Building Trust in Digital Repositories Using the DRAMBORA Toolkit

Joint DCC and DPE Tutorial

The British Library

April 27, 2007

Instructors

- Andrew McHugh
  HATII / DCC, University of Glasgow
- Hans Hofman
  National Archives of Netherlands / DPE
- Raivo Ruusalepp
  DPE / Estonian Business Archives

Programme of the day

11.00 – 11.15 Introductions

11.15 – 12.15 Context and Development of the DRAMBORA Toolkit

12.15 – 12.45 Lunch

12.45 – 14.30 Tutorial I: A Simulated Audit Using DRAMBORA

14.15 – 14.30 Coffee

14.30 – 16.00 Tutorial II: A Simulated Audit Using DRAMBORA

Context and Development of the DRAMBORA Toolkit

Joint DCC and DPE Tutorial

The British Library

April 27, 2007

What do digital repositories do?

- Handle a wide variety of media types
- Guarantee authenticity of the object it holds *over time*
- Protect integrity from intended and accidental harm *over time*
- Enable verification
- Ensure accessibility
- Be self-contained

Trust in repositories

- Trustworthiness is an important characteristic that the repository will have to demonstrate
- How to demonstrate trust in a repository?
- Digital curation is all about taking organisational, procedural, technological and other uncertainties and transforming them into manageable risks

Ten Characteristics of Repositories

- Commits to continuing maintenance of digital objects for its identified community(ies).
- Demonstrates organisational fitness (including financial, staffing, structure, processes) to fulfil its commitment.
- Acquires and maintains requisite contractual and legal rights and fulfils responsibilities.
- Has effective and efficient policy framework.
- Acquires and ingests digital objects based upon stated criteria that correspond to its commitments and capabilities.
- 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.
- Fulfils requisite dissemination requirements.
- Has strategic programme for preservation planning and action.
- Has technical infrastructure adequate for continuing maintenance and security of 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

Repositories must….

- Ensure stuff ingested into the archive can be output (e.g. be accessible) – logically intact, syntactically viable, and semantically accessible.
- Guarantee authenticity of the objects they hold
- Be Secure
- Maintain all documentation in-house
- Have disaster recovery functionality built-in
- Have exit strategies
- In addition……..

…be trusted

- Processes:
  - Workflows
  - Operation (management of integrity, authenticity, intelligibility, and accessibility
  - Automation (e.g. ingest, management, publication)
  - Documentation of procedures
  - Auditability
- Architecture and Implementation
- Organisation…….[and more]

Trust Explained

- Expectations of depositors
- Aspirations of service providers
- Management concerns
- Security
- Authenticity and integrity
- Accessibility
- Documentation, metadata and assets self-contained and accommodated in-house

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?

Attributes and Responsibilities (RLG-NARA): an approach

- Compliance with OAIS
- Administrative Responsibility
- Organisational Viability
- Financial Sustainability
- Technological and Procedural Suitability
- System Security
- Procedural Accountability

OAIS Functional Entities

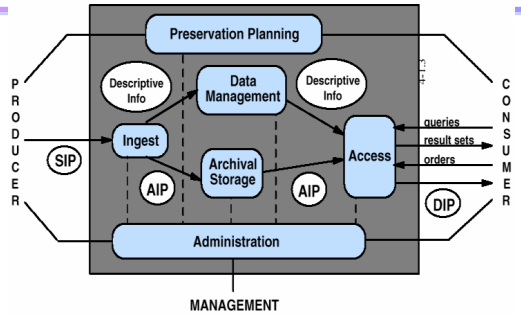


Image from – Reference Model for an Open Archival Information System (OAIS) – CCSDS,2002, http://www.ccsds.org/documents/650x0b1.pdf

Audit and Certification

- Formal means of establishing trust
  - people
  - data
  - processes
  - managing of organisation
  - policies, procedures

How does an audit proceed?

- Peer review?
- Payment? How much?
- Incentives?
- How is independence assured?
- Who is the ideal auditor?

Defining Activities and Context

- UK's Digital Curation Centre (DCC) and Europe's Digital Preservation Europe (DPE)
- Collaboration with:
  - Trustworthy Repository Audit and Certification (TRAC) Criteria and Checklist Working Group
  - Center for Research Libraries' (CRL) Certification of Digital Archives project
  - Network of Expertise in Long-term Storage of Digital Resources (*nestor*)
  - International Repository Audit and Certification Birds of a Feather Group

TRAC Criteria and Checklist

- Outlines best practice criteria for trusted repositories in three distinct areas
- Currently available at: http://www.crl.edu/PDF/trac.pdf
- 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
- More about certification than audit

nestor Criteria Catalogue

- 14 criteria, enriched by detailed explanations and concrete examples http://edoc.huberlin.de/series/nestormaterialien/8/PDF/8.pdf
- Groupings entitled:
  - Organisation Framework
  - Object Management
  - Infrastructure and Security
- Relates specifically to a German context

