Slide 1

D | C | C

Building Trust in Digital Repositories Using DRAMBORA

Seamus Ross, Andrew McHugh,

Raivo Ruusalepp, Hans Hofman & Perla Innocenti

Digital Curation Centre (DCC)

DigitalPreservationEurope (DPE)

HATII at the University of Glasgow & National Archives of the Netherlands

JISC Conference, Manchester, UK, 5-6 June 2007

JISC   Building Trust in Digital Repositories Using DRAMBORA   1

Slide 2

D | C | C

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

JISC   Building Trust in Digital Repositories Using DRAMBORA   2

Slide 3

D | C | C

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   } particularly relevant for DRAMBORA
  - risk management

JISC   Building Trust in Digital Repositories Using DRAMBORA   3

Slide 4

D | C | C

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*

JISC   Building Trust in Digital Repositories Using DRAMBORA   4

Filling a Gap

- 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

Benefits of DRAMBORA

- Following the successful completion of the self-audit, organisations can expect to have:
  - Established a comprehensive and documented self-awareness of their mission, aims and objectives, and of intrinsic activities and assets
  - Constructed a detailed catalogue of pertinent risks, categorised according to type and inter-risk relationships
  - Created an internal understanding of the successes and shortcomings of the organisation
  - Prepared the organisation for subsequent external audit

Objectives

- The purpose of the DRAMBORA toolkit is to facilitate the auditor in:
  - defining the mandate and scope of functions of the repository
  - identifying the activities and assets of the repository
  - identifying the risks and vulnerabilities associated with the mandate, activities and assets
  - assessing and calculating the risks
  - defining risk management measures
  - reporting on the self-audit

What it does not do for you?

- It is not a certifying tool or a OAIS-compliance toolkit, but rather a self-assessment and management tool
- The organization itself sets the benchmark against which it is assessing itself
- The task of DRAMBORA staff is not to audit or assess anyone's result, but to provide the tools for them to do it

Anticipated applications

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

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

Risk Management Model
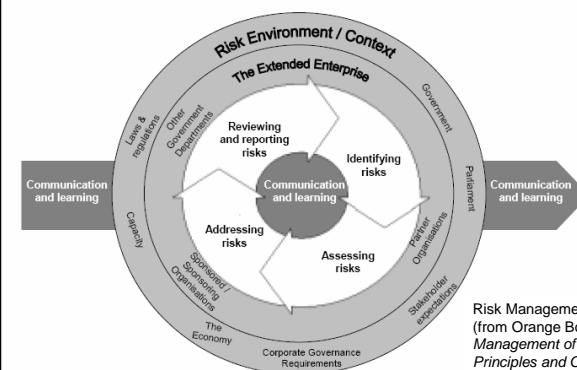
Risk Management Model (from Orange Book. *Management of Risk -- Principles and Concepts*, © Crown copyright 2004)

Slide 13: A Recursive Process

Building Trust in Digital Repositories Using DRAMBORA   13
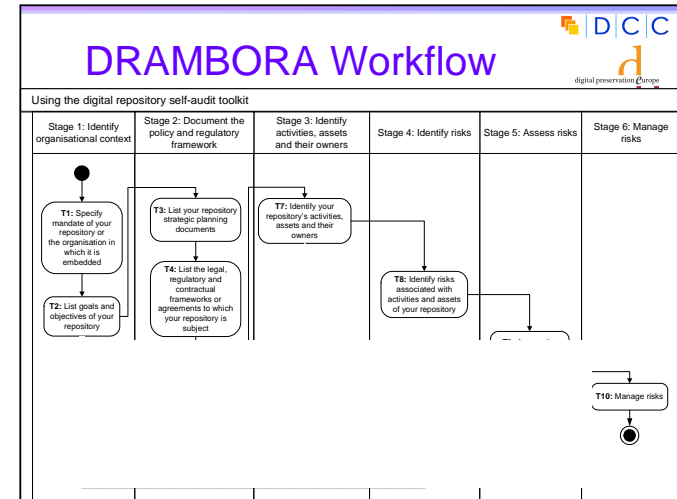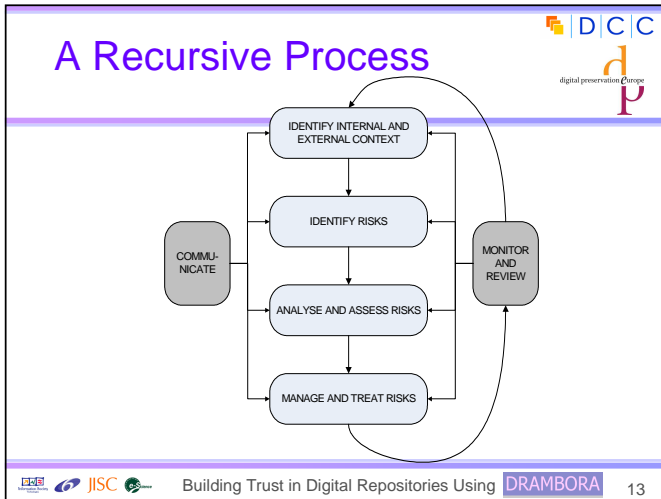
Slide 14: Your role

