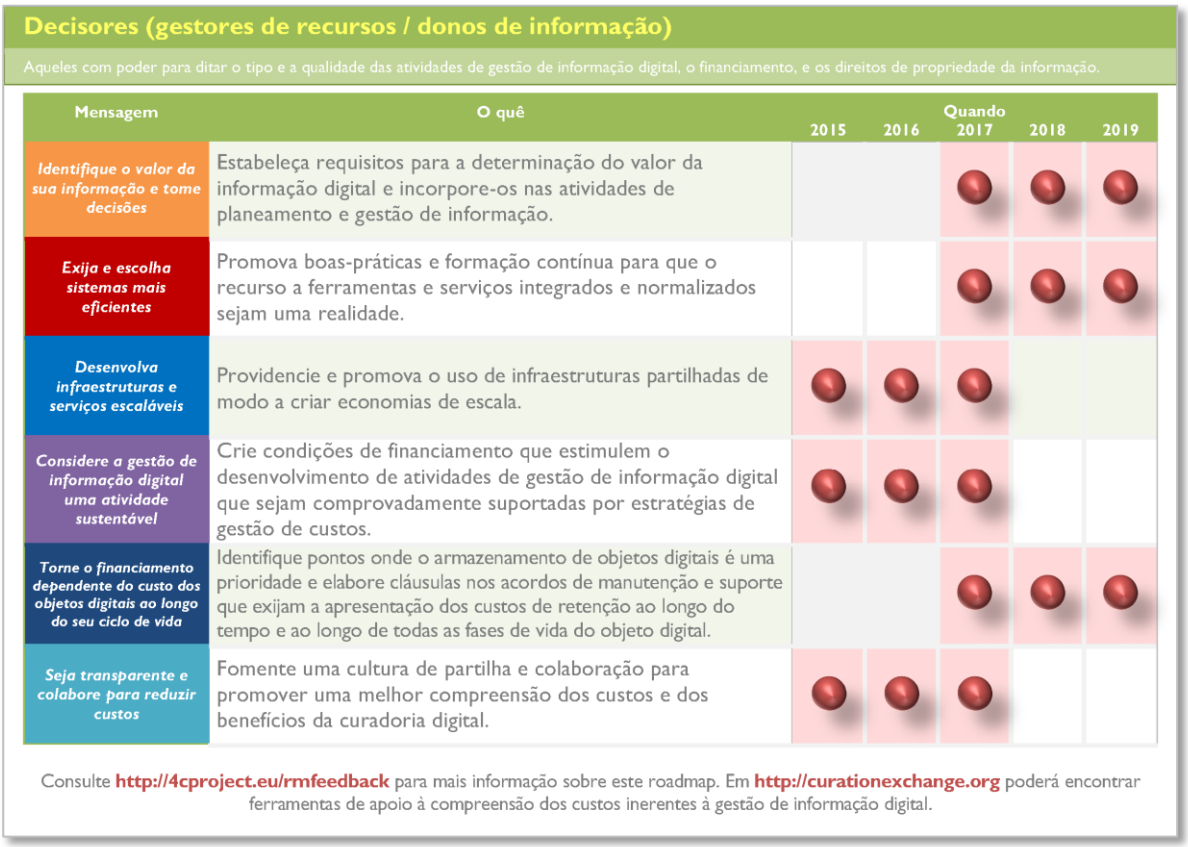


Figure 36—Actions for Policy Makers—Portuguese

Figure 37—Actions for Solution Providers—Portuguese

Stakeholder actions postcards—Dutch

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

Figure 38—Actions for Curation Practitioners—Dutch

Figure 39—Actions for Curation Researchers—Dutch

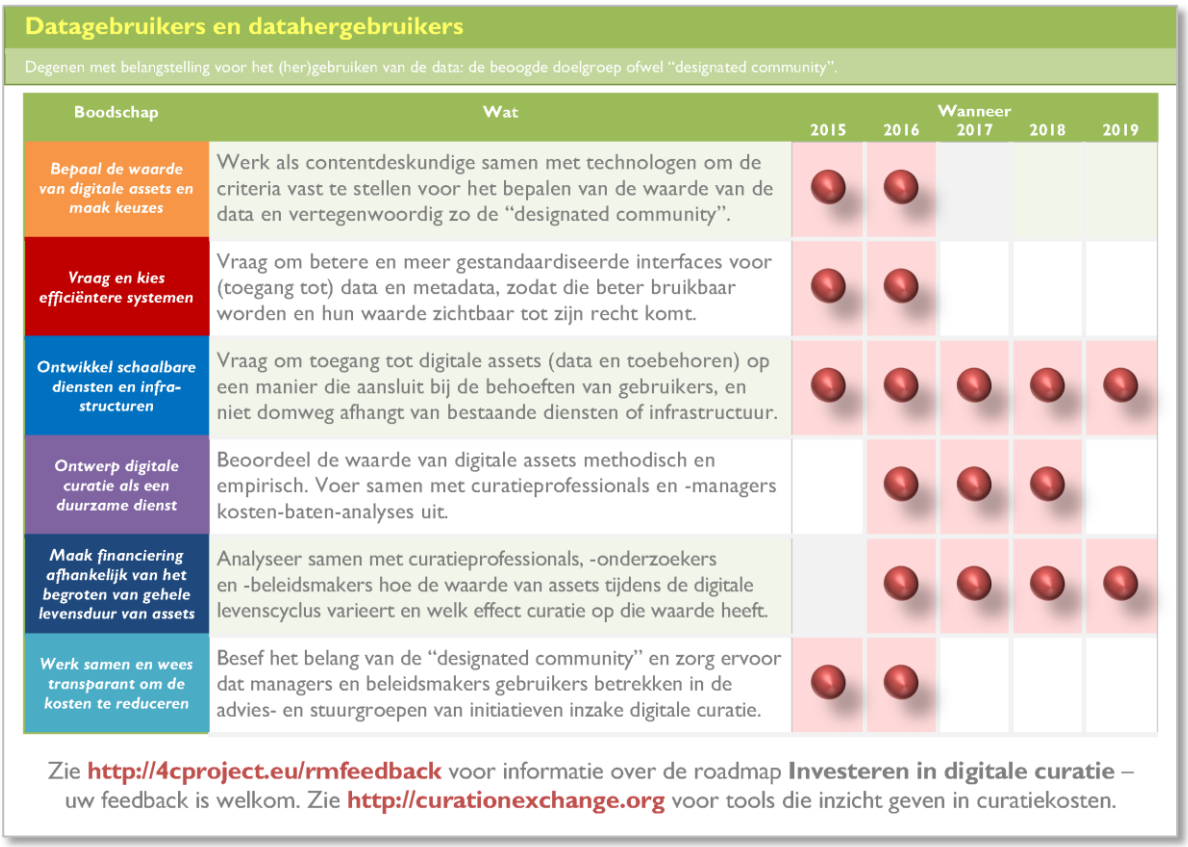


Figure 40—Actions for Data Users—Dutch

Figure 41—Actions for Managers—Dutch

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 42—Actions for Member Organisations—Dutch

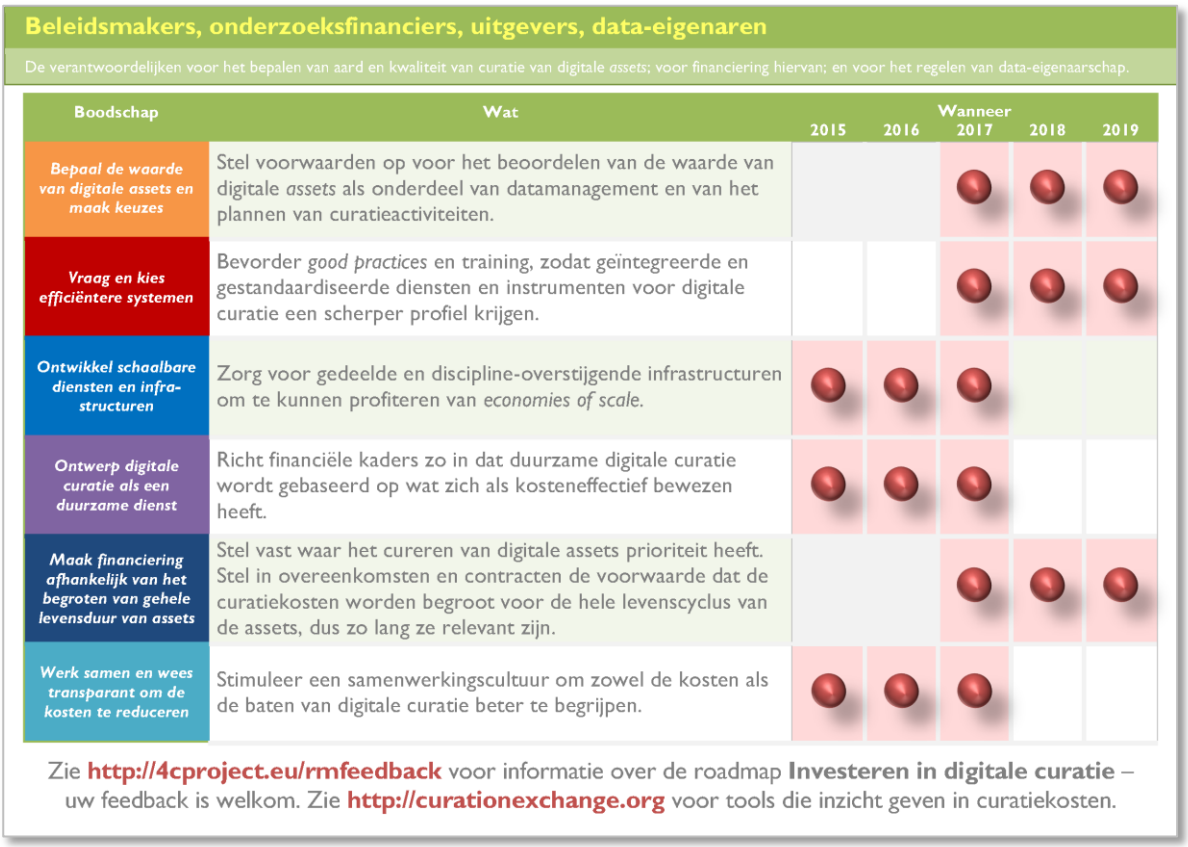


Figure 43—Actions for Policy Makers—Dutch

Figure 44—Actions for Solution Providers—Dutch

Appendix C Structured Interview Template

1. Example for a roadmap interview

*The main question can be seen from a **consumer/provider** perspective. Can we identify the problems which prevent potential customers of digital curation services from investing in existing solutions? It is possible that an organisation is a consumer of an outsourced service and providing services for its own users/consumers at the same time. Therefore it is not a mutual exclusive perspective when speaking of a consumer/provider view and more a service-oriented view.*

*Can we identify the **obstacles** which prevent the interviewee from providing a successful business model for digital curation? These obstacles depend on the stakeholder profile of the interviewee.*

Structure:

- Role
- Digital Preservation (segments of interest)
- Core services of high importance
- Conditions of operation
    - Financial
    - Legal
    - Technical
    - Organisational
    - Policies
    - Etc.
- The 'what if' budget

2. Role

How does the interviewee see his **role** in the context of digital curation? Is he a service provider, a service consumer, an advisor with consulting offers, etc.? (definition of the role)

3. Digital Preservation (segments of interest)

Which aspects of **digital preservation** are of interest for the interviewee?

