# Context and Development of the DRAMBORA Toolkit

ELAG 2007: Library 2.0
Barcelona, 9-11 May 2007

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



# Background

- DRAMBORA developed by Digital Curation Centre (DCC) & DigitalPreservationEurope
- Closely allied with TRAC, nestor criteria, & work of Centre for Research Libraries
- Work conducted by
  - Andrew McHugh (HATII/DCC/DPE),
  - Raivo Ruusalepp (NANETH/DPE/Estonian Business Archives),
  - Seamus Ross (HATII/DCC/DPE), and
  - Hans Hofman (NANETH/DPE)

# Repositories 10 Principles
## CRL/RLG-OCLC/NESTOR/DPE/DCC

- The repository commits to continuing maintenance of digital objects for identified community/communities.
- Demonstrates organizational fitness (including financial, staffing structure, and processes) to fulfill its commitment.
- Acquires and maintains requisite contractual and legal rights and fulfills responsibilities.
- Has an effective and efficient policy framework.
- Acquires and ingests digital objects based upon stated criteria that correspond to its commitments and capabilities.



# Repositories 10 Principles
## CRL/RLG-OCLC/NESTOR/DPE/DCC

- Maintains/ensures the integrity, authenticity and usability of digital objects it holds over time.
- Creates and maintains requisite metadata about actions taken on digital objects during preservation as well as about the relevant production, access support, and usage process contexts before preservation.
- Fulfills requisite dissemination requirements.
- Has a strategic program for preservation planning and action.
- Has technical infrastructure adequate to continuing maintenance and security of its digital objects.

## Critical Services Require Trust

- Task Force on Archiving of Digital Information asserted in 1996:
  > "a critical component of digital archiving infrastructure is the existence of a sufficient number of trusted organizations capable of storing, migrating, and providing access to digital collections."
- RLG/OCLC "Trusted Digital Repositories – Attributes and Responsibilities" (2002)
  - depositors trust information holders
  - information holders trust third party service providers
  - users trust digital assets provided by repositories



## Trust in repositories

- Trustworthiness – a key characteristic that a repository needs to demonstrate
- How can a repository demonstrate this
- Digital curation is all about taking organisational, procedural, technological and other uncertainties and transforming them into manageable risks

# Establishing Trust in a Repository

- How is it established?
- How is it maintained?
- How is it secured?
- What happens when it is lost?
- How can it be verified?
- Can repositories *do* what the say and *show* that they do what they say?
- Have they thought about what they are doing?



# Existing memory institutions

- Are trusted in traditional paper environment
- Why assume their competence in the digital realm?
- New environment requires *all* players to establish trusted status
  - Taxonomy of goods/services (do they belong to same class) do they have similar qualities;
  - we need theory of underlying competence of trustworthy agent for a given task;
  - are the characteristics of that task relevant for a different task

# Legitimacy, Conflicts, & Responsibility

- Weight of legitimacy
  - Who is the accrediting agency?
  - Governmental context?
- Problems with conflicting frameworks
  - Undermine each other
  - People don't want choice
- Legal responsibility of auditors/certifiers



# The Challenge

- Independent measuring of repositories is seen as essential aim
- Taken as axiomatic that audit is a mechanism for establishing the trustworthiness of a repository
- We seek to develop the 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

# Defining Activities and Context

- DCC and DPE collaborations include:
  - Trustworthy Repository Audit and Certification (TRAC) Criteria and Checklist Working Group
    - http://www.crl.edu/PDF/trac.pdf
  - Center for Research Libraries (CRL) Certification of Digital Archives Project
    - http://www.crl.edu/content.asp?l1=13&l2=58&l3=142
  - Network of Expertise in Long-term storage of Digital Resources (nestor)
    - http://edoc.hu-berlin.de/series/nestor-materialien/8/PDF/8.pdf
  - International Audit and Certification Birds of a Feather Group
    - http://www.digitalrepositoryauditandcertification.org



