



a centre of expertise in data curation and preservation

Preservation Pressure Points: Evaluating Diverse Evidence for Risk Management

Plenary II: Certification
iPRES2006
Cornell University, Ithaca NY
October 2006

Seamus Ross & Andrew McHugh
Digital Curation Centre (DCC)
&
HATII, University of Glasgow



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 UK: Scotland License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/2.5/scotland/>; or, (b) send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Funded by:



Preservation Pressure Points: iPRES2006

Introduction

- Independent measuring of repositories is seen as essential aim
- Taken as axiomatic that certification is one marker for trustworthiness of repository
- We seek to develop the debate a debate on the evidence required for objective and transparent assessment
- Two earlier pieces form a backdrop to this talk:
 - S Ross and A McHugh, 2006, 'The Role of Evidence in Establishing Trust in Repositories', *D-Lib Magazine*, July/August, v.12, n7/8 (Also published in *Archivi e Computer*, August 2006), <http://www.dlib.org/dlib/july06/ross/07ross.html>
 - S Ross and A McHugh, 2005, 'Audit and Certification: Creating a Mandate for the Digital Curation Centre', *Diginews*, 9.5, ISSN 1093-5371, http://www.rlg.org/en/page.php?Page_ID=20793#article1



| D | C | C

a centre of expertise in data curation and preservation

Defining Activities and Context

- UK's Digital Curation Centre (DCC)
- RLG-NARA Audit Checklist
- Center for Research Libraries' (CRL) Certification of Digital Archives project
- Network of Expertise in Long-term Storage of Digital Resources (*nestor*)
- Deutsche Initiative für Netzwerkinformation (DINI)

Pilot Audits

- DCC Audits
 - Engage in Three Audits –National Library, Research Centre, and Scientific Data Centre
 - Three more in the pipeline
- Rationale
 - establish evidence base
 - establish list of key participants
 - refine metrics for assessment
 - contribute to global effort to conceive audit processes

Why Evidence?

- It contributes to consistency
- It creates a mechanism that ensures conclusions can be validated and replicated
- It appears insufficient consideration has been given to the evidence base supporting assessment
- Practical, applicability depends on identification of objective means to demonstrate compliance
- Efforts must probe for evidence of *concrete* processes, structures and functionality
- Selection and evaluation of evidence must be regularised.



| D | C | C

a centre of expertise in data curation and preservation

An Evidence Based Approach

- Documentary
- Testimonial (interview based)
- Observation-based



Archive X

- Majority of conclusions drawn from written self-assessment and staff interviews.
- Little documentation available before visit most available on-site
- Time constraints meant that there was little time to subject the documentation provided on-site to formal analysis
- Questions focused on those responses that demonstrate non-compliance with RLG-NARA Checklist
- Visit did not involve demonstration or give auditors chance to see system in operation



| D | C | C

a centre of expertise in data curation and preservation

Archive Y

- Less documentation available than in Archives X or Z
- No self-assessment completed
- Documentation gathered before site visit (e.g. system procedures/functionality, resources)
- Archive did not familiarise themselves with checklist before visit—saw audit as a passive process (had not imbibed the culture of the checklist)
- Demonstrations provided at this archive essential foundation for evidence



| D | C | C

a centre of expertise in data curation and preservation

Archive Z

- Self-assessment completed, range of staff available to auditors
- Abundant documentation gathered before site visit (e.g. system procedures/functionality, resources)
- Level of documentation meant audit could focus on assessing actual day-to-day practice and observation
- Demonstrations provided at this archive essential foundation for evidence
- Audit was investigative
- Checklist provided a pivotal structural framework

Archive Z provided (i)

- Sometimes mere presence will be encouraging, other times content will require scrutiny
- Several example documents
 - Repository Mission Statement
 - Example Deposit Agreements (including legal arrangements)
 - Job Descriptions
 - Organisational Chart
 - Staff Profiles/CVs/Resumes
 - Annual Financial Reports
 - Business Plan
 - Risk Register
 - Policy Documents



| D | C | C

a centre of expertise in data curation and preservation

Archive Z provided (ii)

- System Procedure Manuals
- Technical Architecture
- Maintenance Reports
- Results of Other Audits
- Other Documentation Records
- Document management processes provide insights
- Privacy concerns must be addressed
- Evaluation methods must be refined

Proposed Process & Outcomes

- Pre-visit documentation
 - technical,
 - financial, *and*
 - organisational insights
- Direct subsequent onsite activities
- Three outputs from each pilot audit
 - report for host organisation
 - public report detailing necessary refinements
 - Suggestions for revision criteria to be delivered to RLG-NARA and *nestor*

