Collaboration to Clarify the Cost of Curation

D2.8—Curation Costs Exchange

Deliverable Lead:	Digital Preservation Coalition
Related Work package:	WP2
Author(s):	Sarah Middleton (DPC)
Dissemination level:	Public
Submission date:	15th 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

ersion	Date	Changed pages / reason	Modified by
0.01	23rd January 2015	First draft	SLM
1.00	15th February 2015	Released version	PLSS

Acknowledgements

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

The 4C participants are:

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

Author(s):	Sarah Middleton, DPC

Table of Contents

Figures

Tables

User Testing Figures

User Testing Tables

Executive Summary

There is a sizeable canon of research into cost modelling for digital curation. This research has tended to emphasize the cost and complexity of digital curation. However, the research is in many ways preliminary and there has been limited uptake and limited adoption of the tools and methods that have been developed. In addition there has been no integration into other digital curation processes or tools. The question is why? The 4C Project aims to address this question and provide access to resources and tools which endure.

4C seeks to help organisations across Europe to better understand the costs and benefits that accrue from digital curation and preservation, and the Curation Costs Exchange (CCEx) is just one of the key project deliverables which has been developed to support this aim.

One of the main objectives of 4C is to ensure that where existing work is relevant, stakeholders realise and understand how to employ those resources. An additional aim of the work is to examine more closely how these resources might be made more fit-for-purpose, relevant and useable by a wide range of organisations operating at different scales in both the public and the private sector.

The CCEx draws together existing, useful resources to help users make their own assessment of existing models and provide a foundation for developing their own cost modelling. In addition it provides resources to support smarter investments in digital curation through knowledge transfer and cost comparisons between organisations of all types.

Extensive community building and outreach work has been undertaken throughout the course of the project to optimise engagement with the resource, in order to sustain its relevance and currency—particularly through the further contribution of cost data sets—for as long as possible.

Since the beta launch in August 2014, the CCEx has been well received by the digital curation community and has seen:

- 4,381 hits
- 173 registrations
- 40 cost data sets uploaded

However, there remains some reluctance to publicly share cost data through the platform. When asked why this is, a few reasons were cited more often than any others. Primarily users indicated that the commercial sensitivity of certain data would make it difficult to share. The other reason often mentioned was the resource required to separate just digital curation cost data from some of the often complex accounting systems. Users have also said that they would be worried that their data would be misinterpreted without sufficient context. On the other hand, provision of such context reduces comparability and puts too much administrative pressure on those entering cost data. Whilst the CCEx attempts to address these concerns through anonymity, confidentiality, simplified (but not too simplified) data upload mechanisms and ongoing activities to communicate these features, the 4C Project Roadmap is also presented as a route towards improved sharing in the future, aiming to breakdown the taboo of sharing costs and cost information.

To summarise the majority of comments from those who have been introduced to the CCEx: they are interested—if not keen—to use the CCEx, but would really only feel *compelled* to use it only when it holds more cost data sets for comparison. This summary aptly represents the Catch 22 situation faced by the CCEx and the dependence upon the digital curation community to adopt, use and benefit from its features.

1 Introduction

This deliverable has been described in the Description of Work as follows:

"D2.8) Curation Costs Exchange: Delivery of a functional framework and platform for the exchange of curation costs-related information [month 24]"

1.1 Definition and scope of the Curation Costs Exchange (CCEx)

This document accompanies the actual Deliverable 2.8, the http://curationexchange.org (CCEx), and describes how it evolved into the online resource it has become through the efforts of the 4C project in collaboration with our stakeholders. The CCEx platform, including the Cost Comparison Tool and all related tools and resources, provides a complete online framework for exchange of curation cost knowledge and data.

The CCEx is designed to be a living platform, which will continue to evolve and be added to by the community which will own and adopt it. Even in the time it has taken to produce this report some aspects of the platform have developed further, so this document is a snapshot of the CCEx at the time of writing[1].

The deliverable 'Curation Costs Exchange' represents the whole CCEx platform[2] and all related tools and resources, including the Cost Comparison Tool as described in D3.3 Curation Costs Framework.[3]

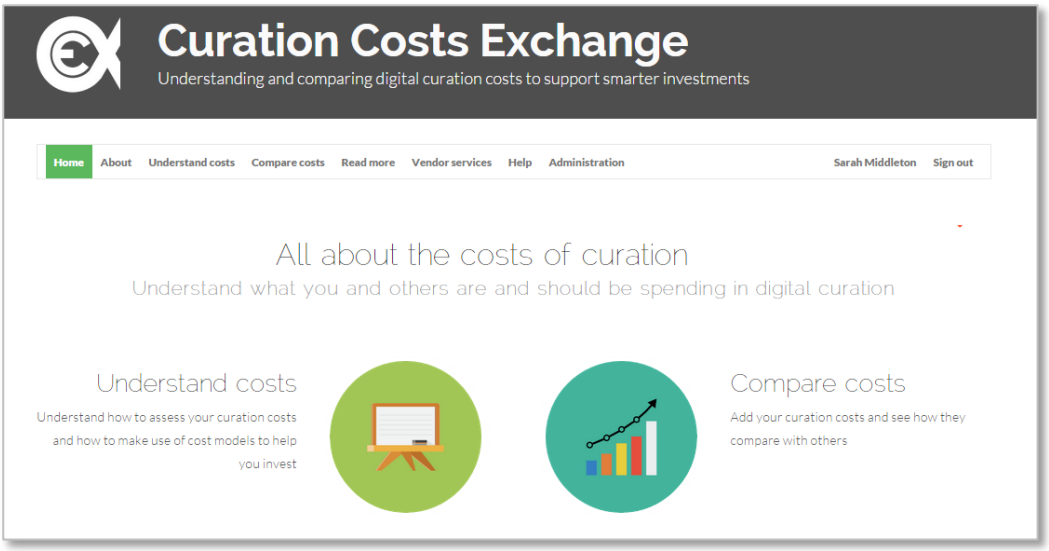


Figure 1—The Curation Costs Exchange (CCEx) platform

1.2 Purpose

Being a 'coordination and support action' project, the most important missions of the 4C project have been to gather, systematise, analyse and disseminate already existing research on the topic of curation costs; and to rally relevant stakeholders around this material and other outputs created by the project.

The development of the Curation Costs Exchange conforms to this project type in two ways:

[1] January 2015

[2] www.curationexchange.org

[3] http://www.4cproject.eu/community-resources/outputs-and-deliverables/d3-3-curation-costs-exchange-framework

1. By making relevant material and analyses available on the CCEx website;
2. By innovating and incorporating functionality, namely the Cost Comparison Tool, with which stakeholders can share and compare empirical cost data. It is through this unique feature that the CCEx aims to attract and sustain broader user engagement.

The overall purpose of the CCEx is to help users understand, collate and assess the costs and benefits of curation within their own organisation, as well as:

1. to help a range of stakeholders understand which digital curation activities might need to be considered and how they might start to assess these using existing cost models and resources
2. to provide digital curators with a platform for communicating with each other in order to clarify costs for curation
3. to help developing a cost concept model and a gateway requirement specification4
4. to enable 4C-project members to understand cost data and user behaviour by assisting them in figuring out how to structure the data and identify which parameters are necessary in order to enable comparison of the data

The CCEx thus provides a starting point for a range of stakeholders with a range of 'maturities' interested in costing curation activities as well as evaluating their benefits within their organisations. It packages and disseminates information gathered through the project including comparisons of current models, tips for making the case for investment in curation through a better understanding of cost determinants, risks, and benefits and examples of what peer organisations are doing.

1.3 Terminology

A full glossary of terminology adopted by the 4C Project is provided on the Curations Costs Exchange.[5]

1.4 Audience and benefits

The stakeholder groups for the CCEx are similar to those of the 4C-project as a whole[6] . However there is scope to modify the functions of the CCEx to cater for additional stakeholder groups should the need arise At present the CCEx caters for 11 key stakeholder groups:

1. Research funders
2. Big data science
3. Digital preservation vendors
4. Government agencies
5. Publisher or content producers
6. Data intensive industryies
7. Memory institutions or content holders
8. Small or medium enterprises
9. Universities
10. Researchers
11. Other

[4] http://www.4cproject.eu/d3-2-ccm-grs. The two project deliverables (D3.2 and D3.3) have been developed concomitantly and influenced each other

[5] http://www.curationexchange.org/read-more/67-4c-project-outputs

[6] http://www.4cproject.eu/d2-1-stakeholders

In order to better understand what each stakeholder group would look for in the CCEx (and in particular the Cost Comparison Tool), use cases were created. The examination of user behaviours provided information on the benefits of using the CCEx and emerge from the table on the following pages.

Stakeholder Group	Possible Roles	What information to find in CCEX	What to be able to do using CCEX	Information we provide initially to meet expectations	What would the main benefits be?	Information stakeholder would be willing to share	Would data be anonymised?
Research funder	Funding programme manager; provide guidelines to peer reviewer of grants	I would expect to find out what kinds of activity costs might be eligible to be covered in-project (what and how much time is "reasonable").	I want to ensure that our guidance helps applicants to understand eligible costs. I want to make sure that the research funds are used in the most cost-effective way and yield impact (Return on Investment).	Distillation of UKDA[7] cost tool (activities and guidance on when costs are most effective—in project or at Ingest to an archive); Overview of KRDS[8] activity models as reference to activities to consider; Link to DCC[9] funders' policy overviews.	The Cost Comparison Tool will enable the identification of activity costs which might be covered 'in-project' e.g. what/how much time is "reasonable," to ensure that funder guidance helps applicants to understand eligible costs, and that research funds yield the best return on investment.	I'd be willing to share our policies on including costs and provide pointers to guidance.	No
Big data science	Research Infrastructure (RI) provider	I want to understand how curation costs should be included in my operational budget and to find and assess suitable third curation party services we might need to procure.	I want to be able to ensure that my RI facility operates efficiently and to be able to plan for sustainability. I want to find the right service provider for my particular needs and be able to understand their charging models.	Comparison with other data stewards (data centres, national libraries). Links to preservation and curation vendors and services.	The Cost Comparison Tool will enable the identification of relevant costs in terms of domains, scale and activities, to assist with effective budgeting for new projects and making reasonable estimates from the outset.	I'd be willing to provide costs relating to specific services offered by our RI facility.	Optionally
						I'd be willing to share feedback on the level of satisfaction with the procured service or product.	Optionally

[7] UK Data Archive.

[8] Keeping Research Data Safe.

[9] Digital Curation Centre.