# TRAC Criteria and Checklist

- Outlines best-practice criteria for trusted repositories in three distinct areas
- Takes OAIS as its intellectual foundation, and the benchmark for measuring success
- Aspiration is standardisation; comparable with what ISO 17799 offers for Information Security Audit.
- Emphasizes certification

# nestor Criteria Catalogue

- 14 criteria, enriched by detailed explanations and concrete examples
- http://edoc.hu-berlin.de/series/nestor-materialien/8/PDF/8.pdf
- Groupings entitled:
  - Organisation Framework
  - Object Management
  - Infrastructure and Security
- Reflects the German context (e.g. juridical, financial, and other legislative environments)



# CRL Certification of Digital Archives

- Certification of Digital Archives Project
- Leverage RLG-NARA (now TRAC) work
- Andrew W. Mellon funded research activities to conduct pilot audits in a series of US and European data archives
  - determine optimum methodologies
  - evaluate cost issues
  - deliver specifications for process
  - outline a business model for certifying agency

# Principles of Trustworthy Repositories

- DCC, DPE, CRL and *nestor* met in Chicago in January 2007
- Conceived a global, united perspective on trustworthiness and digital archives
- 10 General Characteristics of Digital Preservation Repositories
- http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=92



# International Audit and Certification Birds of a Feather Group

- An international effort to conceive an ISO standard upon which a full repository audit and certification can be undertaken
- Possibly synonymous with OAIS certification
- Taking existing work and rationalising to a single document or linked collection of documents
- An open process – you could be involved!
- www.digitalrepositoryauditandcertification.org

# Existing Standards Context

- Efforts must also fit gracefully alongside:
  - ISO 9000 series (Quality Assurance)
  - ISO 17799 & 27001 (Information Security)
  - ISO 15489 (Institutional Records Management)
  - ISO 14721 (Reference Model for an Open Archival Information System)
  - COBIT 4.1 (2007)



# Meeting the shortfall

- Independent measuring of repositories is seen as an essential aim
- It's taken as axiomatic that audit is an appropriate mechanism for establishing repository trustworthiness
- Central to this discussion are issues of:
  - criteria for assessment
  - evidence
  - risk management

} particularly relevant for DRAMBORA

# DCC Pilot Audits

- Digital Curation Centre (DCC) engaged in a series of pilot audits in diverse environments
- 6 UK, European and International organisations
- National Libraries, Scientific Data Centers, Cultural and Heritage Archives
- Rationale
  - establish evidence base
  - establish list of key participants
  - refine metrics for assessment
  - contribute to global effort to conceive audit processes
  - *establish a methodology and workflow for audit*



# Archive U

- Transitional system, so much of assessment was based upon planning and requirements documents for new system
- Audit provided opportunity to determine likelihood that new system would alleviate many of the problems associated with the interim system
- Massive capital investment in development project so planning documentation extremely thorough

# Archive V

- Wide ranging and formalised documentation covering comprehensive range of policy and procedure
- Service based on formal contract which necessitated such evidence
- Made compliance straightforward to determine
- Interviews confirmed that what was documented was actually done



# Archive W

- 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 (then) RLG-NARA Checklist
- Visit did not involve demonstration or give auditors chance to see system in operation

# Archive X

- Little documentation available
- No self-assessment completed
- Documentation gathered before site visit (e.g. system procedures/functionality, resources)
- Archive did not familarise 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



# Archives Y & 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

# Themes

- Need to describe evidence base
  - To contribute towards consistency
  - To create a mechanism that ensures conclusions can be validated and replicated
  - Practical, applicability depends on identification of objective means to demonstrate compliance
  - Efforts must probe for evidence of *concrete* processes, structures and functionality
  - Documentary, testimonial, and observational evidence
- Need to establish 'preservation pressure points' including uncertainties and risks
  - Risk awareness is low within the community



# Documentary Evidence

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

# Documentation (continued)

- 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



# 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)



# Evolving Process & Outcomes

- Pre-visit documentation
  - technical,
  - financial, *and*
  - organisational insights
- Direct subsequent onsite activities (2 – 3 days)
- Outputs from each pilot audit
  - report for host organisation (to be published in collective form)
  - Suggestions for revision criteria to be delivered to RLG-NARA and *nestor*
  - Development of DRAMBORA self assessment toolkit