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?

Building Trust in Digital Repositories Using DRAMBORA   14

Slide: DRAMBORA Workflow

Using the digital repository self-audit toolkit

Slide 16: Stage 1 — Identify organisational context

Stage 1
Identify organisational context

Building Trust in Digital Repositories Using DRAMBORA   16
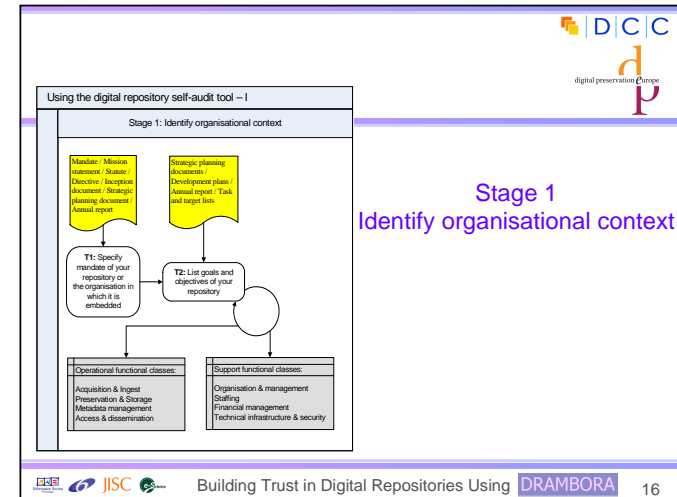
Organisational Context

- The first stage in developing an organisational profile
- Building a platform to facilitate risk awareness
- Success reflects organisational characteristics and aspirations

Stage 1: Tasks

- Identify organisational mandate
  - derived from mission statement or enacting instrument
- Identify organisational goals
  - why does organisation exist?
- Well established means for subsequent risk definition and assessment
- Success demands access to personnel and documentation

Organisational Goals

- Associated with one of 8 functional classes
  - Acquisition & Ingest
  - Preservation & Storage
  - Metadata Management
  - Access & Dissemination
  } operation classes

  - Organisation & Management
  - Staffing
  - Financial Management
  - Technical Infrastructure & Security
  } supporting classes

Stage 1: T1 examples

What is the mandate of your repository or the organisation in which it is embedded?

- To provide a cost-effective, long-term preservation repository for digital materials in support of teaching and learning, scholarship, and research in Scotland
- To collect, list and preserve STM e-thesis as well as making it available to the public
- To focus and strengthen the National Library's efforts to create digital content, and to collaborate with others to ensure that citizens have barrier-free access to the record of their heritage

Stage 1: T2 examples

List goals and objectives of your repository
(Operational functions: Acquisition & Ingest)

- Restrict authorization to deposit materials and withdraw materials
- Ingest of all SIPs delivered to the repository from the user community
- Provide dataset usage statistics for data depositors

Stage 1: T2 examples

List goals and objectives of your repository
(Operational functions: Preservation & Storage)

- Preserve original files exactly as submitted, with demonstrated integrity, viability and authenticity
- Achieve and maintain certification as a Trusted Digital Repository

---

Stage 1: T2 examples

List goals and objectives of your repository
(Operational functions: Metadata management)

- Ensure that data handling within the repository is efficient
- Maintain referential integrity between metadata and archival content

---

Stage 1: T2 examples

List goals and objectives of your repository
(Operational functions: Access and dissemination)

- Provide appropriate preservation strategies to maintain a renderable version of the file at all times
- Provide value-added services to the users within the resources available

Stage 1: T2 examples

List goals and objectives of your repository
(Operational functions: Organisation & Management)

- Provide appropriate reports to associates for management purposes
- Promote the repository and its data collection through regular representation at scientific meetings and provision of appropriate publicity materials