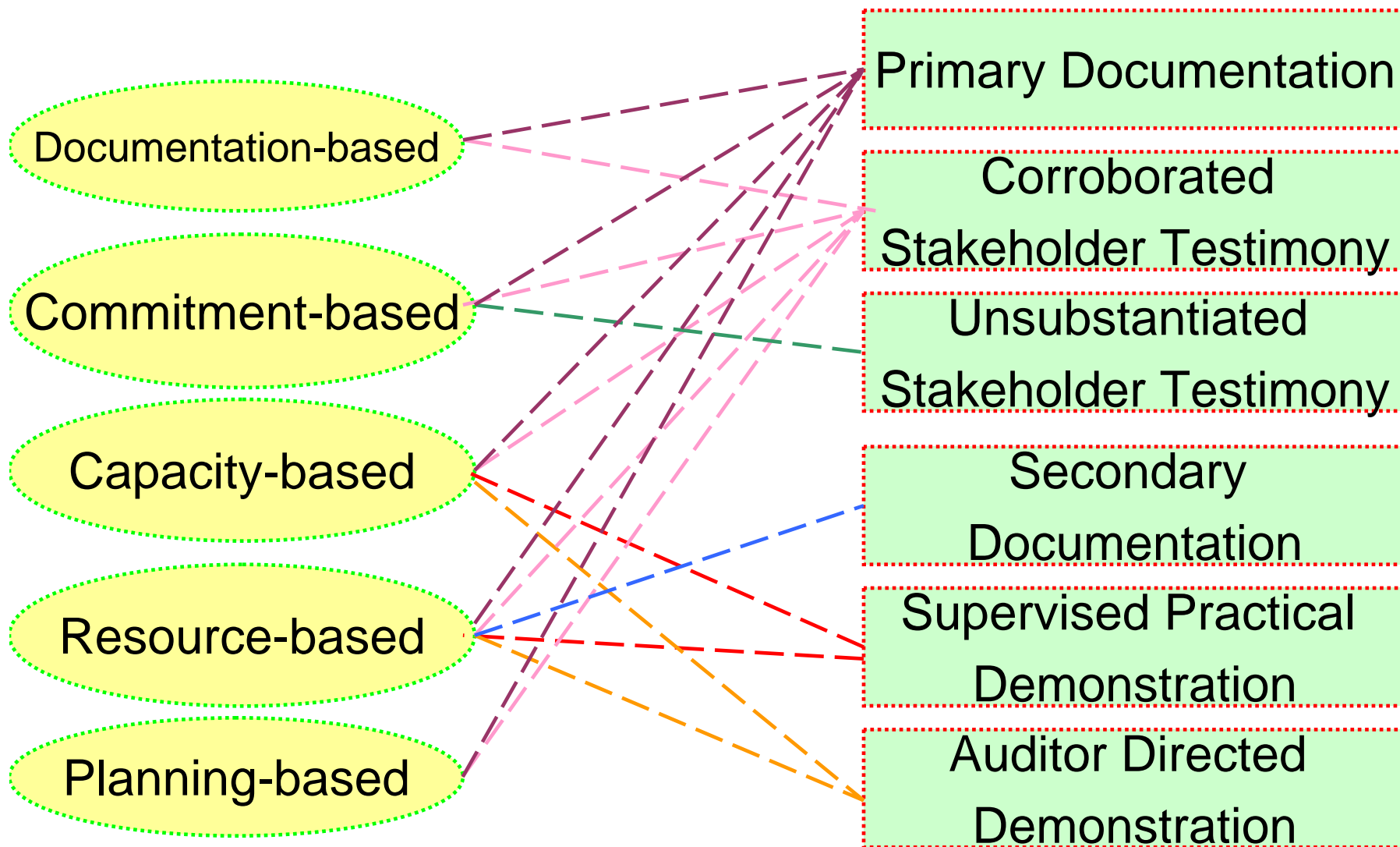


Methodology Shortcomings

- Shortcomings reflect pilot nature of process
- None of the operations of any of the repositories were investigated sufficiently
- Pilot nature has created a 'balance of power' problem for auditors (e.g. Auditor gratitude)
- Length of audits insufficient

Evidential Impact

- What is documented?
- What documentation is available?
- What do staff or stakeholders believe, think, or know happens within the archive?
- What actually happens within the archive?



Testimonial Evidence

- Useful means to:
 - highlight where omissions exist in documentation
 - validate whether documented aspirations are realised in reality
- Roles for interview:
 - Repository Administrators
 - Hardware and Software Administrators
 - Repository Function-specific Officers
 - Depositors
 - Information Seekers
- Questionnaire templates being formulated by DCC

Observation of Practice Evidence

- Less objectively quantifiable, but nevertheless important
- Especially appropriate in terms of procedure and workflow
- Might include
 - walkthroughs
 - testing and measurement of characteristics of objects after preservation action
 - deposit and assessment of test objects (perhaps incrementally over several audits)



| D | C | C

a centre of expertise in data curation and preservation

Where are we?

- Evidence will play a crucial role: an agreed base is necessary for consistency and replication
- Will also assist new repository development
- Require evidence base and metrics for evaluating it
- Downside: highlights cost of process
- Need to better understand how to combine, collect, and use evidence
- Need to conduct more audits



Other Outstanding Issues

- In what circumstances can *quantitative* metrics be established for assessing whether individual criteria have been adequately satisfied?
- How might the qualitative merits of evidence be assessed consistently by different auditors?
- What document procurement powers should be conferred to auditors?
- Will the loss of transparency arising from use of non-disclosure agreements and the lack of detail about how auditors reach their trustworthiness decisions impact on public acceptance of the audit outcome?

Issues to consider

- Process of Audit is effected by the culture of the organisation being audited. All three here different.
- Self-assessment good starting point, but must not be taken at face value
- Likely variation in audit process needs to be recognised reflecting organisation culture and context
- What evidence does your organisation have in place to support audits?