# Risk

- Are repositories capable of:
  - identifying and prioritising the risks that impede their activities?
  - managing the risks to mitigate the likelihood of their occurrence?
  - establishing effective contingencies to alleviate the effects of the risks that occur?
- If so, then they are likely to engender a trustworthy status – if they can demonstrate these capabilities



# Approach to Assessment

- Four key principles lie at the heart of our assessment methods:
  - It should be a self-audit that repositories do themselves, based on the provided tools
  - Self-audit could be a preparatory step for external audit
  - It should be flexible and be valid for repositories of all shapes and sizes and of different contexts
  - It should be assessing how well the repository is managing the risks it is facing when it does what it does
  - It should offer advice on how to overcome the risk situations and what other repositories have done in similar situations

# DRAMBORA

- Easy to say establish evidence and recognise risk, but how do you do this and then take advantage of this knowledge
- *Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)*
- Provides mechanisms to facilitate internal self-assessment & reporting
  - Validates appropriateness of repository's efforts
  - Provides means to generate appropriate documentation
- External certification less of a priority currently, and less immediately viable



# Developing DRAMBORA

- Follows lessons learned from DCC pilot audits
- A collaborative exercise between DCC and DigitalPreservationEurope
- Development will continue with a further period of pilot assessments, training workshops and the release of subsequent versions during 2007 and 2008
- You can download the toolkit at http://www.repositoryaudit.eu

## Not Yet Another Checklist?

- Existing methods are:
  - too static – 'one size fits all' approach
  - too much fixed on the OAIS reference model
  - too little emphasis on evidence in the auditing process
- Audit results should help to manage the repository better continuously, not just give a one-time evaluation



## Core Aspects

- The Authentic and Understandable Digital Object
- Based upon established risk management principles
- Bottom-up approach to assessment (in contrast with TRAC and *nestor methodologies)*
- *Not about benchmarking, but could be used alongside benchmarking standards or criteria*
- *Proactive and retroactive applications*

# Risk and Digital Preservation

- Transforming uncertainties into manageable risks
- ERPANET Risk Communication Tool
  - http://www.erpanet.org/guidance/docs/ERPANETRiskTool.pdf
- Cornell University Library VRC
  - http://irisresearch.library.cornell.edu/VRC/methods.html



# Principles

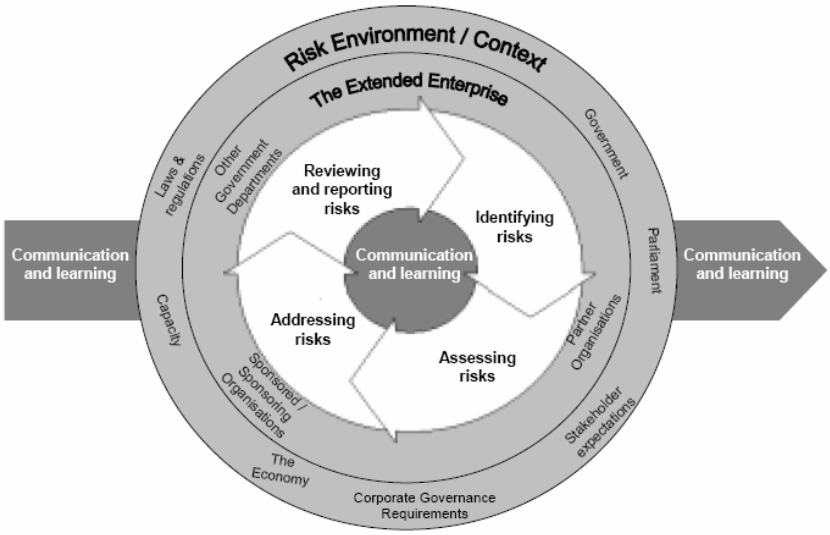
- Appropriateness of auditor
- Measurability of assessment
- Documentation (evidence)
- Flexibility/fluidity to suit a diverse range of repository environments