- Pre-ingest,
- Ingest,
- Data Management,
- Archival Storage,
- Preservation Planning,
- Administration,
- Access

4. Core services of high importance

Are there **core services** which are of high importance to be offered or outsourced?

- AIP generation (business model example: for small memory institutions) and consulting Examples:
    - Creation of AIP including necessary descriptive information. Concept, temporal storage until deployment of the customer system

- Higher preservation services (migration, emulation, etc.)

- Archival Storage (business model example: as service provider including exit strategies, liability issues, etc.) no transformation of content information permitted or limited to exceptional circumstances
  Examples:
  - Creation of technical concept for AIP
  - Receive AIP submissions for archival storage
  - Configuration and creation of AIP
  - Error checking, media replacement, and disaster recovery (definition of SLAs)

5. Conditions of operation

Under which **conditions** does the interviewee operate?

Which conditions (financial, legal, technical, and organisational) are **not** met or remain a problem for achieving a successful business model for digital curation?

- What are the assumptions for the financial calculations/estimations? Do current solutions (services) include a sound cost model or are current solutions too expensive?
  Who else has responsibility for paying for the curation activities over time (i.e., other stakeholders) and how the costs be shared upon the various stakeholders?
  (**financial** conditions)
  Examples:
  - Estimates are based on amount of assets with additional volume for metadata
- What are the Quality of Service (QoS) parameters? Which liabilities are defined or required? How detailed are the definitions of these liabilities? How detailed are descriptions for compliance to the requirements? (**legal** conditions)
  Example:
  - Penalties and extra charges defined as part of service contracts
  - Acquisition of suitable insurance coverage
- Which services are required and why are available solutions not usable? (**technical** conditions)
  - Archival Storage serving multiple customers: heterogeneous implementations of AIP
  - Longevity of service
  - Number/volume of assets exceeds limits of services
  - Heterogeneous content of assets
  - Digital assets don't meet qualitative requirements (digitisation quality, resolution, choice of file format, missing metadata)
- Additional conditions: lack of DP knowledge/personnel or lack of other resources etc.

6. What if

As a **concluding** question an open question about optimal conditions or missing requirements to stimulate the market for digital curation services could offer an outlook for future innovations and developments.

- Into which services would the organisation invest a hypothetically budget in the size of 1-5 million Euros?
- Which future services are planned or desirable?

Appendix D Webinar Agenda

Webex Agenda

A Roadmap for digital curation

Intended to gain an overview of the problem space from various stakeholder points of view along with varied perspectives of potential paths towards a widely accepted solution.

Item	Time		Who
1	2:00-2:05	Intros Who is in attendance and what stake holder group(s) do they represent?	SM/NG/PS
2	2:05-2:25	Overview of the problem space: Addressing the questions (always taking into account the representative group) • What is the desired outcome that the roadmap – if acted upon - will deliver? • What period should the roadmap cover? • What (if anything) is preventing take up of / investment in existing solutions? What is lacking in those solutions? Where are the gaps? • Why do these gaps exist? How can they be addressed? • What is/are the top priorities to be addressed? • What is an acceptable time scale for the top priorities to addressed within?	NG/PS/All
3	2:25-2:45	Possible solutions: • How can the identified problems be sorted? • Who should be responsible for /pay for addressing those priorities • Where are the opportunities?	All
4	2:45-3:00	Analysis and critique of the 4C roadmap as it stands. • Do you recognise the challenges? • What has been missed	All
5	3:00	End	

Appendix E On-line questionnaire

4C Roadmap Feedback

We'd like your opinion about the Draft Roadmap we've produced. Just in case you've stumbled upon this questionnaire without seeing the form you can find it **here**.

We've designed this to be short. It shouldn't take you more than about 10 to 20 minutes to complete (you can of course spend longer if you've got a lot to say). Please stick with it as we really value your thoughts and impressions.

This survey needs to be completed in one pass. While it is possible to save partial surveys, you will not be able to pick up where you left off if you leave the page and restart. In addition, if you use the 'back' button on your browser or stop halfway though for a long time and the survey times out, previously entered responses may get lost.

When you've completed the survey we send you a summary of your responses using the email address you provide. We won't use your email for anything apart from this summary (unless you give us explicit permission to do otherwise)

Name:

Organisation:

Email:

1. Do you share the Roadmap's 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 curation planning and management activities of all digital repositories."

○Yes ○No

a. Please can you tell us why/why not?

2. Message 1:

"Identify the value of digital assets and make choices"

Is this message meaningful to you?

If yes, please tell us a bit more by answering the other questions below.

If no, skip straight to the next message.

○Yes ○No

a. If this message applies to you, are you prepared to act on it?

○Yes ○No

b. Do you agree with this message?

○Yes ○No

c. Is this message aimed at the right audiences

○Yes ○No

d. What are your reasons for the choices you made above?

3. Message 2:

"Demand and choose more efficient systems"

Is this message meaningful to you?

If yes, please tell us a bit more by answering the other questions below.

If no, skip straight to the next message.

○Yes ○No

a. If this message applies to you, are you prepared to act on it?

○Yes ○No

b. Do you agree with this message?

○Yes ○No

c. Is this message aimed at the right audiences?

○Yes ○No

d. What are your reasons for the choices you made above?

4. Message 3:

"Develop scalable services and infrastructure"

Is this message meaningful to you?

If yes, please tell us a bit more by answering the other questions below.

If no, skip straight to the next message.

○Yes ○No

a. If this message applies to you, are you prepared to act on it?

○Yes ○No

b. Do you agree with this message?

○Yes ○No

c. Is this message aimed at the right audiences?

○Yes ○No

d. What are your reasons for the choices you made above?

5. Message 4:

"Design digital curation as a sustainable service"

Is this message meaningful to you?

If yes, please tell us a bit more by answering the other questions below.

If no, skip straight to the next message.