Stakeholder Group	Possible Roles	What information to find in CCEX	What to be able to do using CCEX	Information we provide initially to meet expectations	What would the main benefits be?	Information stakeholder would be willing to share	Would data be anonymised?
Digital preservation vendor	Preservation service provider	I would expect to see real cost data related to operating a digital preservation service (either as a third party service or in-house solutions)	I'd like to compare my expenditure with others and learn how others are spending their budget. I want to see how I compare with competitors and peers and understand the reasons for significant disparities in costs, so I can implement possible remediation for my high cost activities	Access to real cost data.	The Cost Comparison Tool will show real cost data related to operating a digital preservation service (either as a third party service or in-house solutions) and will enable comparisons of expenditure with that of competitors and peers, providing a greater understanding of the reasons for significant disparities in costs, and enabling the implementation of possible remediation.	I'd be willing to share my own cost data to get a fairly accurate result in the comparison.	Yes
Government agency	Preservation manager of a National Library	I would be most interested in browsing through cost model descriptions to compare my expenditure with others. I'd like to know which cost model might best meet my organisation's needs. I'd like to see how other National Libraries define curation related risks and benefits.	I'd like to be able to support proper planning of a new content stream being taken into the archive. I'd like to be able to compare the costs of preserving different types of collection content (e.g., an e-journal collection). I want to find out if other National Library are spending their budgets more efficiently (e.g., preserve larger data holdings with fewer resources) so I can decide if we need to look into our own workflows to make them more efficient.	Overviews of current cost models. Suggestions on which models are suitable for specific types or organisations and/or content. Comparison with other data stewards (data centres, national libraries).	The CCEx will provide cost model descriptions that will help to inform which cost models to use. It will enable comparison of expenditure between government organisations and assist with planning and budgeting.	I would be willing to share information on cost models we employ.	No
						I would be willing to share feedback on suitability of cost model for my organisation/collection type.	No
						I'd be willing to contribute our own curation cost data to help develop more accurate comparisons between National Libraries.	Optionally

Stakeholder Group	Possible Roles	What information to find in CCEX	What to be able to do using CCEX	Information we provide initially to meet expectations	What would the main benefits be?	Information stakeholder would be willing to share	Would data be anonymised?
Publisher or content producer	Repository manager	I would expect to find out how I might be able to reduce my curation costs and increase efficiency in service delivery.	I'd like to compare my costs with similar repositories.	Access to real cost data.	The Cost Comparison Tool will enable comparison of cost structures with those of other companies, highlighting areas of potential cost-reductions in day to day activities.	I would be willing to input my operational and /or cost data.	Yes
Data intensive industry	Financial department of large broadcasting company	I want to find out what other large national organisations are spending to preserve access to their digital collections.	I'd like my company to be able to guarantee access to our data for as long as it is required whilst reducing risk in a cost-effective way. If possible, I'd like to see a return on our investment.	Access to cost comparison tool and short case studies.	The Cost Comparison Tool will show what large national organisations are spending to preserve access to collections so that access may be guaranteed to data for as long as it is required, while reducing risk in a cost-effective way.	I would be willing to share information on our curation services and the costs or operation.	Optionally
Memory institution or content holder	Community archive manager	I would expect to see general information that would help me to manage limited budget resources.	I've been tasked with doing some form of planning or analysis exercise which involves examining the costs / benefits / business case for future digital curation activity and I don't know how to start. I'd like to know what larger institutions are spending in certain infrastructure areas as reference but I'm more interested to find out what other smaller institutions are spending. I'd like to be able to find information that will allow me to benchmark our approach.	Getting started guidance. Overview of cost models, risks and benefits. Access to cost comparison tool and short case studies.	The Cost Comparison Tool will offer controlled contact and comparisons between large and small institutions with a view to assisting in managing limited budget resources, creating business cases for future activity and benchmarking—with accuracy indicators to manage the different scales of operation.	I'd be willing to cooperate with and share advice and guidance with other smaller institutions	Possibly
						I have some historical costs data and/or emerging estimates of costs that I'd be willing to share to help other smaller institutions.	Optionally

Stakeholder Group	Possible Roles	What information to find in CCEX	What to be able to do using CCEX	Information we provide initially to meet expectations	What would the main benefits be?	Information stakeholder would be willing to share	Would data be anonymised?
Small or medium enterprise	DP practitioner that is relatively new to the field	I would like to browse through cost models to find out what is out there and is potentially useable in my organization.	I'd like to be able to propose a viable approach to senior management at my organisation outlining a feasible approach to implement (make the case).	Access to a list of risks and benefits relating to digital curation. Advice on how to develop and integrate cost modelling into organisational budgeting and accounting methods.	The CCEx will provide an assessment of different available cost models in order to identify which is the most suitable for a particular type of organisation.		
University	Research administrators, IT managers, library staff	RAs—I would like to know what to budget into new grant proposals to cover curation costs during the active phase of research. IT and library staff—we would like to make sure we can afford to provide and sustain curation services and storage in line with researchers' needs.	I would like to be able to cost curation activities into new proposals. I would like to be able to budget for sustainable service provision.	Access to activities that should be costed into proposals. Access to cost information for peer organisations.	Our institutions would be better able to plan for and sustain curation costs across the research lifecycle.	RAs—I would provide details on financial systems we use within the institutions. IT and library—we would share our costs so we can see how we compare with our peers.	Optionally
Researcher	Principal Investigators	I would expect to find out what kinds of activity costs might be eligible to be covered in-project (what, and how much time is "reasonable"). I want to see costs relevant to me in terms of domains, scale and activities.	I want to be able to make a good case for justifying additional resources in my application, so I can ensure I can get funded and I meet funders' expectations about Research Data Management and sharing.	Access to guidance on curation activities to consider when developing new project proposals.	To compare with other, similar projects; to see where I could bring my costs down; to comply with funders' expectations, both regarding data management, preservation and expenses	I'd be willing to share information about the kinds of curation activities that I'm planning to undertake. I might be willing to share my data management plan (DMP).	Optionally

Table 1—Use cases for the CCEx

2 Rationale and Methodology

This rational and methodology section describes the process by which the CCEx framework was developed, tested and implemented, taking into consideration the needs and requirements of project stakeholders.

2.1 Consultations

4C is an 'open and social' project and the key to its success has been its ongoing engagement with a wide range of stakeholders. Starting with the idea that greater transparency around the supply and demand of curation services could change in the way that all organisations think about and sustainably manage their digital assets, the project team undertook a range of consultation activities prior to collating the CCEx platform to ensure that this resource met users' needs and requirements, and would continue to be used after the end of project.

Consultation methods included an online survey, focus groups and workshops (see D2.4—Final Report on Outreach Events[10] for a full list of activities) which gathered input from representatives from across the project's identified stakeholder groups.

The results of these consultations were collated into the following early 4C Project deliverables and used to inform the direction of the CCEx development:

- D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives[11]: An Initial report identifying project stakeholders and work relevant to the purpose of the CCEX
- D3.1—Evaluation of Cost Models and Needs & Gaps Analysis[12]: An analysis of existing research to identify strengths and weaknesses of current cost models, identifying emerging good practice and any gaps in provision and community requirements
- D4.1—A prioritised assessment of the indirect economic determinants of digital curation[13]: An evaluation of the relative importance economic determinants and how these influence decision making in digital curation investments

2.2 Examination of existing resources

Building on the concept of an 'open and social' online crowd-sourced database of curation cost information (and using the results of the consultations in the deliverables listed above) the project identified the information which the digital curation community should find useful and hence which ought to be accessible through the platform.

[10] http://www.4cproject.eu/d2-4-final-report-on-outreach-events

[11] http://www.4cproject.eu/d2-1-stakeholders

[12] http://www.4cproject.eu/d3-1

[13] http://www.4cproject.eu/d4-1-ied

In addition to a facility for the exchange and comparison of actual cost data, the CCEx needed to provide a curation costs *information* exchange. To satisfy this requirement, the 4C looked to resources from within and beyond the project. Some of the resources identified for inclusion were delivered at different times and had different opportunities for validation due to the project timetable. This process resulted in the shortlist of deliverables and categorised material for inclusion as follows:

'Understand Costs'

- D3.1—Evaluation of Cost Models and Needs & Gaps Analysis: An analysis of existing research to identify strengths and weaknesses of current cost models, identifying emerging good practice and any gaps in provision and community requirements.
- D3.2—Cost Concept Model and Gateway Specification[14]: A meta-model against which current and future cost models can be benchmarked. A common model based on common concepts and a generic specification (a gateway specification) that can be used in follow-up Research & Development projects.
- D4.1—A prioritised assessment of the indirect economic determinants of digital curation: An evaluation of the relative importance economic determinants and how these influence decision making in digital curation investments.
- D4.2—Draft Economic Sustainability Reference Model[15]: Support the design of strategy; to influence and standardise terminology; and to assist with the declaration of roles and responsibilities in relation to curation and preservation. The focus on sustainability rather than cost per se is designed to encourage stakeholders not to think just about the expense of long-term information management, but more holistically about the value and benefit of that information.
- D4.3—Quality and trustworthiness as economic determinants in digital curation[16]: 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.
- D4.4—Report on Risk, Benefit, Impact and Value[17]: An examination of a range of inter-related issues, including the role of risk and risk assessment, benefit, impact and value, both terminologically and by sector.

'Read more'

- D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives: The Initial report identifying project stakeholders and relevant work developed in to a project bibliography and registry of resources for CCEx users.

'Curation Services'

- Addendum to D3.1—Evaluation of Cost Models and Needs & Gaps Analysis: Following the Year 1 review it was recommended that a list of service providers and solutions be incorporated into the 4C Project's stakeholder considerations. This list has been developed into a registry of resources for CCEx users.

[14] http://www.4cproject.eu/d3-2-ccm-grs

[15] http://www.4cproject.eu/d4-2-esrm-2

[16] http://www.4cproject.eu/d4-3-quality-and-trustworthiness

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

These sections would also contain links to other existing and relevant work on the subject of digital curation costs.

It should be noted that some deliverables, such as the Community Validation of ESRM (and the associated Digital Curation Sustainability Model [DCSM]), were unavailable at the time of writing—they were still in development. It is anticipated that these will be incorporated in the near future.

2.3 Beta testing

On the basis of the material identified, plus the Cost Comparison Tool (see D3.3[18] for full description), the CCEx was launched for beta testing on 5th August 2014, running as a beta service until 15th September 2014.

During this time feedback was elicited through:

- Targeted invitations by email to specific stakeholder contacts
- Workshops and focus groups
- Conference papers, posters and presentations
- General invitation to the digital curation community extended through email announcements, articles on the 4C and project partner websites
- Blogs on the 4C website inviting use and comment
- Tweeting the CCEx link inviting use and comment

Feedback was collated through an online survey sited on the CCEx platform, through the Usersnap tool[19] and direct comments. During this period alone, the CCEx platform received 1,341 hits and 8 registrations.

During beta testing, users were also encouraged to share their cost data through the CCEx as a means of testing the submission workflow. See D3.3 for more details.

2.4 Soft launch and release

Following beta testing, the CCEx was launched to the digital curation community as follows:

- iPRES conference in Melbourne, Australia on 6th – 10th October 2014
- 4C Conference in London, UK on 17th – 18th November 2014

These two events provided the opportunity for stakeholders to see the CCEx demonstrated live, to ask questions and to provide further feedback.

Refinements to the platform continued to be made based on feedback received following the soft launch, until 31st January 2015. After this date the platform was taken over by a mini-consortium of organisations related to the project (as outlined in the continuity plan). It remains on the same server and there has been no break in service. Improvements continue unabated.

[18] http://www.4cproject.eu/d3-3-curation-costs-exchange-framework

[19] https://usersnap.com/

2.5 Usability and Testing

Following the soft launch and prior to the end of the project, targeted usability and testing was carried out over the months of November and December 2014, using the using the criteria outlined on the Usability Test Plan Dashboard shown below:

Product Under Test	What's being tested? The Curation Costs Exchange tool (CCEx)

Test Objectives	What are the goals of the usability testing? What specific questions will be answered? What hypotheses will be tested?
	• This usability test is intended to determine the extent to which an interface facilities a user's ability to complete routine tasks:
	  ○ Do people understand the concept of the tool?
	  ○ Do people trust the tool?
	  ○ What are the common problems when using it

Participants	How many participants will be recruited? What are their key characteristics? 5-8 participants recruited through the University of Edinburgh. For example Curation Practitioners (for example archivists, records managers) and Managers (such as IT/finance managers, head of library)
Test Tasks	What are the test 'critical' tasks? *Examples*:* • Sign up for the CCEx • Enter information about your dataset and the costs associated with it • Compare costs with peers • Contact support • Amend your dataset *A complete list of tasks with scenarios will be circulated later as it will need to reflect development changes made to the website/tool.
responsibilities	Who is involved in the test and what are their responsibilities? • Magdalena Getler, UEDIN-DCC (facilitator, observer) • Diana Sisu, UEDIN-DCC (facilitator, observer)
Business Case	Why are we doing this test? What are the benefits? What are the risks of not testing? • The test will address several key questions that the design/development team need answers to for the next iteration • Failing to answer these questions now increases the risk of developing the wrong tool

Equipment	What equipment is required? • Computer and a voice recorder • We will log usability problems, measure task completion rate, time on task and list errors that will be later mapped to usability problems
Locations & Dates	Where and when will the test take place? When and how will the results be shared? • Date and time (TBC) • Sessions will be recorded and analysed to identify potential areas for improvement to the tool. • Findings with recommendations will be shared with the tool developers and 4C WP3 group
Procedure	Main steps in the test procedure: • 0-4 min: Welcome/Consent Form • 5 min: Pre-test interview and questionnaire • 3 min: Homepage Tour • 35 min: Carry out the test tasks • 15 min: Post-test questionnaires and interview • 5 min: Debrief/Incentive

Table 2—Usability Test Plan Dashboard

The project invited five participants at the University of Edinburgh to complete six key tasks, under observation, and give their opinions on the user interface as well as the look and feel of the web site and in particular the Cost Comparison Tool (CCT).