# Assessing risk

- Most risk assessment exercises are based on a benchmark that is established first
- By defining what success means first it is easy to assess how far from this measure you currently are
- Enterprise risk management is emerging
- Australian Risk Management Standard AS/NZS 4360, latest version is from 2004



# Risk Management Model



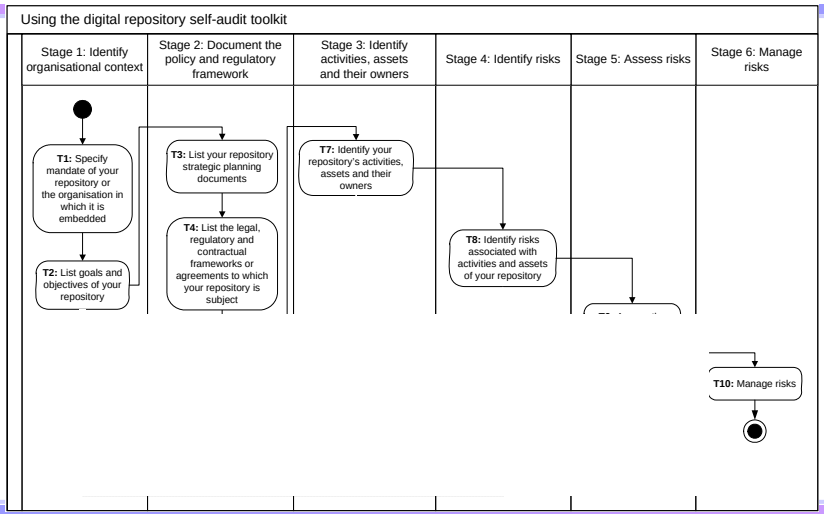
# DRAMBORA Stages

DRAMBORA requires auditors to undertake the following 6 stages:

1. Identification of objectives
2. Identification of policy and regulatory framework
3. Identification of activities and assets
4. Identifying risks related to activities and assets
5. Assessing risks
6. Managing risks



# DRAMBORA Workflow



# Ten Tasks

- What is the mandate of your repository?
- What are the goals and objectives of your repository?
- What policies does your repository have in place to support and regulate how these goals and objectives are to be achieved?
- What legal, contractual and other regulatory requirements / confines does your repository operate in?
- What standards and codes of practice does your repository follow?
- Any other things that influence how your repository does the what it is supposed to be doing?



# Ten Tasks

- What are the activities that your repository does to achieve its goals and objectives within the context and confines set by the regulatory environment, and what assets do you use and produce in the course of these activities, including staff, skills, knowledge, technology?
- What are the risks associated with all of the above?
- How would you assess these risks?
- How do you manage these risks?

# DRAMBORA Outcomes

- Documented organisational self-awareness;
- Catalogued risks;
- Understanding of infrastructural successes and shortcomings;
- Preparation for full scale external audit.



# Interpreting Results

- The self-audit produces a composite risk score for each of the eight functional classes.
- This numeric result can be compared with risk scores of other functional classes and allows the identification of the areas of repository work that are most vulnerable to threats.

# Anticipated applications

- Anticipatory: A process preceding the development of the repository or one or more of its aspects
- Preparatory: A precursor to extended, possibly external audit (based on e.g., TRAC)
- Validatory: Internal self assessment to confirm suitability of existing policies, procedures and infrastructures



# Your role

We would like you to:

- Learn today how to use the audit toolkit
- Use it in a test-audit on any digital repository
- Tell us:
  - what results did you get?
  - where do you think the methodology should be improved and how?
  - what functionality should the on-line tool have?

# Certification Framework

- Levels of certification
- Self-certification
- Certifiable elements
- Accredited Certifying body
- How long will it last?
- When will re-certification take place?
- Surprise audits?
- Consequences of revocation?



# DRAMBORA Future

- Test audits and feedback on the methodology – Spring-Summer 2007
- Version 2.0 to be released in September, as an interactive on-line tool
- Produce a formal audit report at the end of the self-audit
- Version 3.0 in Spring 2008
- Certification of self-auditors in 2008 (?)