Stage 1: T2 examples

List goals and objectives of your repository
(Support functions: Staffing)

- Define staff roles, responsibility and their relationship
- Provide adequate and regular training

Stage 1: T2 examples

List goals and objectives of your repository
(Support functions: Financial management)

- Maintain financial viability after funding from XXX ceases after 2007
- Organize and monitor fund-raising activities

Stage 1: T2 examples

List goals and objectives of your repository
(Support functions: Technical infrastructure & Security)

- Continue to develop and enhance the infrastructure of the repository
- Package and release the repository software under the Open Source General Public License
- Ensure data security by a combination of physical security and cyber-security

Slide 29:

Using the digital repository self-audit tool – II

Stage 2: Document the policy and regulatory framework

| D | C | C
digital preservation europe

Stage 2

Document Policy and Regulatory Framework

JISC — Building Trust in Digital Repositories Using DRAMBORA — 29

Slide 30:

Document policy and regulatory framework

| D | C | C
digital preservation europe

- Aimed at ensuring the repository:
  – operates correctly with respect to regulatory frameworks
  – has an efficient and effective policy framework
  – is aware of societal, ethical, juridical and governance frameworks
  – is aware of legal, contractual and regulatory requirements to which it's subject

JISC — Building Trust in Digital Repositories Using DRAMBORA — 30

Slide 31:

Strategic Planning Documents

| D | C | C
digital preservation europe

- Identified within:
  – procedural or operational manuals
  – intranet or shared network storage
  – wikis
- Includes
  – Policies
  – Procedures

JISC — Building Trust in Digital Repositories Using DRAMBORA — 31

Slide 32:

Stage 2: examples

| D | C | C
digital preservation europe

Strategic Planning Documents

- Action Plan for the file format (2007)
- Disaster and succession plan (2006)
- Repository X: core activities (2006)
- Repository X risk register (2007)

JISC — Building Trust in Digital Repositories Using DRAMBORA — 32

Legal, regulatory, contractual frameworks

- Including:
  - Statute, case law and regulations
  - Mandatory standards of practice
  - Domain specific regulations
  - Contractual obligations and service level agreements
- Inferred by determining:
  - nature of repository; its domain area; relevant legislation (e.g. enacting legislation); third party contracts

---

Stage 2: examples

Legal, regulatory, contractual frameworks (Acquisition and Ingest)

- Standards
  - ISO 9000:2000 Quality Management Systems Series
  - ISO 27001:2005 Information technology — Security techniques — Information security management systems — RequirementsAgreement between IDLDS and the associates

---

Stage 2: examples

Legal, regulatory, contractual frameworks (Acquisition and Ingest)

- UK Acts of Parliament:
  - Legal Deposit Libraries Act 2003
  - Copyright, Designs and Patents Act 1988
- European Directives, Regulations and Decisions:
  - Directive 2001/29/EC (European Copyright Directive)
  - Fourth and Seventh Company Law Directives on annual and consolidated accounts

---

Stage 2: examples

Legal, regulatory, contractual frameworks (Acquisition and Ingest)

- Agreement between repository X and the associates
- Deposit agreement between with Depositor X

Voluntary codes & other documents

- Voluntary codes:
  - Standards imposed upon or adopted by repository
  - Standards forming the basis for other audits
  - Formal compliance programmes
  - Existing risk management programmes
- Other documents
  - e.g., Internal memorandums

Stage 2: examples

Voluntary codes & other documents (Acquisition and Ingest)

- Repository X operations manual (2007)
- Preferred Ingest File Formats (2006)

Stage 2: examples

Voluntary codes & other documents (Preservation and Storage)

- Repository X Disaster Planning (2005)
- Contingency Plan (2004)

Using the digital repository self-audit tool – III

Stage 3: Identify activities, assets and their owners

T7: Identify your repository's activities, assets and their owners

Strategic objectives and goals listed under Tasks 1 and 2 / Policy and regulatory framework from Tasks 3 - 6

Operational functional classes
Support functional classes

Stage 3

Identify Activities, Assets and their Owners

Activities, Assets and Owners

D|C|C

- Building conceptual model of what the repository does
  - split broad level mission and goals into more specific activities or work processes
  - assign to individual responsible actors
  - link to one or more key assets
  - **clues within**: business process re-engineering; imaging & workflow automation; activity-based costing or management; business classification development; quality accreditation; systems implementation