○Yes ○No

a. If this message applies to you, are you prepared to act on it

○Yes ○No

b. Do you agree with this message?

○Yes ○No

c. Is this message aimed at the right audiences?

○Yes ○No

d. What are your reasons for the choices you made above?

6. Message 5:

"Make funding dependent on costing digital assets across their whole lifecycle"

Is this message meaningful to you?

If yes, please tell us a bit more by answering the other questions below.

If no, skip straight to the next message.

○Yes ○No

a. If this message applies to you, are you prepared to act on it?

○Yes ○No

b. Do you agree with this message?

○Yes ○No

c. Is this message aimed at the right audiences?

○Yes ○No

d. What are your reasons for the choices you made above?

7. Message 6:

"Be collaborative and transparent to drive down costs."

Is this message meaningful to you?

If yes, please tell us a bit more by answering the other questions below.

If no, skip straight to the next message.

○Yes ○No

a. If this message applies to you, are you prepared to act on it?

○Yes ○No

b. Do you agree with this message?

○Yes ○No

c. Is this message aimed at the right audiences?

○Yes ○No

d. What are your reasons for the choices you made above?

8. Are the messages complete?

○Yes ○No

a. Can you explain why you think this?

9. Will acting on the messages achieve the vision?

○Yes ○No

a. Can you explain why you think this?

10. Is the vision feasible in 5 years?

○Yes ○No

a. Can you explain why you think this?

11. General Comments

Do you have anything else you would like to tell us about the 4C Roadmap?

12. We'd like to follow up on some responses. Please can we get in touch with you again?

○Yes ○No

What is your role within your organisation?

Appendix F iPRES 2014 Melbourne Roadmap Workshop abstract

The 4C Project (a Collaboration to Clarify the Costs of Curation) is a European Commission funded two year coordination action which has been funded to provide useful, useable resources that provide better support to identify and quantify the cost of digital curation. From the outset, however, the project has taken the view that costs cannot be dealt with in isolation from a number of other related concepts (e.g. benefits, risk, quality, sustainability) and this holistic view might more accurately be described as an economic perspective on digital curation.

Borrowing the language of economics and mapping it onto digital curation needs to be done selectively and carefully. Digital assets do not have the same attributes as other kinds of (financial) assets and equally, it may not be possible to define when digital assets become (economic) liabilities in any objectively quantifiable way. However, there is still terminology from the field of economics that may help to define what the digital curation community might aspire to over the next few years and the starting point for this workshop is the concept of 'economic efficiency'—which might be defined as the optimised situation where it is no longer possible to add quantity or value given a finite availability of resources.

The 4C Project is tasked with delivering a Roadmap report and it is this drive towards 'economic efficiency' in relation to digital curation that will be central to the agenda that it sets out. The consultation, stakeholder engagement, analysis and modelling work that have been done allow some principles to be proposed and some assertions to be made that will form the backbone of the report. The purpose of a Roadmap—particularly where it seeks to set out an action agenda for a range of stakeholders across various communities—is to make politically astute observations and to arrive at plausible conclusions. This is only possible via early interaction with stakeholders and by achieving some level of community validation before publication and this is the purpose of the workshop. One of the guiding principles of the 4C Project is to create a better understanding of the economics of digital curation through collaboration; and also to be an 'open and social' project and to listen to the needs of the community. iPRES 2014 occurs at roughly the three quarter point of the two year project and provides a timely opportunity to check and refine the draft Roadmap.

Early ideas and discussions about the structure and content of the Roadmap have indicated that it will need to address various questions.

- What vision should we advocate and what principles should we espouse to bring about economically efficient digital curation?
- What current economic inefficiencies do we need to eliminate?
- What or who is the most influential mechanism to bring that about and where will that influence most be felt?
- What is the policy, business and regulatory framework for digital curation and how is it likely to change?
- Over what timescales should we advocate action?
- How can we most economically sustain and exploit existing work? (including the 4C Project outputs)
- How are the economic requirements of stakeholders changing?
- Is it possible and economically desirable to try and align digital curation practice (including standards and terminology)?

- How can we most effectively invest in digital curation at the institutional, national and international level?

This workshop is an important opportunity to connect with stakeholders and get input for a critical deliverable of the project. But it is also an opportunity for participants to learn more about the economics of digital curation and to critically assess the efficiency and sustainability of their own services and solutions.

Appendix G iPRES 2014 Melbourne Roadmap Workshop report

Collaboration to Clarify the Cost of Curation

Report of Roadmap Workshop at iPres 2014 on 6th October 2014 at Victoria State Library, Melbourne, Australia

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 2015	First draft	KH
1.00	07 Jan 2015	Finalised version	KH

Attendees

4C:

1. Neil Grindley, Jisc
2. Luis Faria, KEEP Solutions
3. Ulla Bogvad Kejser, KB DK
4. Katarina Haage, DNB
5. Andreas Rauber, Vienna University of Technology
6. Jose Borbinha, Tecnico Lisbon

Participants:

1. Sean Abel, Government of South Australia
2. Emma Barker, RMIT University
3. Deanne Barrett, Curtin University
4. Ed Fay, OPF
5. Juha Hakala, The National Library of Finland
6. Paul Hebbard, Simonn Fraser University
7. Ross King, AIT Austrian Institute of Technology GmbH
8. Steve Knight, National Library of New Zeland
9. Nancy McGovern, Massachusetts Institute of Technology
10. Darryl Mead, National Library of Scotland
11. Clement Oury, Bibliotheque Nationale de France
12. Stephane Reecht, Bibliotheque Nationale de France
13. Barbara Reed, Record Keeping Innovation
14. Seamus Ross, iSchool Toronto
15. Heather Rubinstein, RMIT Publishing
16. Anna Shadbolt, University of Melbourne
17. Barbara Sierman, National Library of the Netherlands
18. Barbara Signori, Swiss National Library
19. Lise Summers, State Records Office Western Australia
20. Helen Tibbo, University North Carolina