At the beginning of the test participants were asked to answer the following questions:

- Are you responsible for managing budgets in your current role? If yes, how much of this information comes from other people within the organisation?
- How big is the budget that you manage?
- Have you ever used an online tool to help plan and manage costs in your organisation?

Participants were also asked to rate the expected difficulty of the task both before completing the task (the expectation ratings) and after (the experience rating). The ratings were on a 7-point scale with measurements ranging from *Very difficult* to *Very easy*. Participants were asked the following two questions:

- Before doing all tasks (expectation rating): "How difficult or easy do you expect this task to be?"
- After doing each task (experience rating): "How difficult or easy did you find this task to be?"

At the end of each session participants were requested to subjectively assess the usability of the tool using a standardised SUS (Software Usability Scale) questionnaire on a 5-point Scale with endpoints of *Strongly disagree* (1) and *Strongly agree* (5). Post-task scenario subjective measures included:

1. I think that I would like to use Cost Comparison Tool (CCT) frequently
2. I found CCT to be simple
3. I thought CCT was easy to use

4. I think that I could use CCT without the support of anyone else (my colleagues / support staff such as librarians / IT staff, etc.)
5. I found the various functions in CCT were well integrated
6. I thought there was a lot of consistency in CCT
7. I would imagine that most people would learn to use CCT very quickly
8. I found CCT very intuitive
9. I felt very confident using CCT
10. I could use CCT without having to learn anything new

The results have been collated into a summary report which can be found in Appendix 1. The findings were incorporated into developments for the final version of the CCEx platform. As a result of the changes made to the final CCEx there may be some features mentioned in the report which have been updated or removed in the current version.

3 Description of the CCEx Platform

The CCEx functions are divided into two main areas:

- Compare costs—for those who have costs to compare with others for the purposes of benchmarking and assessment
- Understand costs—for those who would like to find out more about digital curation costs

Figure 2—The primary functions of the Curation Costs Exchange (CCEx) platform

3.1 'Compare Costs'

The Cost Comparison Tool[20] (CCT) or 'Compare Costs' sits at the heart of the CCEx and enables the exchange of potentially sensitive cost data. In doing so it provides users with the opportunity to identify greater efficiencies, better practices and to communicate with their peers.

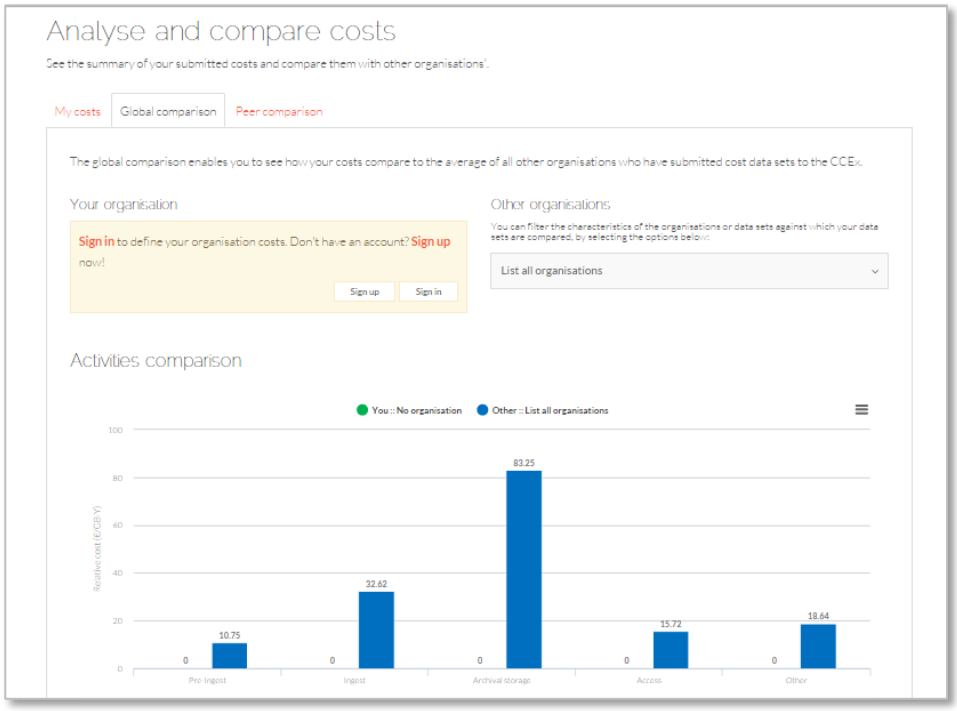


Figure 3 The Cost Comparison Tool landing page within the CCEx platform

[20] http://www.curationexchange.org/compare-costs

Key concepts that help those investing in digital curation compare their costs with other CCEx users and similar organisations include:

- Organisational Context
- Resources
- Services/Activities
- Digital assets

These core concepts provide the basis upon which the CCEx Cost Comparison Tool presents its cost comparison. They also provide a framework of reference for cost concepts models as well as methods and tools for repository managers, cost model theorists and tool developers.

When sharing cost data through the CCEx Cost Comparison Tool, users are asked to describe their organisation's 'stakeholder context' (although they may instead choose to remain anonymous) by providing profile information. This is followed by self-identifying costs as either 'resources' or 'service/activities.' Mapping information onto these categories helps the CCEx Cost Comparison Tool provide comparisons with other CCEx users and similar organisations.

Organisational Context

Users are asked to describe their organisation in terms of:

- Type
- Purpose
- Mission
- Location
- currency used

For each data set submitted, users are also asked to describe their collections in terms of:

- the scope to which the data pertains (such as whole organisation, department, project, or collection)
- curation staff (associated with the curation of the dataset)
- data volume
- number of copies policy
- asset types

Again, this information helps the CCEx Cost Comparison Tool compare the organisation's data with that of similar organisations.

Resources

Users are asked to indicate whether they break their costs down by purchases:

- Hardware—Machines and media used throughout the whole digital asset lifecycle. Hardware may receive, store, validate, make copies, migrate and disseminate digital assets.
- Software–Programs used throughout the whole digital asset lifecycle. Software may receive, process, validate, create copies, migrate and disseminate digital assets.
- External or third party services—Costs spent to buy services from third party providers. Includes outsourcing, renting and leasing of hardware and software.
- Overhead—All costs pertaining to overhead costs such as building costs, electricity, water, etc.

And human resource costs:

- Producer—Any individual involved in creating digital content. This may include for example researchers generating and managing digital research data or aggregating new subsets of existing data for new analysis; government employees undertaking data collection and analysis. Producers may also include software developers and vendors who are producing code to enable analysis, manipulation and visualisation of digital content.
- IT developer—Staff members who develop software. Software engineers, programmers, system developers, coders.
- Support/operations—Staff members who execute technical tasks, for example testing digital material at ingest, operating the computers when migration occurs, burning optical disks, setting up robots, and so on.
- Preservation analyst—Staff members who execute the preservation planning of the managers; archivists who appraise digital assets, consult at access, execute administrative tasks.
- Manager—Staff members who organise and plan the work of digital curation in their organisation. Make tactical and strategic decisions, have staff responsibility and do budgeting.

Services/Activities

Users may also indicate that they break their costs down by services and activities:

- Pre-ingest—Pre-ingest involves any activity related to the preparation of digital assets for archiving. This might encompass digitisation, extraction of data from databases, metadata enrichment, migration of production formats to preservation formats, etc.
- Ingest—This activity covers processes related to receiving digital assets from an external source and preparing them for storage. Examples of activities that could fit into this activity category are: appraisal, submission agreement, validation of digital assets, metadata enrichment, preparing digital assets for storage within the archive.
- Archival storage—This activity covers processes related to storing, maintaining and retrieving the digital assets. Examples of activities that could fit into this activity category are: error checking, media migration, storage hierarchy management, providing disaster recovery capabilities.
- Access—This activity covers processes related to accessing the stored digital assets.

Digital Assets

Users are asked to indicate the type, size and volume of the digital assets so that these factors may also be considered for comparison.

The tool addresses an acknowledged reluctance to share this data by anonymising and amalgamating data into a user average for the benefit of comparison. Only if an organisation wishes to share its identity will the CCEx reveal these details.

For a full description of the methodology and functionality of the Cost Comparison Tool (CCT), please refer to D3.3 'Curation Costs Exchange Framework.'[21]

[21] http://www.4cproject.eu/d3-3-curation-costs-exchange-framework

3.2 'Understand costs'

Acknowledging that some organisations are not in a position to share digital curation costs[22], the CCEx also draws together a wealth of information designed to help users understand more about digital curation costs.

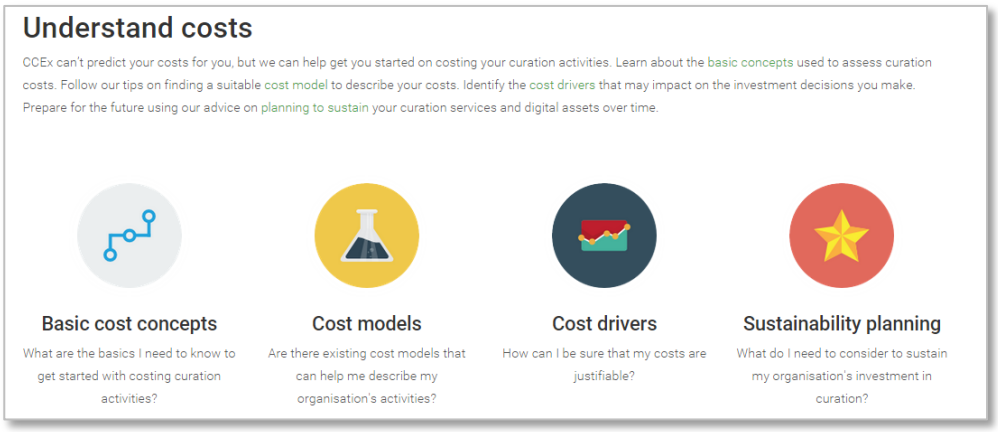


Figure 4—4C Project deliverables accessible through the 'Understand your Costs' section of the CCEx platform

Whether organisations are planning to build a business case for investments in digital curation, carry out a cost appraisal or analysis, or want to know about cost modelling, the 'Understand Costs'[23] section provides toolkit of resources drawn from the 4C project which aim to provide a starting point for achieving clarity about curation cost issues through the digital object lifecycle.

There are four main sections in this part of the site:

- Basic cost concepts
- Cost models
- Cost drivers
- Sustainability planning

They provide the following resources:

a) **Basic cost concepts**
This resource helps users familiarise themselves with key concepts for consideration when scoping a costing exercise. This is based on Deliverable D3.2 Cost Concept Model and Gateway Specification.[24]

[22] in some case this is because they are not yet mature enough—they have no costs to share—and in other cases they may be unable to share because their organisations policy prohibits them from doing so.

[23] http://www.curationexchange.org/understand-your-costs

[24] http://www.4cproject.eu/d3-2-ccm-grs

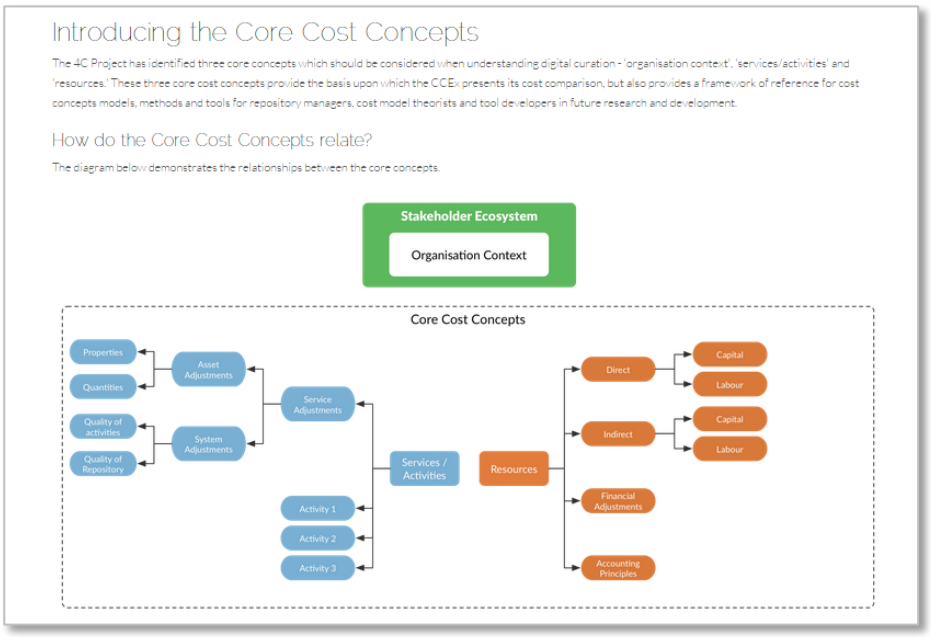


Figure 5—'Core Cost Concepts' introduces the key themes for consideration in understanding digital curation costs

b) **Cost models**

Drawing from the 'Evaluation of Cost Models and Needs & Gaps Analysis' [25]deliverable and providing an overview of selected cost and benefit models, these resources help users to breakdown and describe their costs using existing cost models, or to develop their own cost model using a generic gateway specification.