Organisational Assets

D|C|C

- Includes:
  - information (databases, data files, contracts, agreements, documentation, policies and procedures)
  - software assets
  - physical assets
  - services and utilities
  - processes
  - people
  - intangibles, such as reputation

Instructions for this stage

D|C|C

- Hierarchical analysis
  - breaking up organisation's activities into logical parts and sub-parts
    - charter
    - what makes organisation unique?
    - functions and operations
- Process Analysis
  - look in more detail at how repository conducts its business and what is involved

Stage 3: examples

D|C|C

Identify Activities, Assets and their Owners
(Operational functions: Acquisition & Ingest)

Activity: Verify completeness and correctness of received content
Assets: Digital objects; list of file formats; list of levels of preservation treatment desired for that format by the owner

Activity: Monitor and ingest of SIPs
Assets: Submission package definition; checksums

D | C | C

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Preservation & Storage)

Activity: Implement and review strategies for physical archival storage and migration
Assets: Migration tools; media; digital objects

Activity: Utilise means for functional assessment, including external and internal audit and risk analysis
Assets: Certificate awarded; risk register; disaster planning; organisational reputation

JISC   Building Trust in Digital Repositories Using DRAMBORA   45

D | C | C

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Access & Dissemination)

Activity: Implement authentication and authorization system to reflect agreed access rights and restrictions

Assets: Authentication and authorization systems; Agreement between IDLDS and the associates; Dissemination reports; Withdrawal reports

JISC   Building Trust in Digital Repositories Using DRAMBORA   47

D | C | C

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Metadata Management)

Activity: Acquire preservation metadata for archived content
Assets: Preservation metadata records

Activity: Maintain referential integrity between metadata and archived content
Assets: Digital objects; metadata records; software for maintaining associations

JISC   Building Trust in Digital Repositories Using DRAMBORA   46

D | C | C

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Organisation & Management)

Activity: Negotiate and fulfil legal agreements with producers, depositors and users
Assets: Contracts; legislative or regulatory requirements

Activity: Establish and utilise a mechanism for soliciting feedback from identified community
Assets: Email; other feedback mechanisms; trustworthiness

JISC   Building Trust in Digital Repositories Using DRAMBORA   48

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Staffing)

Activity: Appoint a sufficient number of appropriately qualified staff

Assets: Staff; training library

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Financial Management)

Activity: Define, implement and review short and long-term business plans

Assets: Business document planning; turnover

Activity: Utilise means for financial assessment, including internal and external audits and risk analysis

Assets: Financial audit outcomes; risk register; organisational reputation

---

Stage 3: examples

Identify Activities, Assets and their Owners
(Operational functions: Technical infrastructure & Security)

Activity: Define the information architecture

Assets: System hardware; software; communications infrastructure

Activity: Maintain redundant data and storage and offsite backups

Assets: Backups mechanisms; backup tapes

Using the digital repository self-audit tool – IV

Stage 4: Identify risks associated with activities and assets

T8: Identify risks associated with activities and assets of your repository

Strategic objectives and goals listed under Tasks 1 and 2 / Activities, assets and owners listed under Task 7

Operational functional classes
Support functional classes

Stage 4

Identify Risks

Identifying Risks

- Assets & Activities associated with vulnerabilities – characterised as risks
- Auditors must build structured list of risks, according to associated activities and assets
- No single methodology – brainstorming structured according to activities/assets is effective

---

Kinds of risk

- Assets or activities fail to achieve or adequately contribute to relevant goals or objectives
- Internal threats pose obstacles to success of one or more activities
- External threats pose obstacles to success of one or more activities
- Threats to organisational assets

Anatomy of a Risk

Risk Identifier:	A text string provided by the repository to uniquely identify this risk and facilitate references to it within risk relationship expressions
Risk Name:	A short text string describing the risk
Risk Description:	A longer text string offering a fuller description of this risk
Example Risk Manifestation(s):	Example circumstances within which risk will or may execute
Date of Risk Identification:	Date that risk was first identified
Nature of Risk:	Physical environment
	Personnel, management and administration procedures
	Operations and service delivery
	Hardware, software or communications equipment and facilities
Owner:	Name of risk owner - usually the same as owner of corresponding activity
Escalation Owner:	The name of the individual who assumes ultimate responsibility for the risk in the event of the stated risk owner relinquishing control
Stakeholders:	Parties with an investment or assets threatened by the risk's execution, or with responsibility for its management
Risk Relationships:	A description of each of the risks with which this risk has relationships

