DCH-RP Trust-Building Report

Raivo Ruusalepp
Estonian Ministry of Culture

DCH-RP and EUDAT workshop
Stockholm, June 3rd, 2014

Topics

- Trust in a digital repository
- Trust in a distributed digital repository
- Distributed service models of digital curation
- Trust relationships in distributed service models
- Recommendations for the DCH-RP roadmap
- Does cloud help us mature?

2

Trust or Scepticism?

- Memory institutions enjoy a high level of trust in society
- What happens when we add cloud or another external service provider to the picture?
- Trust tends to turn into scepticism

3

Demonstrating trustworthiness

- How has the DP community understood trust?
- What was the main topic of discussion in DP around 1994?
  - Hardware, storage media, file formats
- Since around 2002 we have been discussing:
  - Functions, workflows, automation
- Digital repository has become the focal point of our discourse on preservation and trust

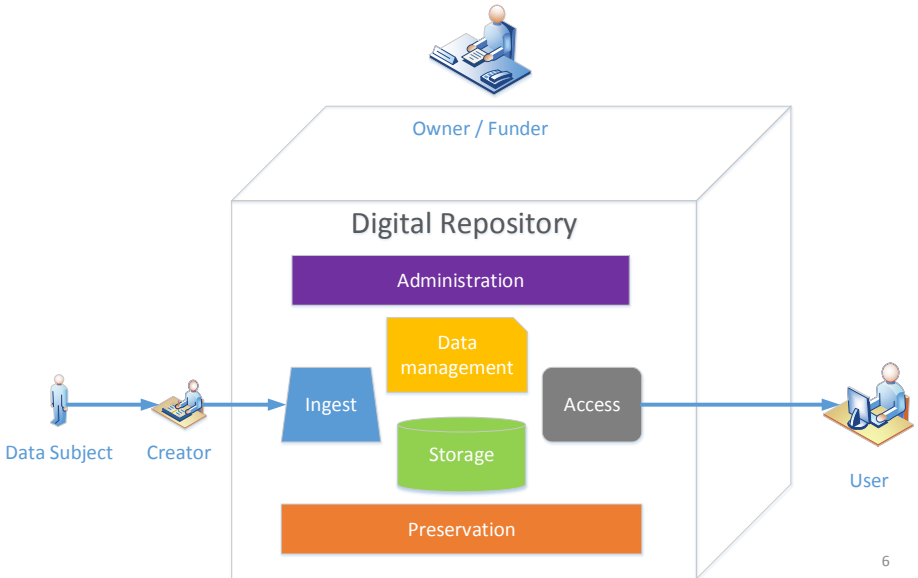
4

TDR – trust or quality?

- *Trusted Digital Repositories: Attributes and Responsibilities,* RLG/OCLC, (2002)
- The 'trusted digital repository' came to be understood as a centralised, single organisation-based preservation service model where the institution that provides the preservation service is also the owner of the digital repository system that houses digital objects
- The practice of applying the TDR criteria over the next decade has demonstrated that the word 'trusted' should more appropriately have been 'quality'
- TDR is essentially about ensuring quality at the operational level of repository work

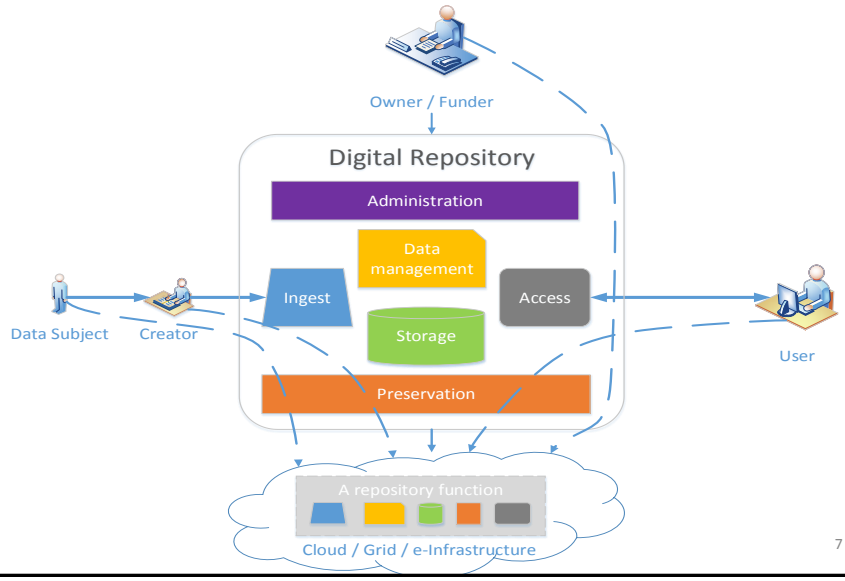
5

Typical trust network of a repository



6

Trust model of a distributed DP service



7

Can TDR audits be applied to cloud?