Agenda

09:00-09:15	Introduction—Neil Grindley, Jisc
09:15-09:30	Reaching out to the Community—Katarina Haage, DNB
09:30-10:30	Presentation of the 4C project outputs—Ulla Bøgvad Kejser, KBDK; Luis Faria, KEEPS; Neil Grindley, Jisc
10:30-10:55	Coffee break
10:55-11:25	Breakout session
11:25-12:00	Presentation of the draft Roadmap—Neil Grindley, Jisc
12:00-12:40	Breakout session
12:40-12:55	Feedback
12:55-13:00	Summing up—Neil Grindley, Jisc
13:00	Lunch

Minutes

Note: Because all presentation slides are available on the 4C website[20], this report focuses on an overall summary of the workshop and its main topic, the 4C Roadmap, and shows the main comments and questions from the audiences during the breakout sessions via mind maps.

The Roadmap workshop was a half day workshop held on 6th October 2014 in the scope of 2014's iPres conference in Melbourne, Australia. It was attended by 20 participants. After a brief but comprehensive introduction to the 4C project, its purposes, approaches and goals by Neil Grindley the workshop went straight off to the heart of the matter—the presentation of the Draft Roadmap that has been circulated beforehand via email and was also available at the event as a printout copy.

The 4C Project was tasked with delivering a Roadmap report and it is this drive towards 'economic efficiency' in relation to digital curation that will be central to the agenda that it sets out. The consultation, stakeholder engagement, analysis and modelling work that have been done allow some principles to be proposed and some assertions to be made that will form the backbone of the report.

Early ideas and discussions about the structure and content of the Roadmap have indicated that it will need to address various questions:

- What vision should we advocate and what principles should we espouse to bring about economically efficient digital curation?
- What current economic inefficiencies do we need to eliminate?
- What or who is the most influential mechanism to bring that about and where will that influence most be felt?
- What is the policy, business and regulatory framework for digital curation and how is it likely to change?
- Over what timescales should we advocate action?
- How can we most economically sustain and exploit existing work? (including the 4C Project outputs)
- How are the economic requirements of stakeholders changing?
- Is it possible and economically desirable to try and align digital curation practice (including standards and terminology)?
- How can we most effectively invest in digital curation at the institutional, national and international level? This workshop is an important opportunity to connect with stakeholders and get input for a critical deliverable of the project. But it is also an opportunity for participants to learn more about the economics of digital curation and to critically assess the efficiency and sustainability of their own services and solutions.

The purpose of a Roadmap—particularly where it seeks to set out an action agenda for a range of stakeholders across various communities—is to make politically astute observations and to arrive at plausible conclusions. This is only possible via early interaction with stakeholders and by achieving some level of community validation before publication and this was the purpose of the workshop. One of the guiding principles of the 4C Project is to create a better understanding of the economics of digital curation through collaboration; and also to be an 'open and social' project and to listen to the needs of the community.

[20] http://4cproject.eu/community-resources/focus-groups/workshop-4-ipres

Comments and questions on the introduction to the project

Question: How do you take care of variable costs over time?

➔ Answer: Cost submissions are tied to specific periods of time and depositors are encouraged to go back to the Exchange and repeat the exercise and update their information over time.

Comment: definitions (of activity) are of critical importance given that it is difficult to compare anything if we are not talking about the same thing

Comment: It would be good to be able to run statistical tools to analyse the costs data over time

Question: Have you done any work on comparing the cost of preserving digital in comparison with print? Lots of organisations are still very much at the stage of dealing with print material.

➔ Answer: Not as such. We have collaborated with relevant projects such as AVPreserve - 'the cost of inaction' initiative.

Idea: We need to ensure that our sustainability plan for the CCEx allows for listening to the requirements that people articulate. It is only by being flexible in what the CCEx provides that it will stay relevant as a tool.

Breakout Session 1 and 2

The following two mind maps reflect the questions, comments and ideas from the audience that was attending the Roadmap workshop:

Figure 45—Challenges identified at the Melbourne workshop breakouts

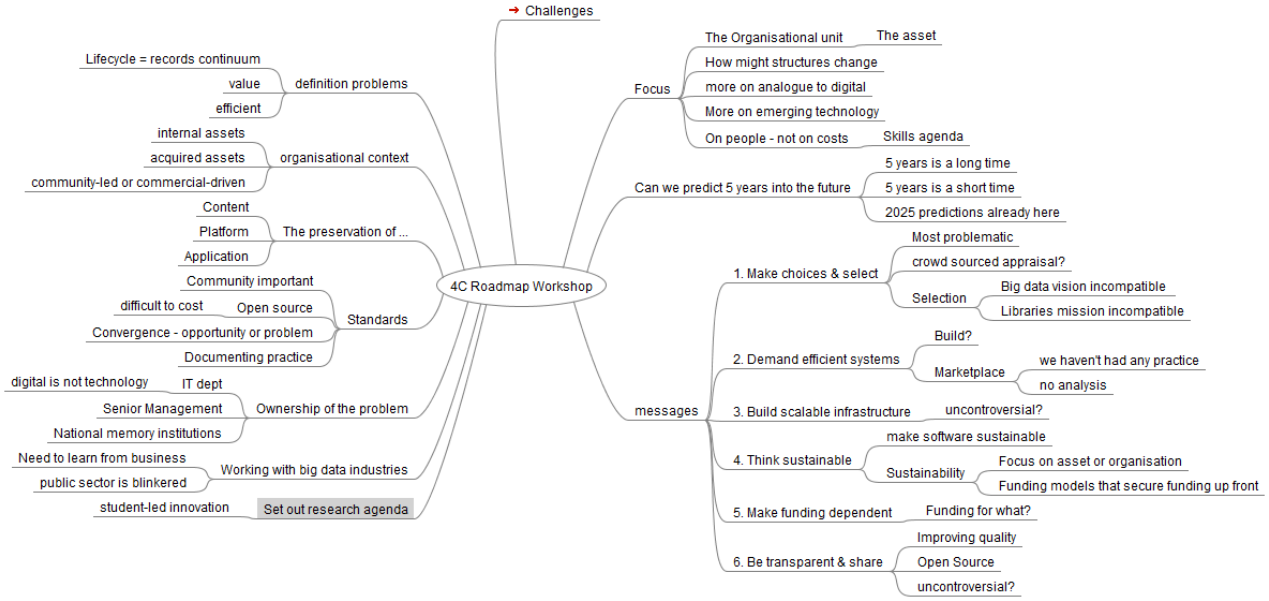


Figure 46—Roadmap issues identified at the Melbourne workshop breakouts

The workshop was concluded by Neil Grindley. He invited the participants to contribute to the development of the 4C Roadmap by providing input and taking part in the online Roadmap feedback consultation: http://4cproject.eu/rmfeedback

Appendix H iPRES Workshop blog[21]

Defining a Roadmap for Economically Efficient Digital Curation

This workshop was the first opportunity to get face-to-face feedback from the community on the draft 4C project roadmap. 'Investing in Curation: a shared path to sustainability' states six messages and sets out a number of actions that various stakeholder groups should act upon to realise a suggested shared vision that could be realised by the year 2020.

The draft Roadmap is available at: http://4cproject.eu/d5-1-draft-roadmap

The vision is as follows:

> *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 curation planning and management activities of all digital repositories.*

The workshop was divided into two main sections. Firstly participants were asked to consider the main challenges they and their institutions faced with curating digital assets (particularly in relation to economic issues). Secondly, they were asked to think about the draft 4C Roadmap messages and to consider how relevant they were to their own local context and to what extent they were plausible and sensible as an agenda for action and change.

The first discussion (challenges) surfaced the following issues:

- The scale and type of issues that will need to be faced is difficult to predict but international collaboration and knowledge exchange will mitigate the impact of that uncertainty
- There are important stakeholders (e.g. certain areas of government and publishing) who don't yet feel that curation planning is their problem or who don't yet understand that 'digital is not technology'. Or to put it another way, they haven't yet understood that digital assets are a business issue and not an IT problem.
- We need better models to understand the cost of collaboration; and to understand the scale and costs of the R&D that may be needed
- There are ownership issues that cause problems around the openness (or not) of data; about how to define the costs of distributed costs centres; who actually owns digital collections; and monolithic IT budgets that can't be broken down into departmental figures.
- Human & managerial issues (rather than technical) require additional focus and resource
- Joined up infrastructure is expensive but is a requirement
- Sustainability is a big challenge and this has to be tackled by robust business and use cases; through automation rather than manual curation processes; and by making the activities (and the assets) more visible and apparent to the organisation
- Selection is happening but techniques need to evolve to cope with appraisal at scale
- The current software solutions are inadequate so demand and requirements need to be better articulated and tools need to be more carefully specified

[21] This blog formed the core of the subsequent iPRES proceedings entry. The proceedings can be found at http://ipres2014.org/sites/default/files/upload/iPres-Proceedings-final.pdf

- There is a lot of inertia and inflexible legacy working practices within organisations that slow down ingest; limit file format choices; hinder policy development and changes to working practices
- Finding properly qualified staff and the right kind of curation expertise is hard

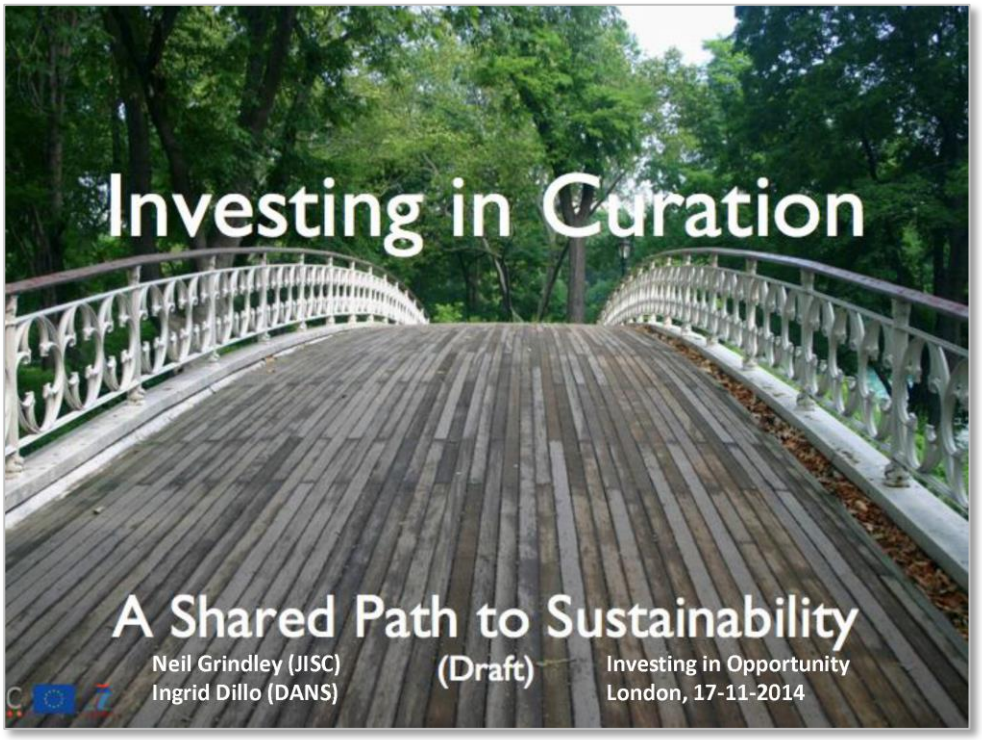
The second discussion prompted the following thoughts in response to the Roadmap:

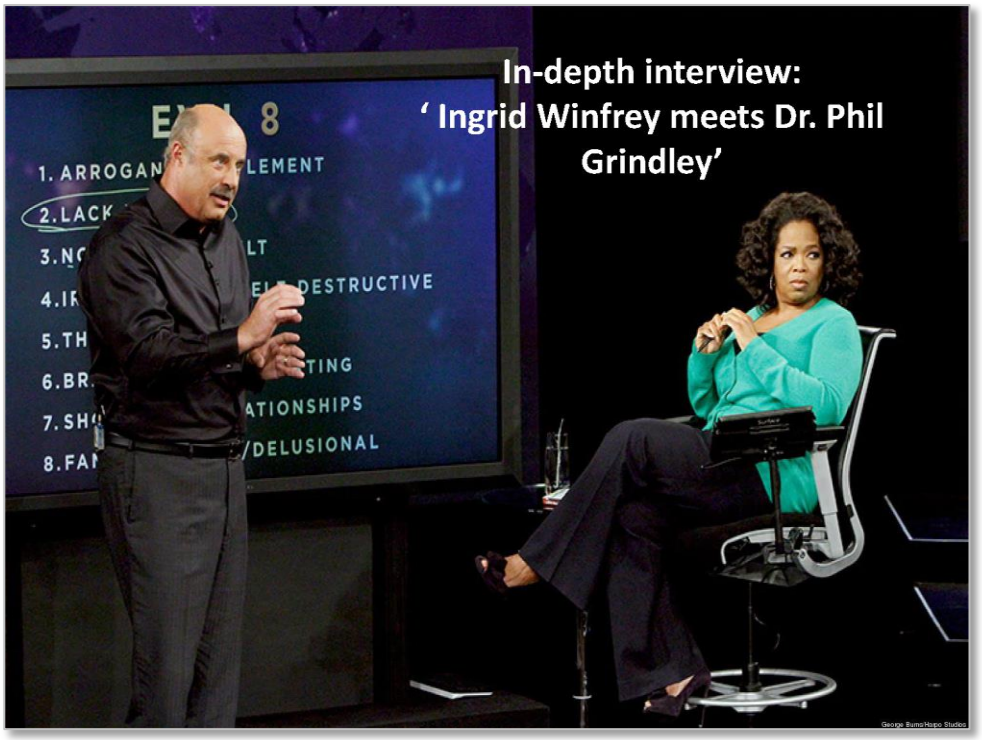
- The focus of the Roadmap is very much on the 'asset' nested within an 'organisation'; structures may change over time and an alternative or additional focus might be on people and skills and emerging technologies
- Many organisations (especially libraries and archives) are still very wrapped up in dealing with analogue collections and the transition to digital and the curation challenges associated with this
- Predicting 5 years into the future is a long or a short time depending on organisational context; the predictions for 2025 in the Roadmap are already being tackled in practice now
- Message 1 ('Make choices and select') was one of the more problematic statements. Selection may be incompatible with 'big data' techniques and may also be in conflict with the mission of some libraries; but it may also be stating the obvious or rehearsing accepted practice in environments where digital curation is established
- Message 2 ('Demand efficient systems') skews activity towards procurement rather than in-house development and assumes that there is already an effective marketplace and market analysis that can be drawn upon
- Message 3 ('Build scalable infrastructure') was an uncontroversial message
- Message 4 ('Sustainability') should extend beyond thinking about organisations and assets and should also include software and applications and embedding sustainability into up-front funding arrangements
- Message 5 ('Make funding dependent on lifecycle costing') should be clearer about what the funding will actually support and be wary of inhibiting activity entirely
- Message 6 ('Be transparent and share') should reference the power of open source and other 'open' concepts and emphasise the potential to improve quality
- There are general issues with definitions throughout the Roadmap, for example it may not be clear to everyone what is meant by, 'lifecycle', 'value' and 'efficient' in the context they are used
- There are important contextual organisational differences that need to be acknowledged, particularly in cases where assets are generated internally or acquired from external sources; and where activity is community-led or where it is commercially-driven
- Curation and preservation thinking needs to happen at the content (assets) level but also at the application (systems) level and at the platform (environment) level and this has economic implications
- The issue of standards alignment and the convergence of practice is complicated and it is not clear whether it is an opportunity or a problem and how the economics work out in terms of community practice and functional markets
- The roadmap needs to be clear about the ownership problem (see 'challenges' above) and who should be taking responsibility and in what context
- There is much that can be learnt and taken from business and big data industries; public sector organisations should be more open to these ideas to introduce more economic practices

- The Roadmap could set out more of a research agenda and provide an innovation platform for students and early-career researchers