DRAMBORA

- DCC and DPE conceived the *Digital Repository Audit Method Based on Risk Assessment* in early 2007
- Based on a number of test-audits conducted by the DCC and an analysis of existing audit criteria
- First version available from http://www.repositoryaudit.eu

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
- Other audit frameworks: COBIT 4.0 (2005, www.isaca.org) on IT governance
  - new version COBIT 4.1 (2007)

COBIT 4.0 (1)

- **Strategic alignment** focuses on ensuring the linkage of business and IT plans; on defining, maintaining and validating the IT value proposition; and on aligning IT operations with enterprise operations.
- **Value delivery** is about executing the value proposition throughout the delivery cycle, ensuring that IT delivers the promised benefits against the strategy, concentrating on optimising costs and proving the intrinsic value of IT.
- **Resource management** is about the optimal investment in, and the proper management of, critical IT resources: applications, information, infrastructure and people. Key issues relate to the optimisation of knowledge and infrastructure.

COBIT 4.0 (2)

- **Risk management** requires risk awareness by senior corporate officers, a clear understanding of the enterprise's appetite for risk, understanding of compliance requirements, transparency about the significant risks to the enterprise, and embedding of risk management responsibilities into the organisation.
- **Performance measurement** tracks and monitors strategy implementation, project completion, resource usage, process performance and service delivery, using, for example, balanced scorecards that translate strategy into action to achieve goals measurable beyond conventional accounting.

What are we seeking to audit?

- Institutional means to manage context to ensure preservation:
  - people
  - data
  - processes
  - management
  - technological means
  - resource

Fundamental Question is of 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

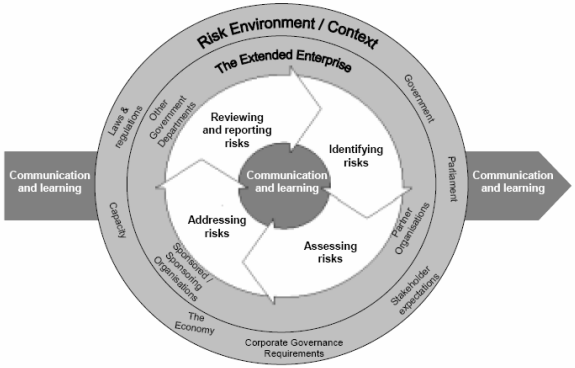
DCC/DPE Audit Principles

- It should be a self-audit that repositories do themselves, based on the provided tools
- Self-audit could be a preparatory step for taking an 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

Assessing risk

- Most risk assessment exercises are based on a benchmark that is established first
  - must be contingent based on the business context
- 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 Core Aspects

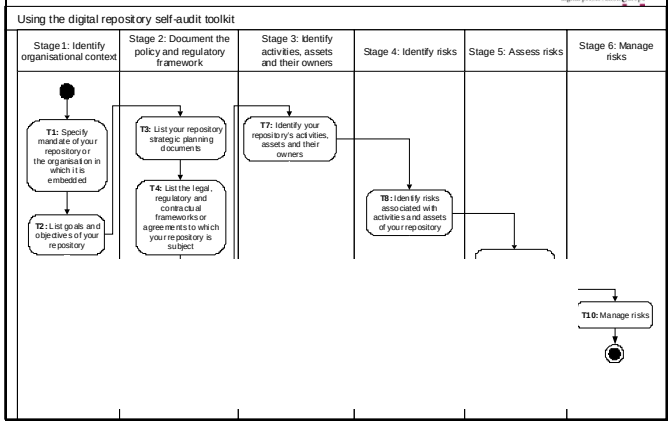
- Authentic and understandable digital object
- Risk based
- Bottom-up approach to assessment (contrast with TRAC and *nestor* methodologies)
- Not about benchmarking, but could be used alongside benchmarking standards or criteria
- Could accommodate different standards, such as ISO/IEC 17799, ISO/IEC 27001, ISO 15489 (RM), ISO 14721 (OAIS), others

DRAMBORA Stages

DRAMBORA requires auditors to undertake the following 6 stages:

- Identification of objectives
- Identification of policy and regulatory framework
- Identification of activities and assets
- Identifying risks related to activities and assets
- Assessing risks
- 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?

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

- 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

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 (?)

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?
  - what other applications of the approach you see feasible?
  - how does this fit into a broader perspective?

Feedback

Please send all your comments, thoughts, suggestions, criticisms, opinions to:

feedback@repositoryaudit.eu

Thank you!