- Repositories have a number of assessment checklists to evaluate their performance
  - ISO16363, ISO/DIS 16919, ISO14641, ISO17068, DIN31645, TRAC, Data Seal of Approval, DRAMBORA, etc.
- Will a cloud or GRID provider want to certify themselves against any of these criteria?
- How can we extend our quality requirements to our service providers?

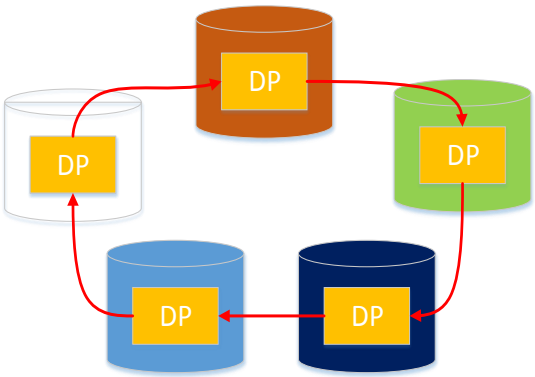
8

Distributed DP service model

- The basic archiving workflow is provided by the OAIS reference model, but it does not articulate clearly how it can cater for distributed archiving architectures
- Cloud, Grid and e-Infrastructure service architectures vary significantly and do not allow for a uniform mapping of preservation services to a single architectural model
- The distributed digital preservation (DDP) model (perhaps as an add-on to the OAIS) is only being developed