---

Risk Relationship	Definition of Risk Relationship
Explosive	where the simultaneous execution of n risks has an impact in excess of the sum of each risk occurring in isolation
Contagious	where a single risk's execution will increase the likelihood of another's
Complementary	where avoidance or treatment mechanisms associated with one risk also benefit the management of another
Contradictory	where avoidance or treatment associated with a single risk renders the avoidance or treatment of another less effective
Atomic	where risks exist in isolation, with no relationships with other risks

Example Risk

- Loss of Trust or Reputation
  - One or more stakeholder communities have doubts about the repository's ability to achieve it's business objectives
- Example manifestation
  - Irrecoverable loss of digital objects provoke community concerns about competence
  - public statement about cut in funding raises concerns about viability of repository's continued operations

Example Risk

- Business policies and procedures are inconsistent or contradictory
  - Rationale and/or practical approach adopted for particular business objectives introduce obstacles to successful completion of other business activities
- Example manifestation
  - Repository requires staff to undertake quality assurance procedures for each object ingested, which takes on average ten minutes, although a further objective is that ingest should take at most eight minutes

Example Risk

- Liability for regulatory non-compliance
  - Repository is liable for failure to conduct its activities in accordance with industrial, business oriented or global regulation
- Example manifestation
  - Repository fails to conform to appropriate jurisdictional health and safety regulations for employees

Example Risk

- Loss of key member(s) of staff
  - Individuals with roles, responsibilities or aptitudes vital to the achievement of business objectives part company with the repository, rendering achievement of those objectives less straightforward
- Example manifestation
  - Repository head systems administrator, the sole individual with knowledge of the system's root password, leaves the organisation to work elsewhere

Example Risk

- Budgetary Reduction
  - Repository's operational budget is reduced
- Example manifestation
  - Local recession provokes budgetary reduction of government financed repository

---

Example Risk

- Incompleteness of submitted package
  - Received packages do not contain information that is necessary to facilitate their preservation
- Example manifestation
  - Submitted package lacks metadata information that, in accordance with contracts, must accompany all deposited content

---

Example Risk

- Media degradation or obsolescence
  - Storage media deteriorates, limiting the extent to which it can be written to and read from
- Example Manifestation
  - Tape stored content is inaccessible or corrupted due to deterioration of magnetic tape
  - Contemporary tape drives are incapable of reading dated storage media which is prolific throughout archive

---

Example Risk

- Unidentified information change
  - Repository is incapable of tracking or monitoring where one or more changes to archived information has taken place
- Example manifestation
  - Repository has failed to record or maintain adequate checksum information to detect where changes have been made to archived content

Example Risk

- Ambiguity of Understandability definition
  - Repository is unable to describe what understandability means with reference to their identified community's expectations or requirements
- Example manifestation
  - Repository preserves information and associated metadata based on a perception of what is required by communities that is not necessarily representative
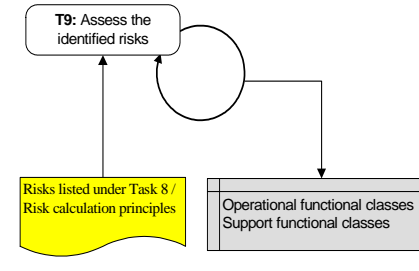

Example Risk

- Authentication subsystem fails
  - Systems for limiting accessibility of information are insufficient, resulting in inappropriate accesses or failures to access
- Example manifestations
  - Individuals who are not entitled to have access to content can access it due to IP based authentication; all local network users connect via a proxy, essentially sharing an IP number and share unrestricted access

---

Using the digital repository self-audit tool – V

Stage 5: Assess risks

T9: Assess the identified risks

Risks listed under Task 8 / Risk calculation principles

Operational functional classes
Support functional classes

Stage 5

Assess Risks

Assess Risks

- Fundamental issues are:
  - probability of risks
  - potential impact of risks
  - Relationships between / groupings of risks
- A risk assessment must be undertaken for each identified risk

Risk Assessment

- For each risk auditors must record:
  - example manifestations of risk
  - probability of its execution
  - potential impact of its execution
  - relationships with other risks
  - risk escalation owner
  - severity or risk (quantification of seriousness, derived as product of probability and impact)

Risk Impact

- Impact can be considered in terms of:
  - impact on repository staff or public well-being
  - impact of damage to or loss of assets
  - impact of statutory or regulatory breach
  - damage to reputation
  - damage to financial viability
  - deterioration of product or service quality
  - environmental damage
  - *loss of ability to ensure digital object authenticity and understandability* is ultimate expression of impact