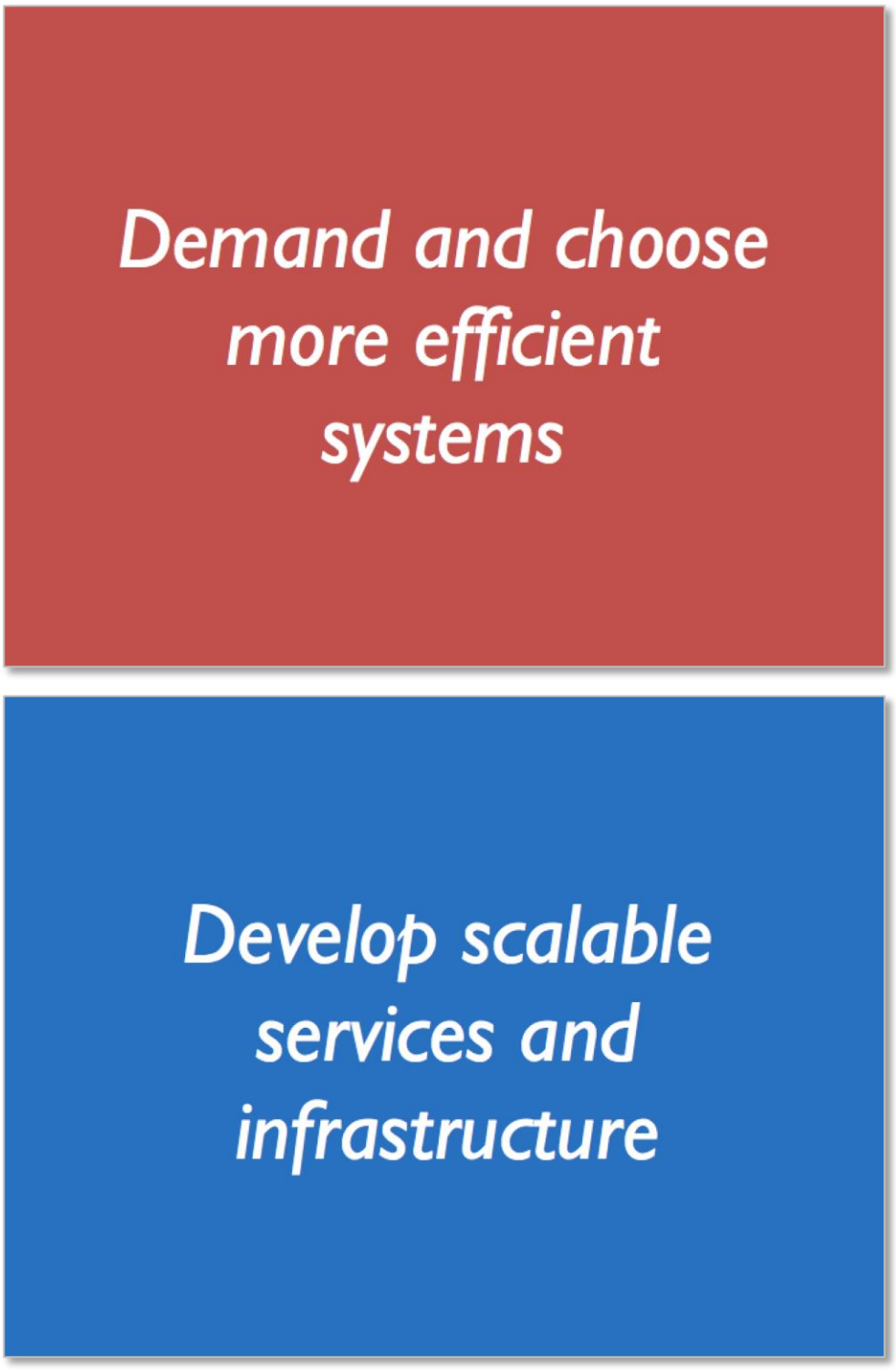
Appendix I 4C Conference Roadmap presentation—Nov 2014

The webcast of the session can be seen at
http://www.dpconline.org/events/webcast4canddpa2014/1324-4cwebcastroadmapdayone





The initial central question
The background
The vision
The Roadmap
Identify the value of digital assets and make choices


Demand and choose more efficient systems
Develop scalable services and infrastructure

Design digital curation as a sustainable service
Make funding dependent on costing digital assets across the whole lifecycle

Be collaborative and transparent to drive down costs

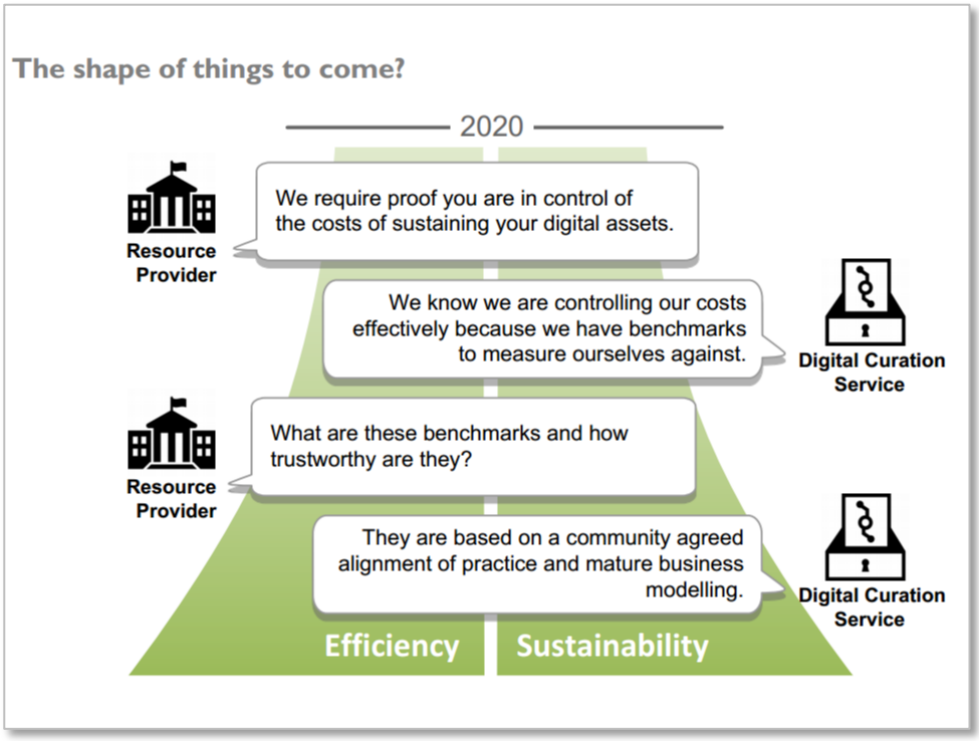


Figure 47—Roadmap presentation from the 4C Confrence November 2014