9

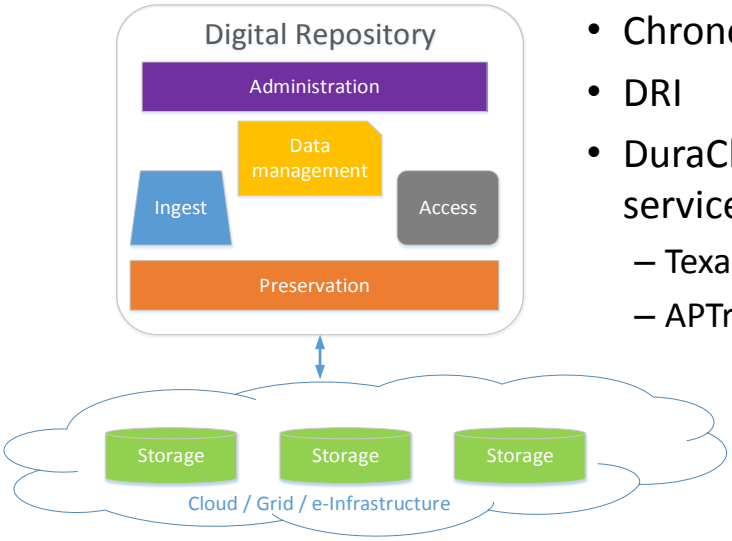
Co-operative service model



- LOCKSS
- Data-PASS
- MetaArchive
- UK LOCKSS Alliance
- LUKII
- DPN
- …

10

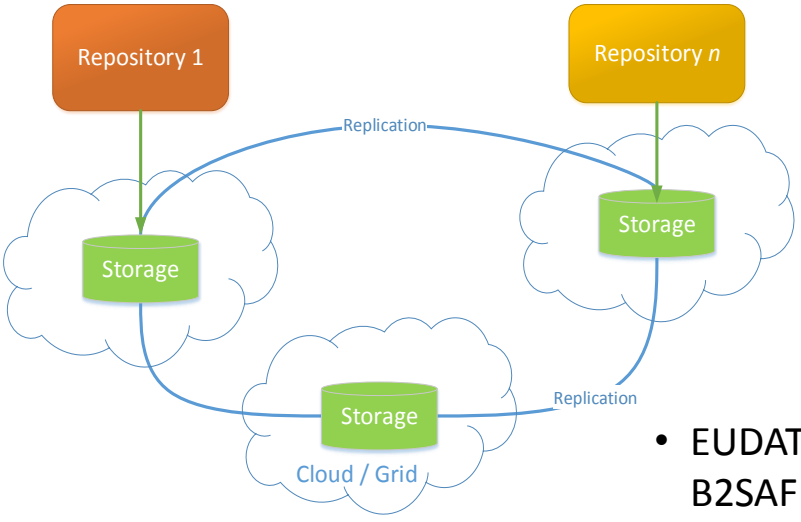
Centralised repository + storage cloud



- Chronopolis
- DRI
- DuraCloud-based services
  - Texas DL
  - APTrust

11

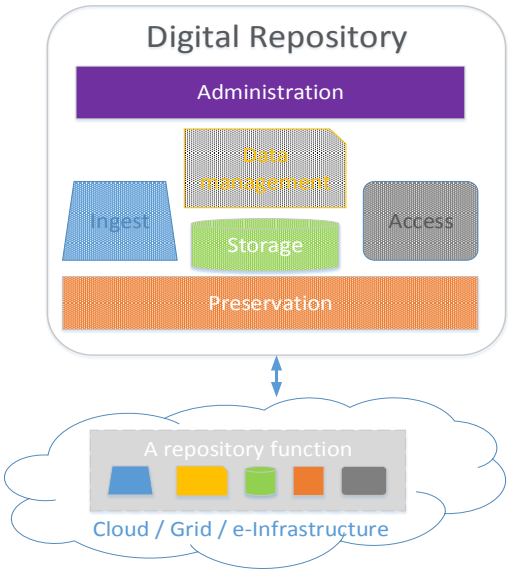
Repository network with shared cloud storage network



- EUDAT B2SAFE

12

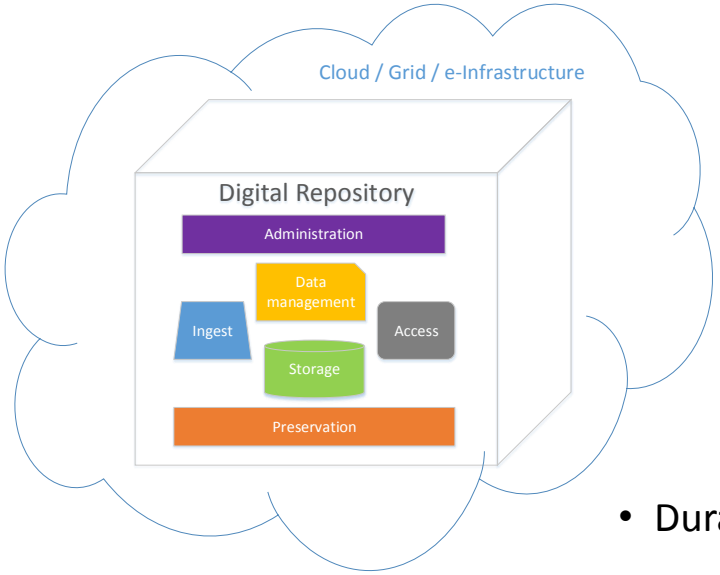
DP service outsourced to cloud



- e-Culture Science Gateway (eCSG) developed in the Indicate and DCH-RP projects

13

Repository in the cloud



- Duracloud

14

Distributed trust relationships

- The distributed digital preservation trust model includes:
  - inter-organisational trust relationships, where both the trustor and the trustee are organisations,
  - individual-organisation type trust relations where an individual (trustor) is interested in trusting an organisation or organisations (trustees)
- Individual-organisation trust relationships are asymmetric
- Inter-organisational trust is governed by contractual agreements (SLA, OLA) and transparent

15

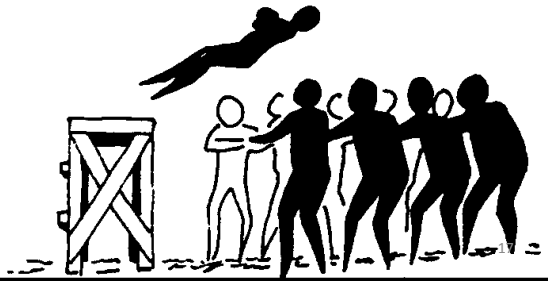
Digital objects and distributed service

- The nature of material that is being preserved
  - an in-copyright e-book will have different preservation and access requirements from open public records or a digitised image of a museum artefact
- The purpose of preserving the object
  - organisations tends to impose less stringent requirements for short term retention of objects for, for example compliance with regulatory requirements than when depositing collections for long-term digital preservation
- The intended users of the preserved object
  - both depositors and users tend to prefer subject or discipline or data type specific repositories to carry out preservation and disseminate data to a specific user group

16

Depositor-Repository Trust

- The trust is of dispositional type
  - when the Depositor consigns material to a Repository he trusts the Repository to carry out active digital preservation actions on the deposited content in the future
  - the Depositor accepts the risk that something can go wrong with maintaining the accessible objects



User-Repository Trust

- Mostly dependent on quality of information provided
- The quality of digital preservation operations of the repository is also considered

18

Funder-Repository Trust

- The main trust driver for owners and funders is customers' (Users, Depositors) satisfaction with the services received from the Repository
- Efficiency of operations, including return on investment, of the Repository is also a significant component but less directly related to trust

19

Repository-Service Trust

- Inter-organisational
- Established through governance and fail-safe mechanisms
  - Generic subcontracting situation
  - Service and operation level agreements
  - Contracts
  - Agreed terms

20

US NDSA report (2013)

The reliability, design, and behaviour of both centralized and distributed preservation networks is just beginning to be understood. It is critical to develop robust trust frameworks that address the risks, because institutions need to be able to measure and evaluate and monitor the reliability and trustworthiness of trustworthy repositories, collaborating organizations and third-party services (such as cloud computing).

