

Collaboration to Clarify the Cost of Curation



Report of the 3rd 4C Advisory Board meeting on June 30th 2014 at the Main Library of University Edinburgh

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

Version History

Version	Date	Changed pages / reason	Modified by
0.01	07 August 2014	First draft	KH
0.02	07 August 2014	Second draft	SLM
1.00	07 August 2014	Final version	KH

Attendees

1. Stephan Abrams, California Digital Library
2. Rachel Bruce, Jisc
3. Ron Dekker, Netherlands Organisation for Scientific Research
4. Sean Barker, BAE
5. David Rosenthal, LOCKSS
6. Barbara Sierman, Koninklijke Bibliotheek
7. Eefke Smit, STM Association
8. Neil Grindley, Jisc
9. Raivo Ruusalepp, National Library of Estonia
10. Katarina Haage, Deutsche Nationalbibliothek
11. Paul Stokes, Jisc
12. Alex Thirifays, Danish National Archive
13. Kevin Ashley, U Edinburgh

Purpose of the meeting

- to review progress made on the project;
- to address any problems or challenges that have arisen;
- to reach a shared understanding about the work that needs doing;
- to discuss and agree changes that we may need to introduce into the work plan;
- to review the draft deliverables due in the latter part of the project, in particular the Roadmap
- to gain data from the Advisory Board for the current data gathering exercises and in the process to review/validate the current tools, in particular the CCEX

Minutes

The meeting started with an introduction by project coordinator Neil Grindley in which he highlighted the “excellent progress” rating awarded to the 4C project by the European Commission reviewers in the first project review in March 2014. He gave an overview of the Curation Costs Exchange, the Core Cost Concept Model and the Roadmap. NG also announced the 4C Conference on November 17th and 18th 2014 in the Wellcome Trust Centre in London and invited the AB members to attend.

The AB members asked about the relation between CCEx and CCCM; how are costs, CCEx, CCM interrelated - conceptually and terminology-wise? The Advisory Board members expressed their satisfaction with the project's progress. They emphasize the good inter-operability between the individual work packages and would be happy to be included in test drives for different outputs.

Particular discussion points between the Advisory Board and the project team included:

- *The Curation Costs Exchange (CCEx)*
- *Trust and the cost of auditing digital archives*
- *The 4C Roadmap*

The introduction and highlighting was followed by an update on the status of all work packages of the project.

WP1

PS gave an overview of milestones and deliverables in the project and the overall progress in the project- no questions arose.

WP2

KH gave an overview of stakeholder engagement activities and where the project has got since the last AB meeting. A first indication of the focus for the rest of the year was made which includes amongst others CCEx promotion activities/promotional material, a CCEx focus group with the RDM Community, Roadmap Engagement, including a 4C Roadmap workshop at iPRES 2014, and promoting and carrying out the 4C conference.

Recommendations/comments/ideas from AB:

- *We ask stakeholders to ...; but what's in it for them? -> make clear the benefits of engaging with the project and using the outputs (CCEx etc.)*
- *What are the triggers for the eagerness to use the outputs (CCEx)?*
- *What about benchmark fear? Early adopters set the benchmark!*
- *Review mentions engagement with "open source" in addition to vendors*
- *Could the CCEx become a standard (e.g. for research funders: if a researcher has estimated/compared costs via the CCEx then his data management section in his/her proposal is eligible for funding)? -> was described in WP3 presentation - trusted process for well costed data management plans - e.g. for H2020*

WP3

AT gave an overview of the latest outputs of the work package, such as Cost Concept Model and explained the target group here is managerial, whereas for the Gateway Requirement Specification it is more theorists/developers. The CCEx shows “everything” about costs in digital preservation in one place; but still we need to explain to the potential users that working out costs takes time and effort.

Recommendations/comments/ideas from AB:

- *How do service providers get on the list for the CCEX? -> We approach them and they can approach us.*
- *SB: Costs are sensitive to the non-functional criteria (e.g. length of store, type of data, etc); each organisation has their own different non-functional criteria -> that's why they want to roll their own cost model start from the outside (the organisational detail and enter the cost concept model via the organisational profile route)*
- *The CCEX could become part of e-infrastructure; could it be publicly funded? -> RB volunteered Jisc to help establish contacts with some of the named vendors*

WP4 – Raivo Ruusalepp

RR reported on the objectives, the architecture and the success indicators of the work package. He presented the approach, next steps and visions for task 4.5 (From costs to business models). The deliverable in task 4.4 (The role of risk, benefit, impact and value as an economic determinant in digital curation) was not yet achieved, in task 4.1 (A prioritized assessment of the indirect economic determinants of digital curation) the definition of risk could be seen different from an engineering point of view.

Recommendations/comments/ideas from AB:

- *The IEDs are non-functional characteristics - How much further does the project want go with them? How is it covered on the CCEX? Who are the business drivers?*
- *Idea: to turn the dial diagram on the website into an interactive presentation*
- *The IEDs connect other business activities in an organization to the curational process; it could be worth it to produce audience specific terminology. (Question: Update the deliverable if possible?)*
- *The [RASSC project](#) (Retention and Access Services in Supply Chains) as an example was a big organisation making sure that the non-functional characteristics were taken into account by smaller organisations. The project “aimed to transform current approaches to data retention and access in data intensive supply chains. The targets were aerospace design and surveying in the oil and gas industry. The objective was to provide third-parties and major supply chain partners with a blueprint for delivering new data retention and access services to the whole community.”*
- *Regarding task 4.4: risk is a way of re-lasting an IED to a (costly) way of mitigating its impact.*
- *The ISO16363 sections 5.1 and 5.2 relate to risk and could be called in for further work on the risk modelling*

WP5 – Neil Grindley

NG stated a slow start in defining the approach to the Roadmap, but also the gaining momentum. Questions like “How do we really engage with the vendors?” came up -> we are in a better place to do that at this stage of the project and it is most likely that the provoking a response will naturally happen more in year 2 of the project. The reviewers wanted us to map onto standards – this can happen with the ESRM towards the end of the project. The project was also invoked to “Take a position” --> communication/marketing from our brand/value proposition; what could be the messages that people can disagree with and engage in a debate with us? Do we have a strategy for developing and communicating our messages? What would drive you (a funder) to engage?

Recommendations/comments/ideas from AB:

- *DR: growing collection with flat IT budgets - cost grows at same speed as collection grows (exponential) See graph at end of <http://blog.dshr.org/2014/02/talk-at-fast-bof.html>*
- *How much that should be preserved is being preserved - less than half see <http://blog.dshr.org/2014/03/the-half-empty-archive.html>*
- *Which 10 organisations do we want to be on our awareness and cost contribution wish list? Medical and vendors mentioned specifically. -> Set up a target list of funders?*

- *The project is very public sector focused - private sector (vendors) is a conduit to companies (-> Representation across the chain feasible?)*
- *RD: perhaps start with research data and then go on from there*