DRAMBORA in Practice: Using the Self Audit Toolkit

Joint DCC and DPE Tutorial

The British Library

April 27, 2007

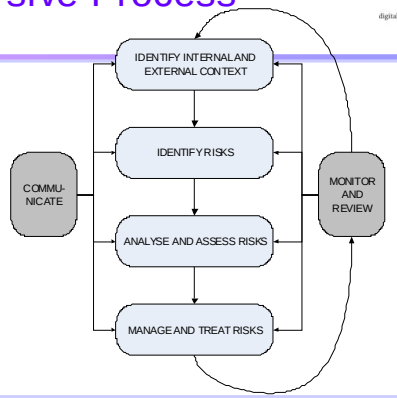
DRAMBORA Outcomes

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

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

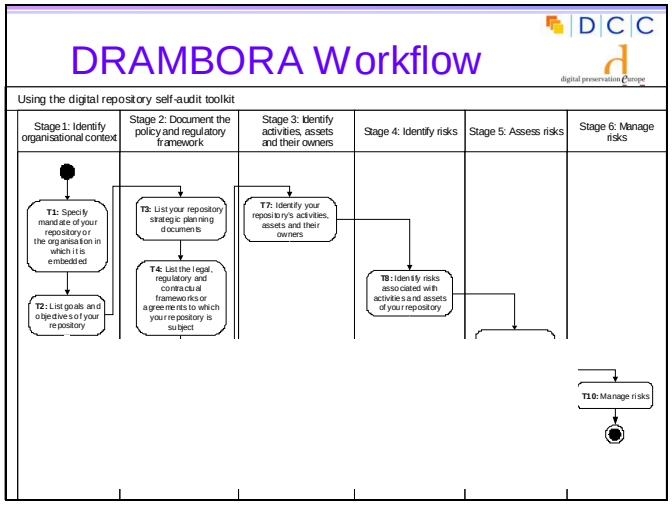
A Recursive Process



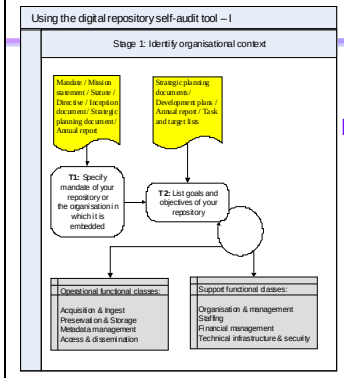
DRAMBORA Stages

- Establish organisational profile
- Develop contextual understanding
- Identify and classify repository activities and assets
- Derive registry of pertinent risks
- Undertake assessment of risks (and existing management means)
- Commit to management strategies

DRAMBORA Workflow



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
Complementry	where avoidance or treatment mechanisms associated with one risk also benefit the management of another
Domino	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



Identify organisational context

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 Mandate

- Example Mandate:
  - The role of *[repository_name]* is to assist researchers to locate, access and interpret *[type_of_data]* produced by *[named_data_creator_group]* and to ensure its long term integrity.

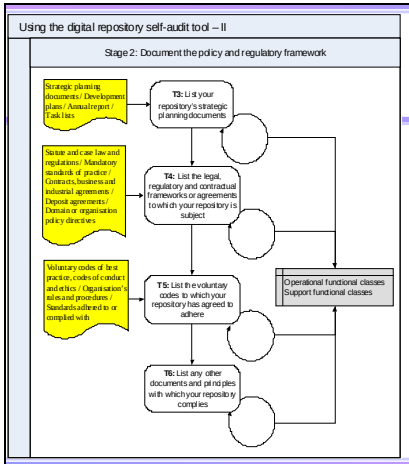
Organisational Goals

- 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



Document Policy and Regulatory Framework

Document policy and regulatory framework

- 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

Strategic Planning Documents

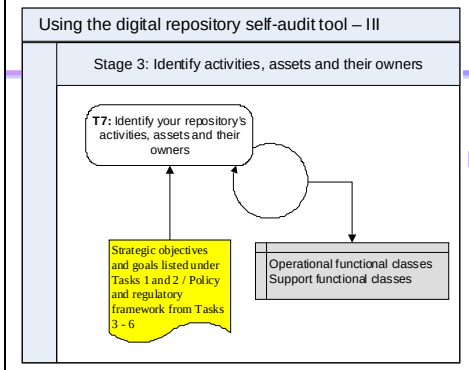
- Identified within:
  - procedural or operational manuals
  - intranet or shared network storage
  - wikis
- Includes
  - Policies
  - Procedures

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

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



Identify Activities, Assets and their Owners

Activities, Assets and Owners

- 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 & work flow automation; activity-based costing or management; business classification development; quality accreditation; systems implementation