21

Trustworthiness through risk analysis

- Situations of trust have come to be described as "a subclass of situations involving risk"
- They are situations in which the risk one takes depends on the performance of another actor
- Risk profiles for third-party cloud services are aplenty
- One example exists where awareness of threats to repository systems and operations, and the ability to deal with the expressed risks are the basis for the claim of being a trustworthy digital repository (DRAMBORA)

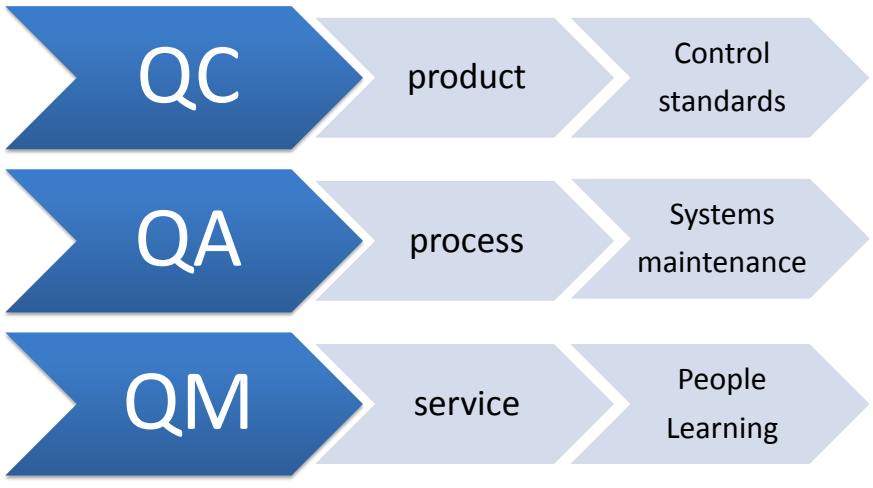
22

What should be done?

- Keep developing distributed digital preservation models and architectures
- Clarify the legal, policy and organisational level issues that are involved when outsourcing DP services
- Extend the TDR audit tools to the distributed architectures and level of services
- Keep track of (and share) incidents and things that go wrong!

23

Maturity: from evidence to learning



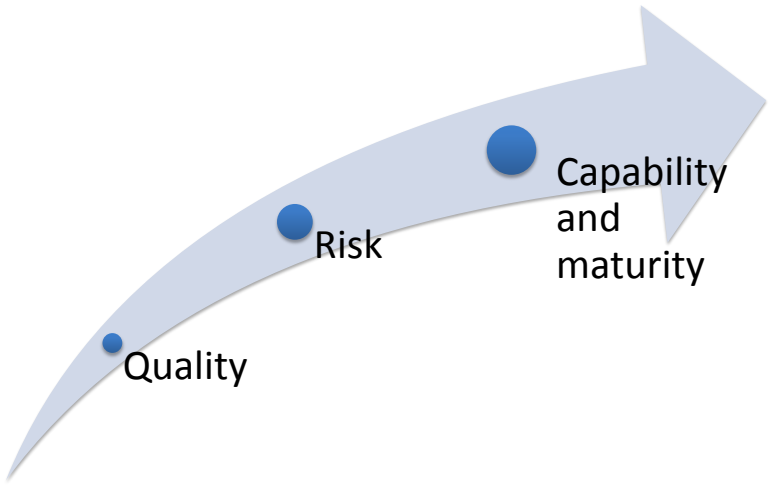
24

Example: Federated Access

- User authentication is required to upload and download data from any repository
- There is a willingness of institutions to be part of federations for managing user authentication
- Federated access provides the technical and policy framework to allow for services to be shared in a trustworthy fashion across borders
  - Users will be able to log in once (single sign-in) using their institutional credentials and access multiple services (sign on)
  - Digital cultural curators and cultural institutions participating will be free of the burden of user name and password administration

25

From Best Practices to Trust



26

Q & A

Raivo Ruusalepp
Raivo.Ruusalepp@nlib.ee

27