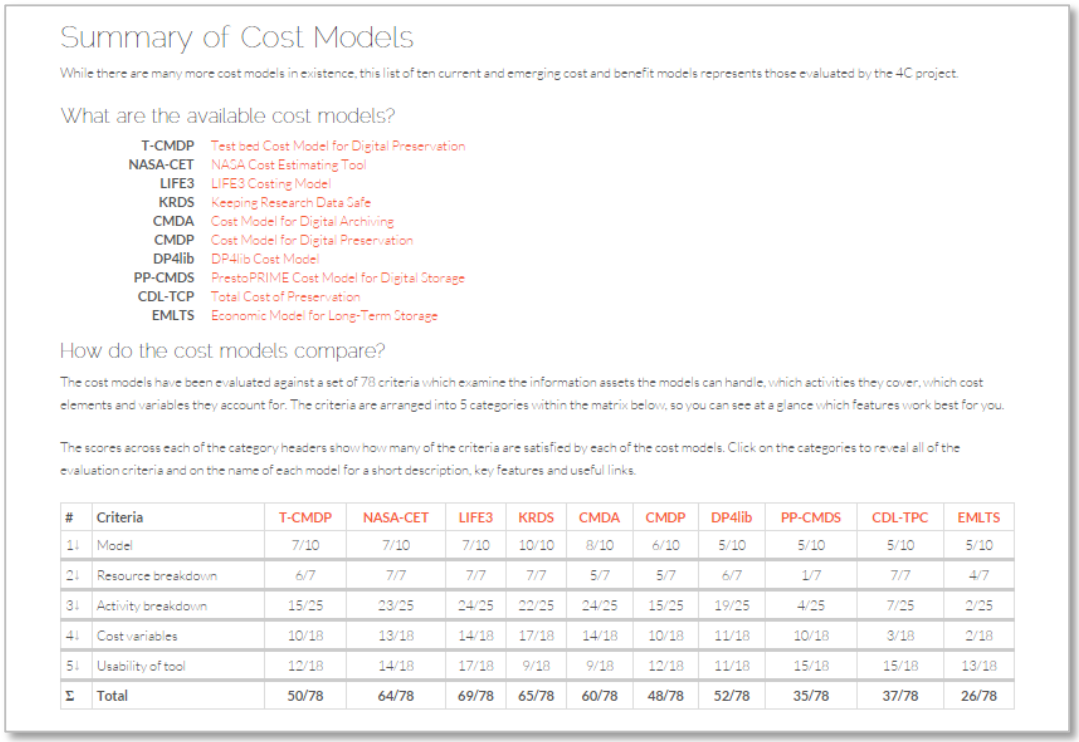


Figure 6—The 'Summary of Cost Models' collates the evaluated cost models into an 'at a glance' view to help users make their own assessments of available resources

[25] http://www.4cproject.eu/d3-1

c) Cost Drivers

Having a breakdown of your costs is very useful, but to really understand whether your spending is appropriate, you also need to contextualise your costs. Linked to Cost concepts, Cost drivers pinpoints a set of values significant in the practice of digital curation today and includes information about costs, benefits and risk.

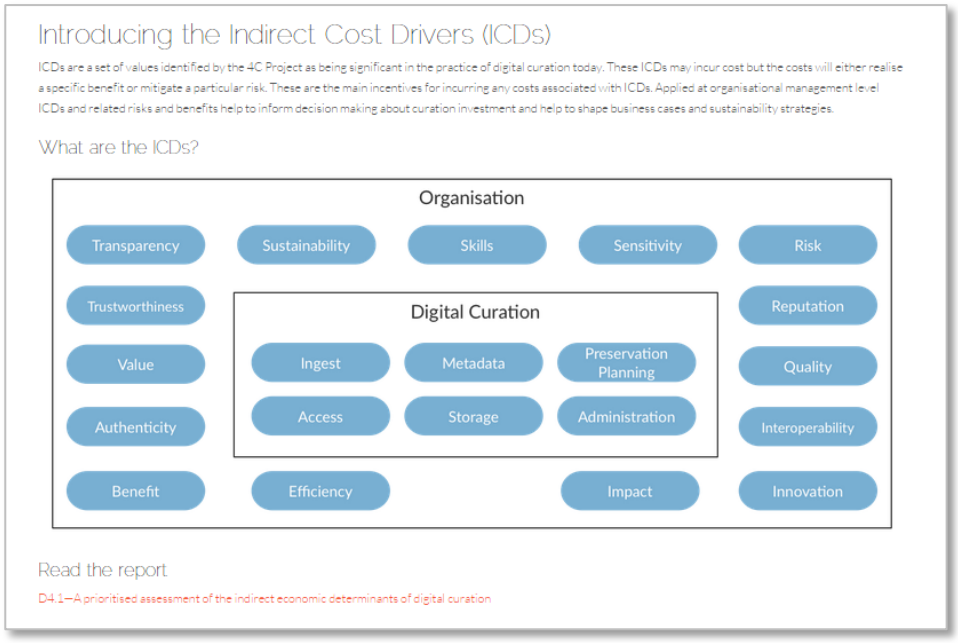


Figure 7—'Indirect Cost Drivers' introduces a set of values significant in understanding digital curation costs

d) Sustainability Planning

This resource helps users develop a business model that reflects their organisation's mission and strategic aims, as well as plan for the long term. The Digital Curation Sustainability Model[26] highlights key digital curation concepts, relationships, and decision points in a complex problem space. This helps users to benchmark and compare their own local models and invest strategically to preserve data for the long term, and also consider the costs and benefits of certification.

[26] http://www.curationexchange.org/make-the-case/15-economic-sustainability-reference-model

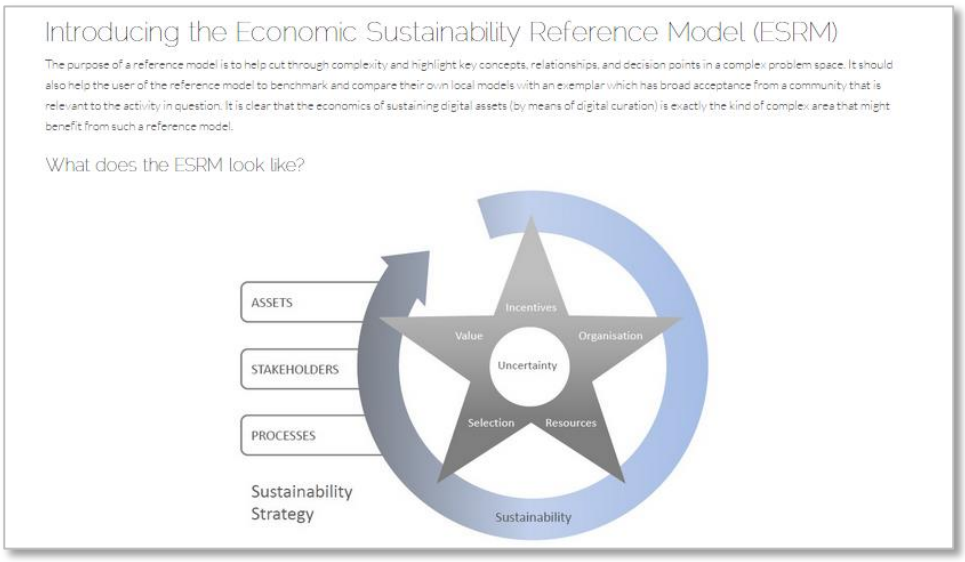


Figure 8—The 'Digital Curation Sustainability Model' highlights key concepts, relationships and decision points in the process of investing in digital curation

3.3 'Read more' and 'Curation services'

'Read more' draws from the deliverable D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives[27] and the bibliography—a registry of literature and articles—generated by the 4C Project.

The registry is by no means exhaustive and is intended simply to provide those interested in the cost of digital curation with a starting point. It is intended that this be built on with updates and new additions by the community, for future development and research.

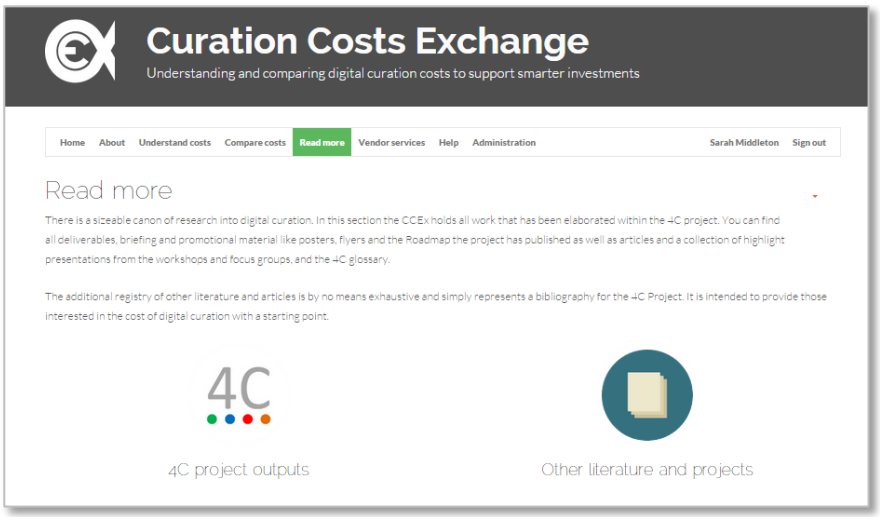


Figure 9—'Read More' provides a starting point for those interested in learning more about the costs of digital curation

[27] http://www.4cproject.eu/d2-1-stakeholders

'Curation Services'[28] provides a link to the Community Owned digital Preservation Tool Registry (COPTR)[29] which contains descriptions of a wide range of tools useful for long term digital curation and collates the knowledge of the digital curation community on preservation tools in one place.

This section also presents a list of some of the systems and service providers—both commercial, open source and hybrid—that are currently available.

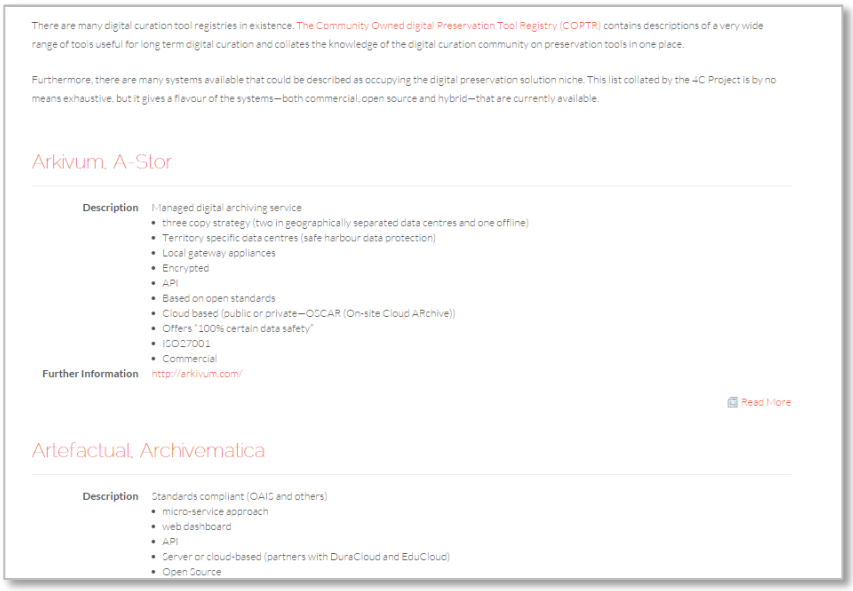


Figure 10—'Curation Services' provides links to the COPTR tools registry as well as a list of some of the available service providers

3.4 Supporting features

Supporting features of the CCEx are provided to assist users making their way around the platform and, in particular, using the Cost Comparison Tool.