---

Risk Impact Score	Interpretation
0	*Zero* impact, results in **zero deterioration** of ability to ensure digital object authenticity and understandability
1	*Negligible* impact, results in **isolated, non-serious and recoverable deterioration** of ability to ensure digital object authenticity and understandability
2	*Superficial* impact, results in **isolated but non-serious and/or fully recoverable deterioration** of ability to ensure digital object authenticity and understandability
3	*Medium* impact, results in **widespread or organisation-wide but non-serious and/or fully recoverable deterioration** of ability to ensure digital object authenticity and understandability
4	*High* impact, results in **isolated, serious and non-recoverable deterioration** of ability to ensure digital object authenticity and understandability
5	*Considerable* impact, results in **widespread, serious deterioration** of ability to ensure digital object authenticity and understandability**, which is unrecoverable or recoverable only by third party intervention**
6	*Cataclysmic* impact, results in **organisation-wide, terminal, and unrecoverable loss** of ability to ensure digital object authenticity and understandability

---

Risk Probability Score	Interpretation
1	Minimal probability, occurs once every **100 years or more**
2	Very low probability, occurs once every **10 years**
3	Low probability, occurs once every **5 years**
4	Medium probability, occurs once **every year**
5	High probability, occurs once **every month**
6	Very high probability, occurs **more than once every month**

Determining impact and likelihood

- Consider:
  - Historical experiences
  - Mitigation/avoidance measures already in place
  - Experiences beyond repository itself
    - Relevant research
    - Expert opinion (e.g. legal, technical, environmental)
    - Experiences of comparable organisations

Manage Risks

- Combination of avoidance, tolerance and transfer
  - avoid circumstances in which risk arises
  - limit likelihood of risk
  - reduce potential impact of risk
  - share the risk
  - retain the risk

Using the digital repository self-audit tool – VI

Stage 6: Manage risks

T10: Manage risks

Risks listed under Task 8 / Risk assessment from Task 9 / Risk management methodologies

Operational functional classes
Support functional classes

Stage 6

Manage Risks

Risk Management & DRAMBORA

- The toolkit refrains from prescribing specific management policies
- Instead, auditors should:
  - choose and describe risk management strategy
  - assign responsibility for adopted measure
  - define performance and timescale targets
  - reassess success recursively

Management Risk: Steps

- Auditors should:
  - identify suitable risk responses
  - identify practical responses to each risk
  - identify owners for risk management activities
  - investigate threats arising from risk management
  - prioritise risks
  - update risk register and circulate information
  - secure approval for planning and allocations

Example Avoidance or Treatment (2)

- Staff skills become obsolete
- Avoidance strategies
  - Establish means for staff training, and for staff to employ skills of limited *frequent* value in test environment
  - Implement staff performance reviews to identify skill levels and training req'ts
- In the event of risk's execution
  - Provide training to reverse obsolescence

Example Avoidance or Treatment

- Legal liability for IPR infringement
- Avoidance
  - Assess preserved materials to determine those to which IPR restrictions may apply
  - Seek legal advice to determine legality of actions
- In the event of risk's execution
  - Establish policies and procedures to follow in the event of IPR challenge

Example Transfer Strategy

- Enforced cessation of repository operations
- Transfer Strategy
  - Establish succession arrangements
  - Establish contingency plans or escrow arrangements
  - Establish exit strategy

Example Tolerance Strategy

- Preservation strategies result in information loss
- Tolerance Strategy
  - Implement policy to define the parameters of acceptable loss resulting from these activities

Interpreting the Audit Result

- Composite risk score enables quantification of risks' severity
  - illustrates vulnerabilities
  - facilitates resource investment
- Online tool will feature rich reporting mechanisms
  - what should this consist of?

After the audit

- Improvement requires ongoing activity
  - are risk management strategies working?
  - are risks within a satisfactory tolerance level?
  - risk exposure must be reassessed on an ongoing basis
  - risk management strategies must be re-evaluated
  - management must be informed of developments

What we'd like to know

- What features would you like to see within the toolkit's online version?
- What have you learned about your repository following DRAMBORA assessment?
- Have you combined DRAMBORA effectively with other tools/check-lists?

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 (?)

Closing Questions?

- If you have any further questions please email us at feedback@repositoryaudit.eu
- We'd be delighted to hear of your own experiences using the DRAMBORA toolkit