Instructions for this stage

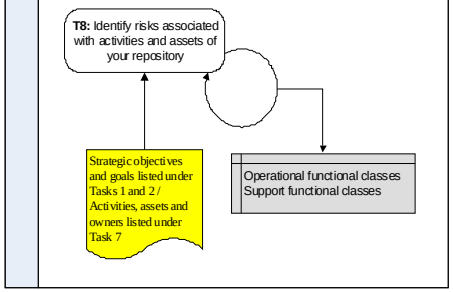
- Hierarchical analysis
  - breaking up organisation's activities into logical parts and sub-parts
    - charter
    - what makes organisation unique?
    - functions and operations
- Process Analysis
  - look in more detail at how repository conducts its business and what is involved

Organisational Assets

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

Using the digital repository self-audit tool – IV

Stage 4: Identify risks associated with activities and assets



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

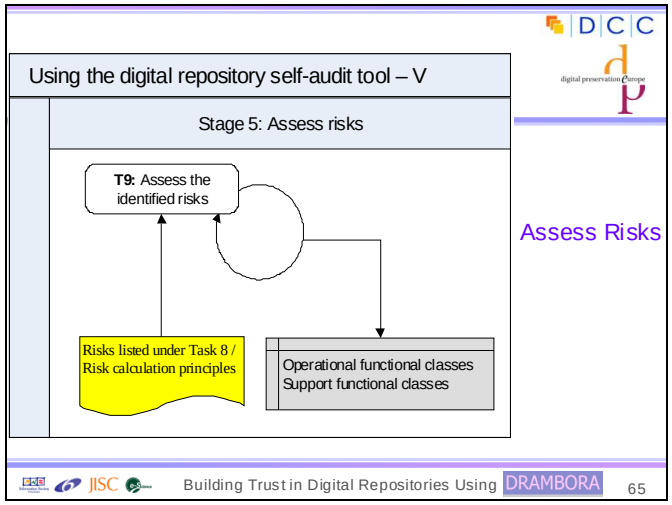
Kinds of risk

- 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 Probability:	*This indicates the perceived likelihood of the execution of this particular risk*
Risk Potential Impact:	*This indicates the perceived impact of the execution of this risk in terms of loss of digital objects' understandability and authenticity*
Risk Severity:	*A derived value, representing the product of probability and potential impact scores*
Risk Management Strategy(ies):	*Description of policies and procedures to be pursued in order to manage (avoid and/or treat) risk*
Risk Management Activity(ies):	*Practical activities deriving from defined policies and procedures*
Risk Management Activity Owner:	*Individual(s) responsible for performance of risk management activities*
Risk Management Activity Target	*A targetted risk-severity rating plus risk reassessment date*



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 Score	Interpretation
0	*Zero* impact, results in **zero loss** of digital object authenticity and understandability [1]
1	*Negligible* impact, results in **isolated but fully recoverable loss** of digital object authenticity and understandability
2	*Superficial* impact, results in **widespread but fully recoverable loss** of digital object authenticity and understandability
3	*Medium* impact, results in **total but fully recoverable** loss of digital object authenticity and understandability
4	*High* impact, results in **isolated loss, including unrecoverable loss** of digital object authenticity and understandability
5	*Considerable* impact, results in **widespread loss, including unrecoverable loss or loss that is recoverable only by third party** of digital object authenticity and understandability
6	*Cataclysmic* impact, results in **total and unrecoverable loss** of digital object authenticity and understandability

[1] Note that we use understandability in its broadest sense to encapsulate technical, contextual, syntactical and semantic understandability.

Risk Impact

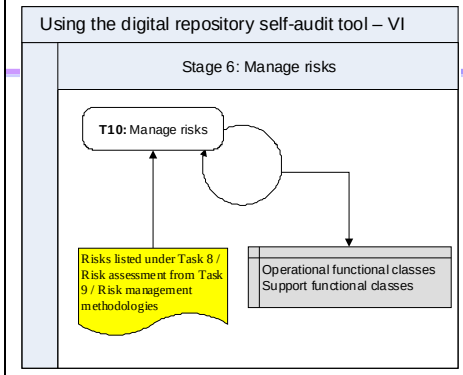
- Impact can be considered in terms of:
  - impact on repository staff or public well-being
  - impact of damage to or loss of assets
  - impact of statutory or regulatory breach
  - damage to reputation
  - damage to financial viability
  - deterioration of product or service quality
  - environmental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

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

Using the digital repository self-audit tool – VI



Manage Risks

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

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

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

Improving DRAMBORA

- Toolkit usability concerns remain
- Can a single individual coordinate an audit?
- Can risks be effectively derived where activities meet or transactions occur?
- We're very interested to hear your thoughts (now, or after you use DRAMBORA)

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?

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