3.4.1 About

The 'About' page is designed to communicate with three kinds of visitor to the CCEx: First time visitors, regular users, and people who want to help with the further development the CCEx. The page aims to provide appropriate information for these groups answering the following questions:

First Time Visitors

- How (often) is content updated?
- How can I register?
- How can I contact you if what I've read here isn't enough?

Regular/registered Users

- How can I engage with/contribute data to the CCEx?
- How can I stay up to date with developments?
- How can I contact you?

[28] http://www.curationexchange.org/vendor-services

[29] http://coptr.digipres.org/Category:Function

People who want to develop the CCEx functionality or use aspects of it in their own work

- Is the CCEx available for adoption/development/investment?
- How can I list my services/tool on the CCEx?
- What's the website traffic/usage like?

3.4.2 Step By Step Tutorial and Video

One of the themes that emerged from early feedback was that using the Cost Comparison Tool (CCT) might be a complex process, especially for those not familiar with "standard" curation or accounting terminology. The 'Compare Costs' landing page was developed to address this. It provides a step-by-step tutorial and video. The video is designed to articulate the benefits of using the CCEx, and to encourage users to sign up. The tutorial talks users through each stage of the 'Compare Costs' process, eliminating confusion and providing helpful hints for every part of the workflow.

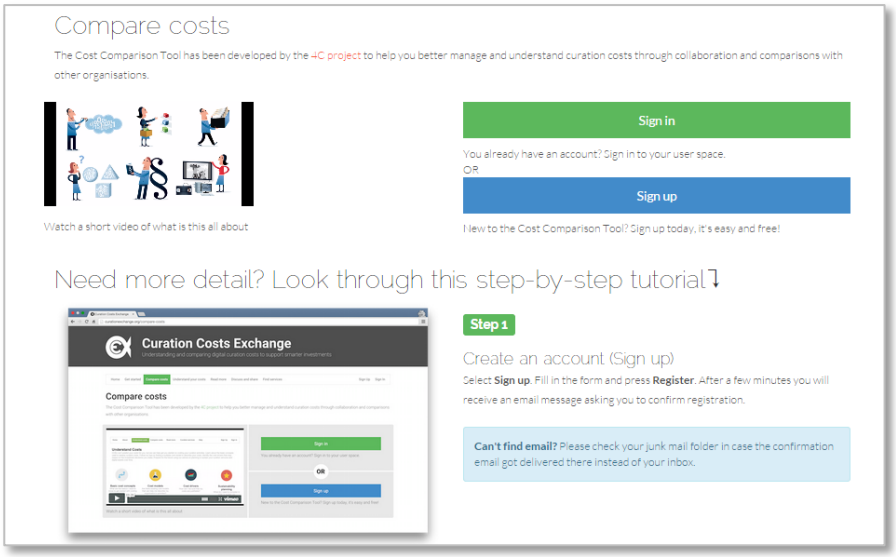


Figure 11—A video and step by step tutorial help users to understand the CCT workflow

Further support is provided through a 'bootstrap' tour on every page of the CCT workflow in pop-up bubbles alongside each cell requiring user input. This may be switched off once the user is confident with the process.

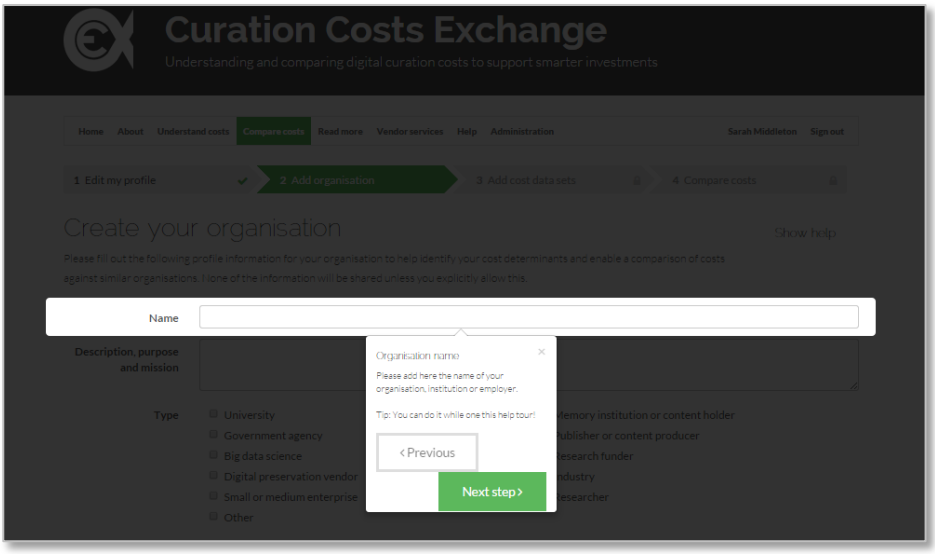


Figure 12—Bootstrap tours guide users through every step of the cost submission process, on each page of the workflow

3.4.3 Help

Some of the other issues from the feedback could not be incorporated into the platform within the project timeframe. These have been addressed by providing a 'Frequently Asked Questions' (FAQs) page within the 'Help' section. This page aims to address questions around the approach of the CCT, providing context and definitions to some of the terminology used. Further answers to questions specific to the handling of data are provided in the CCEx Privacy Policy.

4 The future of the CCEx

As it stands, the platform provides an exchange framework for cost data and cost information as required under the terms of the Description of Work.

4.1 Future developments

In the course of developing the CCEx, the 4C team identified a number of future developments which would enhance the functionality of the platform. However, these developments did not represent 'priority' features and fell outside the initial scope and programme of the project.

Nevertheless, it is hoped that with community adoption further development may be possible in the future, and the team has highlighted the some potential developments in this linked spreadsheet[30].

The most important suggestions for future development of the curationexchange.org are highlighted below. These suggestions are intended as an inspiration for the future care-takers/developers of the website. They are derived partly from the ideas of the 4C consortium partners, and partly from the user feedback received during numerous workshops, conferences, focus groups, and so on.

The future developments revolve around the following topics: Enhanced precision and comparability of the output of the Cost Comparison Tool (CCT); Enhanced integration between the CCEx, the CCT and other tools; More user interaction with the CCEx and enhanced engagement.

4.1.1 Enhanced precision and comparability

The nature of the parameters—and number of parameters[31]—that are used in the CCT are important to enable precise calculation of costs. They are also important in order to increase the understanding of (and refine the comparison of) costs. It is thus essential to know if the data you are curating is for example complex or simple, big or small, frequently or rarely used; if your organisation is large or small; if the staffing number is high or low, and so on. To address this the suggestion is to:

a. Revisit the level of granularity and include more nuanced cost determinants, some of which can:
    i. Offer the user the possibility of submitting different data volumes depending on which activity is being costed (Pre-Ingest, Ingest, Archival Storage, Access). As of now, only one cross-activity data volume can be submitted, but the reality is that organisations curate different volumes depending on which part of the curation lifecycle is addressed.
    ii. Increase number of currencies[32]. Being obliged to do several conversions decreases precision and makes the cost data submission more tedious.
    iii. Nuance the submitted information by offering the possibility of commenting each input (textual fields) and making these comments public (if the user has chosen to make their data public).

[30] https://docs.google.com/spreadsheets/d/1lCh9clTDX2G_-4tD582j_yb5WSl0dtUfC128YqLCpcY/edit#gid=0

[31] The counter to this being that the more cost determinants you integrate into the CCT in order to enhance the precision and comparability of the outcome, the less the users are likely to upload their cost data, because the submission process becomes too tedious.

[32] At the moment, the CCT accepts GBP (£), USD ($) and EUR (€). Adding currencies can be done relatively easily on the Administration page.

b. Display the entirety of the information that organisations have been willing to submit in the attempt to nuance their cost landscape, and make it available even though it is neither quantifiable nor suited for automatically generated comparisons.
c. Detect 'bad entries'. If someone submits gibberish, it should be detectable and eliminated (semi-)automatically so the authentic cost data won't get polluted.
d. Enable manual peer comparisons instead of automatic ones. At the moment the results page automatically compares your organisation with another which is supposedly similar to yours.

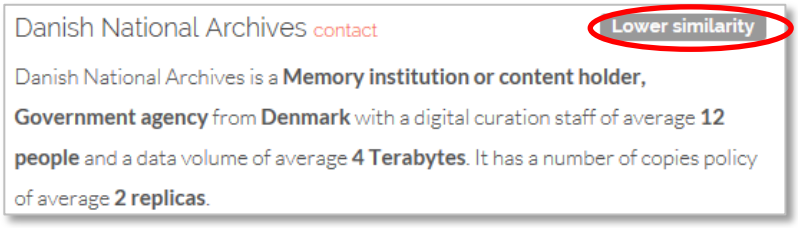


Figure 13—The CCT automatically compares your organisation with other similar organisations

This similarity feature compares a set of pre-chosen parameters defined by the 4C-project. If the feature were manual, each organisation could decide which other organisation it would like to compare costs with, thus enhancing possibilities of actual similarity.

e. Enable other types of cost calculations. The same results page display the costs in Euros per Gigabyte per year (€/GB/Y) only. This way of calculating costs can lead to the drawing of false conclusions—what if, for example, the data is very complex, but has low volume? Then the costs displayed will show a very a high cost for a very low data volume. Examples of other ways of calculating costs could include: Per asset type[33]; per number of data objects.
f. Capture funding sources within the CCT in order to establish an overview of different types of funding and to get a better insight into the potential business models of digital curation[34].

4.1.2 Enhanced integration between the CCEx, the CCT and other tools

The CCEx represents one discipline (the economics of digital curation) in a larger domain (digital curation). As such it is, in theory, possible to integrate the former into the latter, and if you want a better understanding of the domain, then this is what should be done. In practice, it is already possible to do this with the CCEx. During the lifetime of the 4C-project there have been concrete suggestions for how to enhance this. These include:

a. Establish a concrete integration between the CCT and the other, internal resources of the CCEx (see D3.3—Curation Costs Exchange Framework[35]). An example of this would be to enable users to tag themselves interactively with the cost drivers[36] that are the most important to them, enriching their profile and informing others.
b. Establish a concrete integration between the CCT, the CCEx and external tools:
    i. Between the CCEx Curation services page[37] and COPTR[38] to be able to maintain a central and up-to-date tool register.

[33] For example.: Unformatted text; word processing; spreadsheet; sound; graphics; geodata.
[34] Simon Hodson, CODATA
[35] http://www.4cproject.eu/d3-3-curation-costs-exchange-framework
[36] http://www.curationexchange.org/understand-your-costs/cost-drivers
[37] http://www.curationexchange.org/curation-services
[38] http://coptr.digipres.org/Category:Function

ii. Between CCT and research grant application processes[39] to enable useful estimation of research project costs.
iii. Between CCEx and educational material and tools, developing a curriculum that would educate and move organisations forward and even an online course in preservation for students[40].

4.1.3 More user interaction with the CCEx and enhanced engagement

It has proven difficult to get users to submit their cost data to the CCEx for a number of reasons. Explanations often cited are the sensitivity of financial information and lack of central supporting services to breakout detailed information for the different stakeholder groups within an organisation. Added features and services that address this and which may remedy the reluctance and spur more user engagement include the following:

a. Enable cost predictive features, for example make room for input of retention periods and accession plans. This would also enable the CCT to address, for example, emerging data archives in third world countries and the modelling of their future costs. This in turn would help them immensely in the process of establishing themselves as functioning digital curation organisations[41].
b. Create hosted instances of the CCT to allow people to use the tool within their organisations[42]. This could prove to be very useful for example for membership organisations that want to keep their costs to themselves, but don't have the expertise to create and/or install a similar tool. The Vendors Focus Group (described in D2.3 Final Stakeholder Report) discussed the idea of using the CCT in 'local instances' to standardise and track costs over time. This idea was well received.
c. Enable anonymised re-mailing so that organisations wanting to learn more about a result they see on the result's page can easily get in touch with their peers who can remain anonymous if wished.
d. Create a downloadable spreadsheet with all the information that the CCT is going to ask from a user, allowing for easier retrieval of financial information, which usually is acquired from different sources and people within the organisation.
e. Tuned site search[43].

[39] Ron Dekker, Netherlands Organisation for Scientific Research

[40] Kate Wittenberg, Portico. The APA virtual Centre of Excellence is also relevant in this context: http://www.alliancepermanentaccess.org/index.php/community/event/apaconferences/apa-conference-oct-2014/

[41] Simon Hodson, CODATA

[42] The source code and installation instructions are already available on GitHub for those who might wish to install their own instance and/or contribute to future developments—http://my.curationexchange.org/

[43] The underlying CMS already allows for generic site search, however, the results returned can often unhelpful without tuning of the underlying search rules.

5 Community Building and Outreach

The Curation Costs Exchange (CCEx) is intended to be an online, virtual community platform for the exchange of curation cost information.

Whilst consolidating many of the 4C deliverables, the CCEx has been developed as an 'independent brand,' designed to outlast the life of the project and go on to be owned, cultivated and driven by the very user community who employ it. It is therefore entirely reliant on continuous user engagement and feedback.

5.1 Activities

Community building and outreach efforts from within the 4C Project have focused on encouraging users to share their data through the Cost Comparison Tool (please see Annex to D2.5: 4C Project Communications Plan[44]), with a summary of activities below:

Push strategy

1. Personal invitation and direct contact through email/ phone
2. Exhibitions, workshops, focus groups and demonstrations
3. Offer of Incentives

Pull strategy

1. Social media

- Twitter
- Linked in

2. PR

- Mailing lists
- Website news releases
- Publications

5.2 Results

Since the beta launch in August 2014, the CCEx has been well received by the digital curation community with:

- 4,381 hits
- 173 registrations

Whilest 40 organisations have submitted their data to the CCEx, there remains a reluctance to share this data publicly and it is this aspect of user engagement which the 4C Project has tried to address in particular.

Feedback throughout the initial consultation, beta testing and at iPRES in October all indicated a need for and an desire for a platform such as the CCEx, as well as a willingness to use it and become involved in the associated community. Some digital curation practitioners at iPRES even asked whether the CCEx could be expended to incorporate analogue curation, such was its warm reception.

[44] http://www.4cproject.eu/d2-5-project-communication-plan

The benefits of using the CCEx's Cost comparison tool have also been recognised by high level funders, like Ron Dekker of the Netherlands Organisation for Scientific Research (NWO) who said,

> *"The CCEx is* **the** *platform to 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."*

To summarise the majority of comments from those who have been introduced to the CCEx: they are interested—if not keen—to use the CCEx, but would only feel really be *compelled* to use it when it holds more cost data sets for comparison.

6 Sustainability of the CCEx

The CCEx has been designed to outlast the life of the 4C project and go on to be owned, cultivated and driven by the very user community who employ it. It is therefore entirely reliant on continuous user engagement.

Furthermore, the CCEx will rely on a number of representatives from within the project and the project's community to form a 'Post-project consortium' which will review and manage the administration, content and relevance of the CCEx.

The following organisations from with (and outside) the original project consortium have already informally agreed that they will assist in the sustainability of the CCEx:

- Digital Preservation Coalition (DPC)
- nestor
- NCDD (Netherlands Coalition for Digital Preservation)
- KEEP Solutions
- Cyber Media
- UK Data Service

6.1 Roles and responsibilities

Content Management and administration

- DPC to manage the 'day to day' running of the content management system [estimated contribution—2.5 days per month]

Content Refreshing/ Updating:

- Estimated contribution—nestor to provide 2 hours every month to check links and currency of content. [Or alternative: 4 hours every 2 months]
- nestor to work with NCDD on an annual basis to review currency and relevancy of CCEx content (estimated contribution—1 person day = 8 hours)

Technical Support:

- DPC's web host and technical service provider Cyber Media will support the standard Joomla! components within the CCEx. They have observed the other, custom developments which will sit in an MVC/joomla compatible framework, so are familiar and comfortable in supporting these elements once they are handed over.
- Ongoing support from Cyber Media sits easily alongside the arrangements already in place for DPC's own web support and will include bug fixing and standard enhancements.
- If for any reason Cyber Media are unable to support the custom developments, KEEP Solutions have offered to contribute to a shared support function whereby Cyber Media take care of the standard applications and KEEPS manage the custom developments.

CCT data:

- As discovered through 4C's evaluation of existing cost models, there is a lack of suitable costs data available to service providers and cost model developers for testing and refining cost and business models. To help improve on this situation, the 4C project will harvest and clean the cost data submitted via the CCT up to the end of January 30, 2015. This is in line

with the CCEx terms and conditions which clearly states that users agree to allow 'snapshots of anonymised cost data to be collected periodically'. The anonymised data will be fully described, assigned an identifier and deposited with the UK Data Service for future use as sample data for curation costs related research and innovation. The 4C Consortium will harvest and clean any additional data collected by the CCT for a period of 12 months following the end of the project in the first instance. The data will be described and assigned an identifier and linked to the initial CCT data set. A longer term decision on collecting data submitted via the CCT will be made by the 4C Consortium in late 2015.

The first review of the CCEx will assess the currency and relevance of not just of the platform's content, but of the platform as a whole. This will determine actions for the following year(s).

7 Conclusion

The value that may be derived from the CCEx relies heavily on the willingness of organisations to submit their cost data. This in turn relies on their understanding of the benefits that sharing will bring about. The more costs that are shared, the more we can all learn about making smarter investments in digital curation.

The digital curation community recognises the potential of the CCEx, most particularly the Cost Comparison Tool, to demystify the costs of digital curation and to bring about greater cost efficiencies all round. The degree of success of the CCEx platform will thus be measured in the readiness of stakeholders to share sensitive cost data and by the Cost Comparison Tool's adequacy and pertinence: Does it work properly? Is it intuitive? Are the results interesting? Are the results useful?

The 4C Project has provided a tool that is fit for purpose and pertinent for a wide range of stakeholders through a high degree of user involvement in the development of the tool, and by performing an in-depth examination of user behaviour around the costing of digital curation. Hence the CCEx is a product of the user requirements that the project has identified through its deliverables, user consultations and iterations in the form of workshops, webinars, focus groups, beta-testing, individual usability tests, Advisory Board consultations, internal meetings, and conference sessions.

However, two key challenges of the CCEx remain: to gather more cost data, and to ensure these translate into figures that organisations of all kinds find meaningful and engaging. The meaningfulness of these figures does not just depend on *how* results are presented, but on *how many* cost data sets the CCEx is able to attract in order to generate statistically reliable and useable figures. The meaningfulness of the results *is* the compelling argument for the submission of new cost data sets.

In short, the CCEx needs further and ever increasing community support, it needs digital curators, creators, vendors and funders to use it, add to it, take care of it—and above all share their data with it to make it a sustainable community resource.

Appendix 1 CCEx Cost Comparison Tool, Usability Testing Report

A1.1 Executive summary

This report presents the results from the evaluation of the Cost Comparison Tool (CCT) developed for the Curation Costs Exchange (CCEx) platform, as part of the 4C[45] project funded by the European Union.

In November-December 2014 user feedback was collected via a focus group and usability tests.

The results of the tests were overall positive and showed that the tool is simple to use, as reflected in high task completion rates and low number of errors made by users. The majority of users (60%) agreed that the various functions in the tool were well integrated, found CCT intuitive and consistent, and thought that most people would learn to use it very quickly.

However the tool scored rather low (Grade D) in the post-test questionnaire, where users were asked to rate the usability of the tool (see section 5 under Results). None of the participants stated they would like to use the tool frequently, and only one user thought the tool was easy to use and could be used without some learning effort.

The biggest challenge is how to effectively communicate the purpose of the tool. None of the users surveyed clearly understood what the tool does. This was true of both participants to usability tests and participants to the focus group, and this is very worrying since the focus group consisted of the tool's target audience.

The CCEx website must be revised to provide a clear description of what the tool does and how it does it. Information on how much time it takes to enter data, how much data is required, how comparisons are made must be apparent from the moment user lands on the website. In the words of one of the participants: "I think that's the value of the service, to be able to compare notes, to engage with the rest of the community…what do I need to do to make a cost saving. To clearly articulate that on the homepage is critical".

Users frequently complained that web pages were too wordy and full of jargon. Filling in key forms such as add/edit cost data set, add/edit cost unit is very time consuming. The similarity in look and content between the two sets of forms causes users to lose track of the process.

A list issues highlighted by users and suggestions for improvement are summarised in Annex 1: Recommendations for Improvement.

Some of the problems highlighted could be easily solved with a thoughtful redesign of web pages, while others will require more profound changes to the workflow. This is challenging given the short time remaining before the project ends.

[45] Collaboration to Clarify the Costs of Curation, http://4cproject.eu/

A1.2 Overview of methodology

Five people who had never seen nor used the tool were recruited for usability tests. Test users had to complete six key tasks, under observation, and give their opinions on the user interface as well as the look and feel of the website. Users were tested individually. Each session lasted approximately one hour.

Participants were asked to rate the expected difficulty of the task on a 7-point scale with measurements ranging from 'Very difficult' to 'Very easy', by answering the following questions:

- Before carrying out each task (expectation rating): "How difficult or easy do you expect this task to be?"
- After completing the task (experience rating): "How difficult or easy did you find this task to be?"

In addition to completing the tasks, participants were also asked to subjectively assess the usability of the tool using a standardised Software Usability Scale (SUS) questionnaire on a 5-point Scale with endpoints ranging from 'Strongly disagree' (1) to 'Strongly agree' (5). Post-task scenario subjective measures included:

1. I think that I would like to use CCT frequently
2. I found CCT to be simple
3. I thought CCT was easy to use
4. I think that I could use CCT without the support of anyone else (my colleagues, support staff such as librarians, IT staff, etc.)
5. I found the various functions in CCT were well integrated
6. I thought there was a lot of consistency in CCT
7. I would imagine that most people would learn to use CCT very quickly
8. I found CCT very intuitive
9. I felt very confident using CCT
10. I could use CCT without having to learn anything new

To assess how familiar users were with online budgeting/costing tools, they were asked to answer the following questions:

- Are you responsible for managing budgets in your current role? If yes, how much of this information comes from other people within the organisation?
- How big is the budget that you manage?
- Have you ever used an online tool to help plan and manage costs in your organisation?

Preparation: Three Stages

1. Critical tasks that users must be able to perform using CCT were identified.
2. Scenarios were created for each task.
3. Each scenario came with a set of testing objectives and definition of what success in task completion looks like (see 'Metrics recorded' section below).

About the test users

Five participants were tested in one week, three males and two females. All participants worked at the University of Edinburgh, see User Testing Table 1 below. All had been responsible for managing budgets in their current or previous roles.

Book keeper	School of Informatics
Research officer (previously, Tissue culture facility manager)	Digital Curation Centre/Institute of Stem Cell Research
Project manager	Information Services
Data librarian/RDM services coordinator	Library & University Collections
Senior business development executive	School of Informatics/The Data Lab

User Testing Table 1—Participant details

Tasks

Test participants attempted completion of the following tasks:

1. Find the CCT tool.
2. Sign up for CCT.
3. Enter information about your dataset and the costs associated with storing it. Save it.
4. Compare costs with peers at California Digital Library.
5. Please contact support for the website.
6. Go back and amend your data set.

In addition to carrying out these tasks, users were asked to describe the Home page, to explain what they saw and what they thought they could do on the web site. The purpose of this test was to check whether the design of the page had introduced any usability issues, and if it had could the users still figure out how the web site worked?

Biases to data

Bias in usability testing was taken into consideration when analysing the results:

- Hawthorne effect: the act of observing or studying someone can alter their behaviour
- Task-Selection Bias: if we ask users to find something or attempt to accommodate something, they assume it is possible. Outside the usability test, users have their own goals and don't know if a function is even possible
- Social Desirability: users generally tell what they think facilitators want to hear and are less likely to say disparaging things about products. This often means they would blame themselves rather than the product if they encounter a problem. It can also impact preferences; most users would say yes if asked if they would use the product.

- Note Taking: users become aware the facilitator is taking notes and observing their behaviour may become more self-conscious, less confident, about the actions they are taking

Metrics recorded

- Completion Rates—collected as binary measure of task success (coded as 1) or task failure (coded as 0)
- Usability Problems—a list of problems encountered by a user with a description and severity rating (prevents task completion/causes a significant delay or frustration/has relatively minor effect on task performance/is a suggestion)
- Task Time—how long user spends on activity, the amount of time user devotes to activity until it completes task (correctly or incorrectly) or gives up
- Errors—any unintended action, slip, mistake or omission. Where possible these were mapped to usability problems.
- Satisfaction Ratings—after the test, users were asked to complete a standardised usability questionnaire

A1.3 Results

1. Task Completion Rate

All participants successfully completed Task 2 (sign up for CCT), Task 5 (contact support for the website) and Task 6 (go back and amend your dataset). Four of the five (80%) completed Task 1 (find the CCT tool on the website). The lowest completion rate 60% (but still more than half) was for Task 3, which asked users to enter information about a dataset and the costs associated with storing it (the core task in CCEx).

Participant	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	1	1	0	1	1	1
4	0	1	0	1	1	1
5	1	1	1	0	1	1
Success	4	5	3	4	5	5
Completion Rates	80%	100%	60%	80%	100%	100%

User Testing Table 2—Task completion rate

2. Usability Problems

The usability problems encountered by users are listed below. See User Testing Table 3 for the frequency with which they occurred and for an impact score. Impact score was calculated by combining four levels of impact (4 = prevents task completion, 3 = causes a significant delay or frustration, 2 = has a relatively minor effect on task performance, 1 = is a suggestion) with four levels of frequency (4 = frequency >90%; 3 = 51-89%, 2 = 11-50%; 1 = <10%). Problems with the highest impact score are marked in red.

Problem 1: 'Related services' category misunderstood for Help/report a problem.

Problem 2: Users struggle to find the tool on the homepage. They take various paths to find it: look in 'Understand your cost'; click on ESRM model and Indirect cost drivers; look in 'Get started' which seems like a better option but 'Compare costs' are featured more prominently; go back and forth, not entirely sure whether they found it.

Problem 3: 'Create an account' on 'Get started' pages still an option even for registered users.

Problem 4: Each form (edit organisation, add cost data, add cost unit) is asking for 'Name' and 'Description' without offering any help what exactly we mean by this. Other fields are also not clear, such as 'Scope', 'Number of copies', 'Curation staff' etc.

Problem 5: Users comment how 'fiddly' the slider is especially when entering shorter periods (do we anticipate projects for up to 50 years?)

Problem 6: Users don't understand that they are asked to allocate the same cost across asset types, activities mapping and purchases and staff.

Problem 7: Users can't see text very well on the page**.**

Problem 8: Users comment that 'Analyse and compare costs' page is too busy / overwhelming with text. They struggle to spot 'Compare with other peers' button.

Problem 9: 'Get started' has a rocket as an icon, while the text underneath it reads 'get acquainted with CCEx', a discrepancy between expectation and reality.

Problem 10: Users click on both 'Report a problem' and 'Contact' not understanding the difference between them. They also comment that it sends email to the same address.

Problem 11: Users expect FAQ/Help page, not just an email to generic info@4cproject.eu when looking for help. Email could be changed to support@4cproject.eu

Problem 12: Users expect to find contact details at the top right-hand side of the homepage.

Problem 13: **'**Edit' button in 'Manage cost data sets' is easy to miss; its positioning suggests it only allows users to change the name of the project; the button is also red in colour, which is usually reserved for errors.

Problem 14: Users expect an overview of the tool before deciding whether they want to use it or not.

Problem 15: Users want to see links to more information and explanations of the various options under 'Information sharing'. Users are particularly concerned about privacy and authority: *"Hmm… I'm thinking, one may need some authority to allow the organisation to be linked to the data I provide"* (also see results from focus groups).

Problem 16: Users find 'Information sharing' options full of jargon and struggle to understand what they mean.

Problem 17: Users complain about the jargon on the website (i.e. cost unit, pre-ingest, CCEx registry etc).

Problem 18: Users don't understand what it means to add a new cost unit. One says: *"I assume that my cost unit maps onto my project, it's the project cost unit or something but it seems very vague, this information is quite simple but if it was more complicated then I might start wondering how much detail I should be going into, how many cost units I should be putting in. It's not apparent"*.

Problem 19: Positioning of the 'Edit' button is the same (always next to a name) but its function is different depending on the page. On 'Compare costs' page it takes the user to 'Edit your organisation' but on 'Manage costs data sets' it allows to edit a cost data set. Users found it confusing, they automatically assumed it would only allow them to edit their organisation's details, and therefore ignored the link, resulting in longer time to complete the task.

Problem 20: Any changes made in cost data set (i.e. changing volume) automatically alter graphs etc. This may not be what users want: One user *comments "It has moved everything proportionally, it moved everything up by 50% I guess …now that would be massively annoying if I got nine of these things…to unpick all the other stuff…that would drive me mad…what did I put, how have I done that and it's going to be annoying and fiddly job"*.

Problem 21: User thinks Compare costs and Understand costs are tools and everything below is information.

Problem 22: Saving the form, unlike in most familiar applications (such as MS Word) alters the form.

Problem 23: Users still don't understand what the tool is about even half way through the tasks: *"I'm confused now I thought this was going to tell me how much this was going to cost"*; *"I think I've spent all my money already… why am I buying hardware from you guys?"*; . *"Perhaps I've misunderstood the whole thing"; "Already, I'm lost"*.

Problem 24: Users expect a search functionality to find desired peer organisation in the list.

Problem 25: Returning users expect their 'own' space on the website. In other words something g like "XY's things" or "YX's details".

Problem 26: Stage 2 (Add cost data set → Add cost unit) in the process is not obvious.

Problem 27: Users clicks on 'Analyse and compare costs' button instead of 'Add new cost data set'.

Problem 28: Users struggle to read chart. For example fail to see the other organisation's costs in the comparison chart.

Problem 29: System automatically creates a new data set when user uses the 'back' button on the browser to go back from 'Add cost unit' page to 'Add cost data' page.

Problem	User 1	User 2	User 3	User 4	User 5	Total	Proportion[46]	Impact Score[47]
1	X	X				2	0.4	4
2	X	X	X	X	X	5	1	16
3	X					1	0.2	6
4	X	X	X	X		4	0.8	9
5	X	X	X	X		4	0.8	3
6	X		X		X	3	0.6	12
7	X		X			2	0.4	2
8	X	X	X			3	0.6	9
9		X				1	0.2	2
10		X	X		X	3	0.6	3
11			X		X	2	0.4	2
12		X	X	X	X	4	0.8	6
13	X		X		X	3	0.6	9
14		X				1	0.2	8
15		X		X		2	0.4	6
16			X			1	0.2	6
17		X	X	X	X	4	0.8	9
18		X		X		2	0.4	8
19		X	X			2	0.4	6
20		X				1	0.2	6
21			X			1	0.2	6
22			X			1	0.2	6
23			X	X	X	3	0.6	12
24			X			1	0.2	2
25			X			1	0.2	2
26					X	1	0.2	8
27				X	X	2	0.4	4
28				X	X	2	0.4	6
29				X		1	0.2	6

User Testing Table 3—Usability problems with frequency and impact score

[46] Frequency of occurrence of a problem

[47] Impact Score is calculated by combining four levels of impact (4-prevents task completion, 3-causes a significant delay or frustration, 2-has a relatively minor effect on task performance, 1-is a suggestion) with four levels of frequency (4-frequency >90%; 3-51-89%, 2-11-50%; 1-<10%)

3. Time on Task

User Testing Table 4 shows the time spent on each task, irrespective of whether the user completed the task successfully or not. Some tasks were inherently more difficult and longer to complete than others, and this is reflected by the average time on task.

Task 3 required participants to enter information about their dataset and the costs associated with storing it, and took the longest time to complete (mean = 755 seconds, a bit over 12 minutes). However, completion times ranged from 268 seconds (approximately 4.5 minutes) to 1392 seconds (more than 23 minutes).

Tasks 2, which asked users to sign up for CCT was the second longest (mean = 176 seconds, approximately 3 minutes), closely followed by Task 4 to compare costs with peers at California Digital Library (mean = 145 second s, more than 2 minutes).

	User 1	User 2	User 3	User 4	User 5	Average Total
Task 1	158	70	122	120	30	100
Task 2	300	189	171	122	95	**176**
Task 3	680	882	1392	550	268	**755 (12min)**
Task 4	245	122	98	118	141	**145**
Task 5	58	54	39	30	61	49
Task 6	112	114	314	37	73	130

User Testing Table 4—Time on task

4. Errors

In this context an error means any unintended action, slips, mistakes or omissions made by participants made while trying to complete the task scenarios. User Testing Table 5 displays a summary of the test data. Tasks with low completion rates and high errors and time on tasks are highlighted in red. The most complex tasks, Task 3, received the highest number of errors.

Task	Task Completion	Errors	Time on Task
1	4	5	100
2	5	0	176
3	**3**	**12**	**755**
4	4	4	145
5	5	5	49
6	5	9	130

User Testing Table 5—Summary per task in terms of completion rates, errors and time

5. Post-test questionnaire

After each session participants were asked to rate the usability of the tool using SUS questionnaire on a 5-point Scale with endpoints of 'Strongly disagree' (1) and 'Strongly agree' (5).

Please note that, in general, users who complete more tasks tend to rate tasks as easier. Some users, however, fail tasks and still rate them as being easy, while others complete tasks quickly and rate them difficult. That is why collecting multiple metrics in a usability test is advantageous because it provides a better picture of the overall user experience than any single measure, or for that matter any single research method, can.

Overall participants didn't rate the usability of the tool very high. Not one participant would like to use the tool frequently and only 20% (1 participant) thought it was easy to use and could be used without having to learn anything new. Only one participant felt very confident using the tool. However, approximately half of participants (60%) agreed that the various functions in CCT were well integrated, found it intuitive and consistent, and would imagine that most people would learn to use it very quickly.

	Strongly disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly agree (5)	Mean rating	Percent agree[48]
I think that I would like to use CCT frequently	1	2	2	-	-	2.2	0
I found CCT to be simple	-	2	1	2	-	3.0	40%
I thought CCT was easy to use	1	-	3	1	-	2.8	20%
I thought that I could use CCT without the support of anyone else (colleagues / IT staff, etc.	1	-	1	3	-	3.2	60%
I found the various functions in CCT were well integrated	-	-	2	3	-	3.6	60%
I thought there was a lot of consistency in CCT	-	-	2	2	1	3.8	60%
I would imagine that most people would learn to use CCT very quickly	-	-	2	2	1	3.8	60%
I found CCT very intuitive	-	2	-	3	-	3.2	60%
I felt very confident using CCT	-	2	2	1	-	2.8	20%
I could use CCT without having to learn anything new	-	1	3	1	-	3.0	20%

User Testing Table 6—Results of post-test questionnaire

The overall SUS score for CCT is 53.5 (Grade D), which is poor. This was calculated by subtracting 1 from the raw item score for each participant and multiplying by 2.5. It is worth nothing that there was high

[48] Percent agree (%) = Agree & Strongly Agree Responses combined

variation in individual SUS scores, which ranged from 30 to 72.5.

What does this score mean? Based on data from 446 studies and over 5,000 individual SUS responses, the overall mean SUS score across software applications as a whole is 68. Usability Researcher Jeff Sauro uses percentiles to show how usable products are relative to other products, and to develop the curved grading scale shown in User Testing Table 7 (this allows us to compare an individual application's scores with others). A SUS score of 53.5 has a percentile range of 15-34%. This means that that a score of 53.5 is considered more usable than 15-34% of the products in the Sauro database and less usable than 66-85%.

Curved Grading Scale Interpretation of SUS scores		
SUS Score Range	**Grade**	**Percentile Range**
84.1-100	A+	96-100
80.8-84	A	90-95
78.9-80.7	A-	85-89
77.2-78.8	B+	80-84
74.1-77.1	B	70-79
72.6-74	B-	65-69
71.1-72.5	C+	60-64
65-71	C	41-59
62.7-64.9	C-	35-40
51.7-62.6	D	15-34
0-51.7	F	0-14

User Testing Table 7—SUS scores

6. Expectation Ratings

Participants were asked how difficult or easy they expect it to be both before (expectation ratings) and after (experience ratings) completing the task.

	Very difficult (1)	Rather difficult (2)	Difficult (3)	Neutral (4)	Easy (5)	Fairly easy (6)	Very easy (7)	Mean rating
Find the CCT tool	-	-	-	1		3	1	5.8
Sign up for CCT	-	-	-	-	-	4	1	6.2
Enter information about your dataset & the costs associated with storing it. Save it	-	-	-	1	3	1	-	5
Compare costs with peers at California Digital Library	-	-	-	-	3	1	1	5.6
Please contact support	-	-	-	1	1	-	3	6
Go back and amend your data set	-	-	-	1	2	1	1	5.4

User Testing Table 8—Expectation rating

	Very difficult (1)	Rather difficult (2)	Difficult (3)	Neutral (4)	Easy (5)	Fairly easy (6)	Very easy (7)	Mean rating
Find the CCT tool	-	1	-	1	2	-	1	4.6
Sign up for CCT	-	-	-	-	-	4	1	6.2
Enter information about your dataset & the costs associated with storing it. Save it	-	1	3	-	-	1	-	3.4
Compare costs with peers at California Digital Library	1	-	-	-	1	2	1	5
Please contact support	-	-	1	-	-	1	3	6
Go back and amend your data set	-	1	-	2	-	2	-	4.4

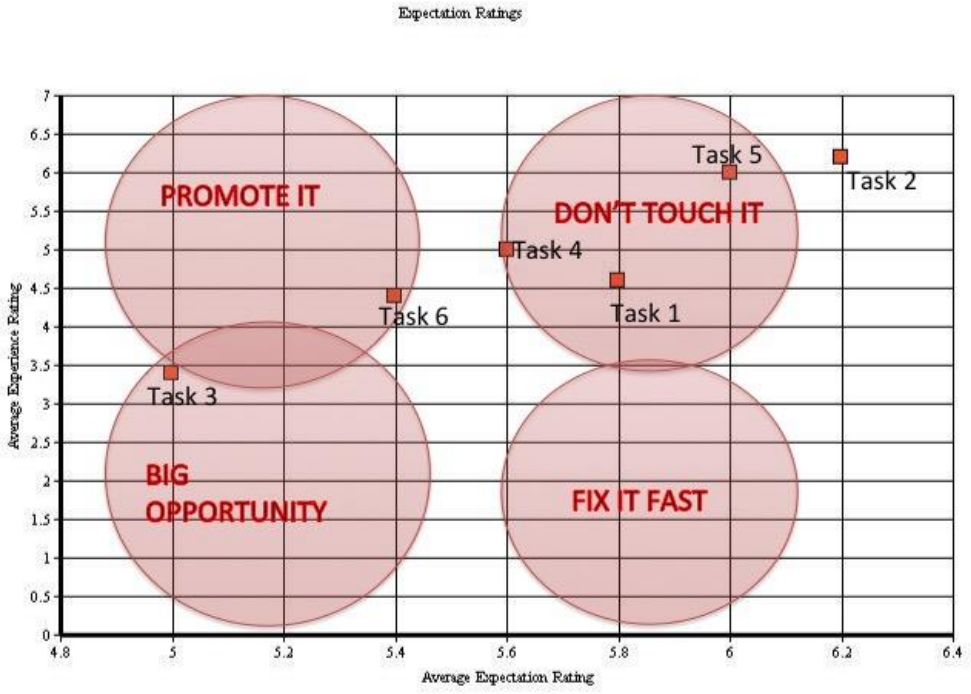
User Testing Table 9—Experience ratings

The results were used to create a scatterplot and mapped onto four quadrants (see User Testing Figure 1 below) as follows:

- Upper left (Promote it): These are the tasks that participants thought would be difficult but turned out to be easier than expected, so these are features that could be further developed and advertised as key strengths of the tool. Results show that Task 6 (Compare

costs with peers) is to one of strong points of the tool, which is great because it coincides with the core function of the tool.

- Lower left (Big opportunity): these are tasks which participants perceive as difficult before and after performing them so these are areas for improvement. As far as CCT is concerned, Task 3 (Enter information about your data set and the costs associated with storing it) is a function that must be improved.
- Upper right (Don't touch it): This quadrant contains the tasks perceived as easy before and after a task performance, so it's advisable to leave these areas alone. It's good to note that four tasks (Task 1, 2, 4 and 5) fell into this category.
- Lower right (Fix it fast): These are the tasks that participants though would be easy but turned out to be difficult—a potential source of user dissatisfaction—making this the quadrant to focus on. None of the features tested in CCT fell into this category.



User Testing Figure 1—Plan of action according to ratings

Participant Likes, Dislikes

While carrying out tasks participants would often volunteer comments on what they liked most/least about the tool, and make recommendations for improvement.

Liked Most

- The homepage *("It's nicely set out, the colours are good"*)

Liked Least

- Breaking down costs is complicated and open to interpretation
- Purpose is unclear: *"What would I use it for? I might want to compare costs out of curiosity but how much data am I going to put in? Is it going to be quite accurate data about my most recent project, will I be more likely to just throw something not junkie but kind of rough guess?"*
- Reliability of data: *"How much data do you need to have in there for people to be able to make reasonable comparison as there is not much in there, needs to be seeded."*

- Categories unclear: *"If there is some ambiguity about what's going where then there is going to be natural variation."*
- Return for effort: *"As a researcher how does is help me? Is it going to make my proposal more likely to succeed?"*
- Chart—one user struggled to spot the bar representing the other organisation on the chart
- One of users said he would like to see the expenditure model for the California Digital Library. It wasn't clear to him how the system made comparisons entities as varied as his own project and California Digital Library. The fact that the system doesn't explain the economies of scale to the user was a *"huge design flaw"*.
- Cluttered pages. One user suggests getting rid of *"second narrative"* by which he meant all the icons below Compare costs and Understand costs
- Homepage looks busy *("It's very difficult to absorb all that text"*) and doesn't explain the purpose of the web site *("Digital curation—it's a fairly nebulous sort of subject"*)
- Tool looked daunting to some users. They felt that before entering data, they would need a good grasp of the inner workings and the concepts behind the tool *("I couldn't imagine you doing this without having done a lot of preparation"*).

A1.4 Conclusion

Overall CCT is a good tool but this is overshadowed by the fact that its purpose and the benefits of using it are not immediately apparent. Entering cost data is a laborious process and the cost comparison graphs are not easy to understand. These are more a matter of concept rather than interface, so not strictly-speaking usability issues. However, clear explanations of how the tool works and why it should be used, coupled with good help information (explanations on mouse-over, FAQs, step-by-step guides) can turn reluctant users into committed followers.

A1.5 Annex 1: Recommendations for Improvement

Issue	Explanation	Possible solutions
Home page		
• No clear starting point for the tool (for both new and registered users) • Too much content detracting from the purpose of the site • Word 'tool' not mentioned anywhere on the homepage • Bubbles replicated on the top menu add no value • Help not visible enough—both 'contact' and 'Related services' mistaken for help	• Users tend to go for either 'Get started' or 'Compare costs', neither of which seem to provide them with an obvious starting point • Users assume (based on the homepage and even halfway through the tasks) that CCEx is a website like moneysupermarket.com, a comparison service of sort, or a website, which offers cost forecast similar to Planforcloud.com. Others thought it would help them create cost model.	• Mention the word tool • Remove bubbles such as 'read more', 'discuss and share', 'related services', which are replicated in the menu • Add better text on benefits of the tool and motivation: *"I think that's the value of the service, to be able to compare notes, to engage with the rest of the community... What do I need to do to make a cost saving? To clearly articulate that on the home page is critical."* • Rename current 'Get started' to 'About' and introduce a 'Get started' page which consists of step-by-step wizard (the one currently used in compare costs) • Create a video tutorial that explains what the tool does, how it does it, how much time it will take to enter data, and how much data will you need. Provide example of costs from anonymous organization. Interviews with those who already submitted data, and what are the benefits of doing so. • Introduce Help page, accessible from main menu and containing FAQ and possibly a form inviting user to describe problem
Language		
• Menu items not understood ('Related services') • Too much jargon	• People did not understand concepts 'cost data set', 'number of copies', scope', etc. • Explanatory text under bubbles not clear, for example, 'Related services' misunderstood for 'Other', users even clicked on it when looking for support. • Not knowing what to put in 'Curation staff' one participant comments: *"I might think that's one person, I don't know if I include myself… It's not apparent to me what you need there. So if you are thinking about comparisons it means that people will be giving you quite a range when they all describe the same thing…"*	• Simpler English needed, for example 'Please use the sliders below to give a full breakdown of your current data' instead of 'Allocate the data volume to the various asset types within the scope that you indicated above'

Issue	Explanation	Possible solutions
• Forms (add/edit organization, add/edit cost data set, add/edit cost unit) too similar in content and too similar in look and contain buttons to functions, which are not available yet, thus adding to confusion about the workflow. • Buttons such as 'edit' are too small, red in colour (usually reserved for errors), and its function changes depending upon the page it appears on. • Any changes made in cost data sets (changing the volume, for example) alter the graphs: *"It has moved everything proportionally, it moved everything up by 50% I guess…now that would be massively annoying if I got nine of these things…to unpick all the other stuff…that would drive me mad…what did I put, how have I done that and it's going to be annoying and fiddly job"*.	• Forms are easy to complete but the help accompanying the various fields isn't clear enough (too much jargon). • 'Add cost data set' and 'Add cost unit' forms are too similar in both content and look, which leads users to wonder whether they are repeating the process by mistake. One user says: "*I don't quite understand why I'm having to put a description in for this and I had to do that on a previous page? Have I put the right description on the previous page I'm asking myself now because is that my overall thing or is that my project thing".* • Users add cost data set then click on 'Analyse and compare costs' which is not available before adding cost units. • Users fail to complete task 6 (amend volume of data) because they assume the 'edit' button on 'Manage cost data sets' form was for changing the name of project or organization rather than the whole cost data set. • The sliders are intuitive and easy to use, except the one for duration of project (on 'Add cost data set' form) where users struggle to select two years.	• Forms to be redesigned. As a bare minimum, add help accessible via 'i' icon • Employ a designer to improve page layout and give user visual clues through the process from add a cost data set through to adding cost unit and saving; • It must be made clear to the user that adding a cost unit is part of the process of adding a cost data set and this can be done through design; • Keep forms clear of anything that is not necessary, white space is fine! • It is not clear where the process of adding costs begins and where it ends so narratives would be useful—'If you would like to add more costs, please create a new cost unit. If not, you can now go and compare your costs with those of other universities'.

Issue	Explanation	Possible solutions
Workflow		
• It is still not very clear to users that adding cost units is part of the process of adding a cost data set, they think it is the next step in the process of comparing costs • For existing users, what should they see when they login? • Inconsistencies in workflow: 'Create an account' is still an option to those who already registered for an account • When filling institution information users were worried about being asked to represent their organisation. They frequently commented they felt they had no authority to do so.	• Step 3 in the wizard can be one of 3 things: 'Manage cost data sets', 'Add/edit cost data set', 'Add/edit cost unit'	• New users should perhaps be taken to the wizard of creating a cost data set • Existing users should perhaps be taken to a dashboard style page, a combination of user profile (with organisation, etc.) and my costs • Make it clear to user that they can edit any information in their profile (including items under 'Information sharing') at any point. • Edit cost unit page should have either a 'Save and close' button or an automatic save function with visual clues (Saving…); some users save part-way through completing a cost unit and they don't like being thrown out of the form and taken back to the 'Edit cost unit' page, they even worry that they have lost their partly completed cost unit.
Comparing costs page		
• Busy with text • Overwhelmed by charts, which are not properly explained • Page layout, split in two columns not very user friendly	• Users moved between global/peer comparison tabs easily but didn't see the 'Compare with other peers' button, they expected to choose the peer from the drop-down list on the left hand side, which is currently a selector for user's cost data sets	• Link to contact details for peer institution should be more visible • Include explanation for how the system makes the comparison, access to the peer's expenditure model, user wanted to judge for themselves whether the comparison was valid, fair • Improve readability of charts, for example, institution names that appear at the top should be bigger

User Testing Table 10—Recommendations for Improvement