



Digital**Preservation**Coalition



# Risk Assessment of Digital Holdings

**Angela Dappert**

Digital Preservation Coalition

The TIMBUS Project



# Overview

## Risk management

In general



In Information Management



In Digital Preservation

## Status of RM in Digital Preservation

- Examples
- Guidelines
- Applications
- Tools



# Motivation: Risk Impact

- Damage to or loss of our digital assets
- Loss of access, understandability and authenticity
- Statutory or regulatory breach
- Deterioration of product or service quality
- Damage to reputation
- Damage to financial viability
- On public well-being
- On repository staff
- Environmental damage





# Risk

is uncertainty of outcome





# Digital Preservation

The series of managed activities necessary to ensure continued access to digital materials for as long as necessary.

Beagrie & Jones

How do you determine which action to take?





# Digital Preservation

to our digital assets



**Keep risks from becoming issues**

**Proactive**  
preservation

**Risk Management**

**Deal with issues when they arise**

conservation  
**Reactive**

Risk: may happen

- ❖ negative impact - threat
- ❖ (positive impact - an opportunity)

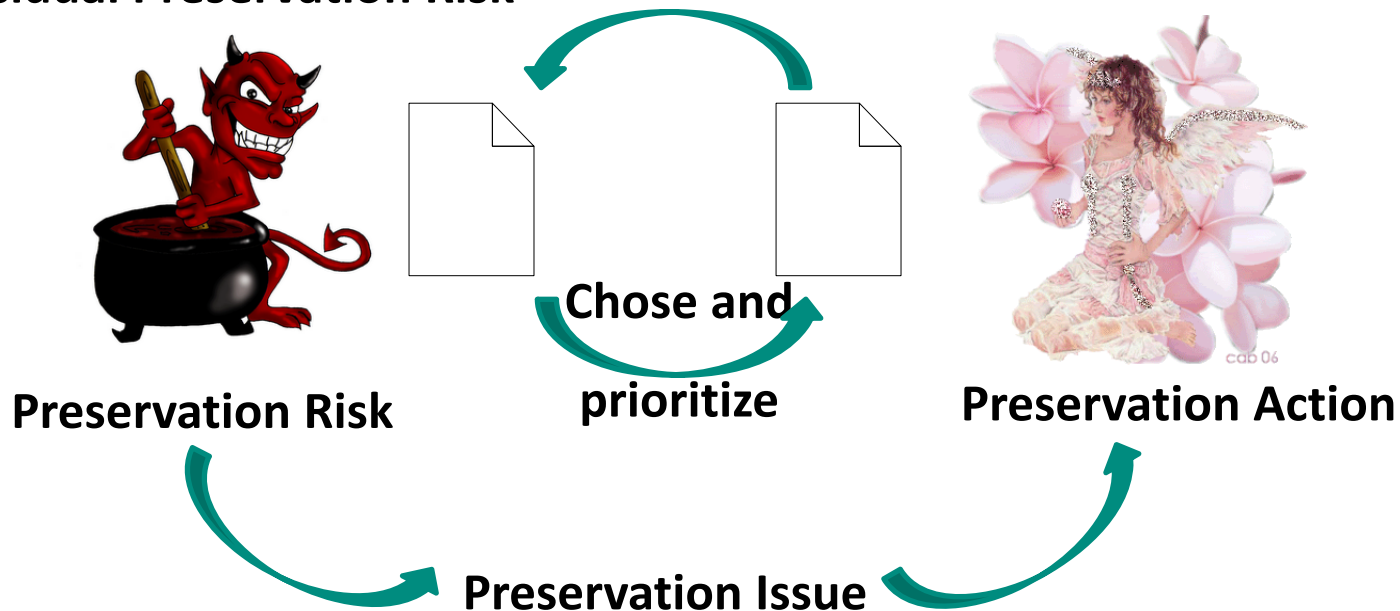
Issue: has happened



# Digital Preservation

- Central function: Risk Management

Residual Preservation Risk



- A support function for the overall organization
- Integrated into the organizational flow



Digital**Preservation**Coalition

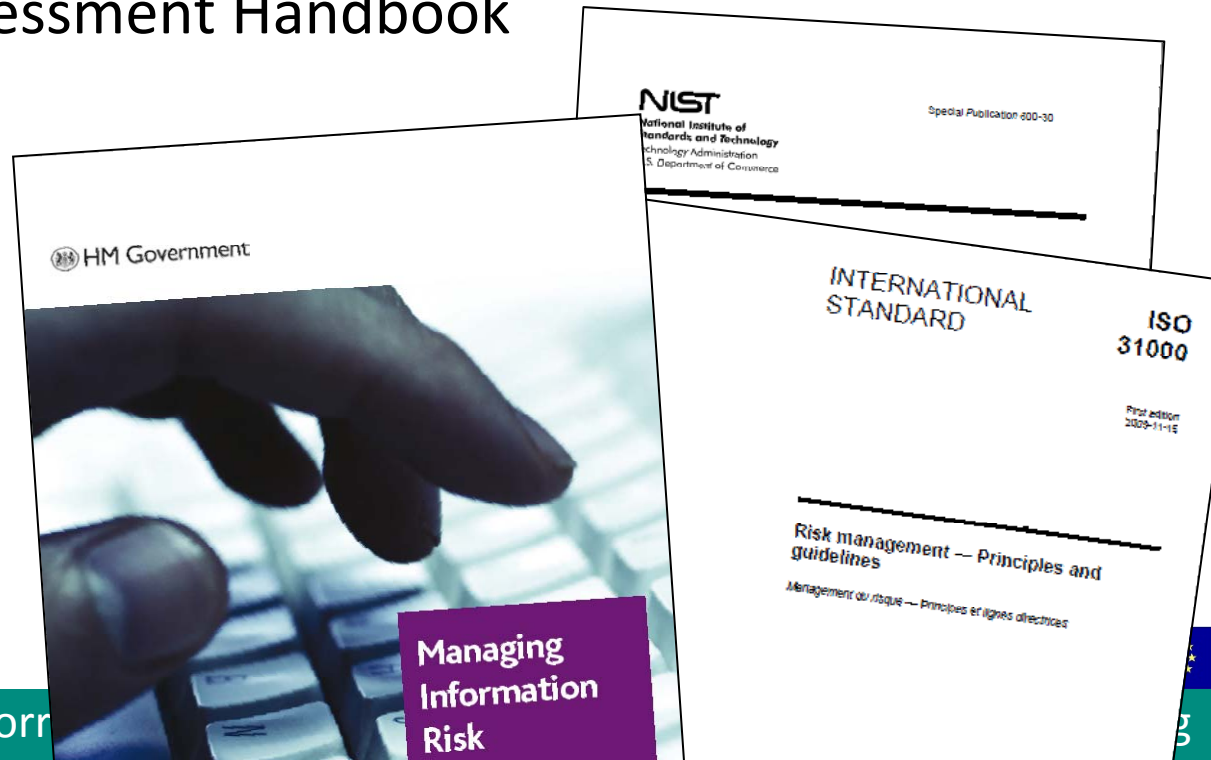
# Risk Management – Familiar Terrain

**Risk Management** – Principles and Guidelines: e.g. ISO 31000

**Information Risk Management** &  
Information Assurance Maturity Model IAMM

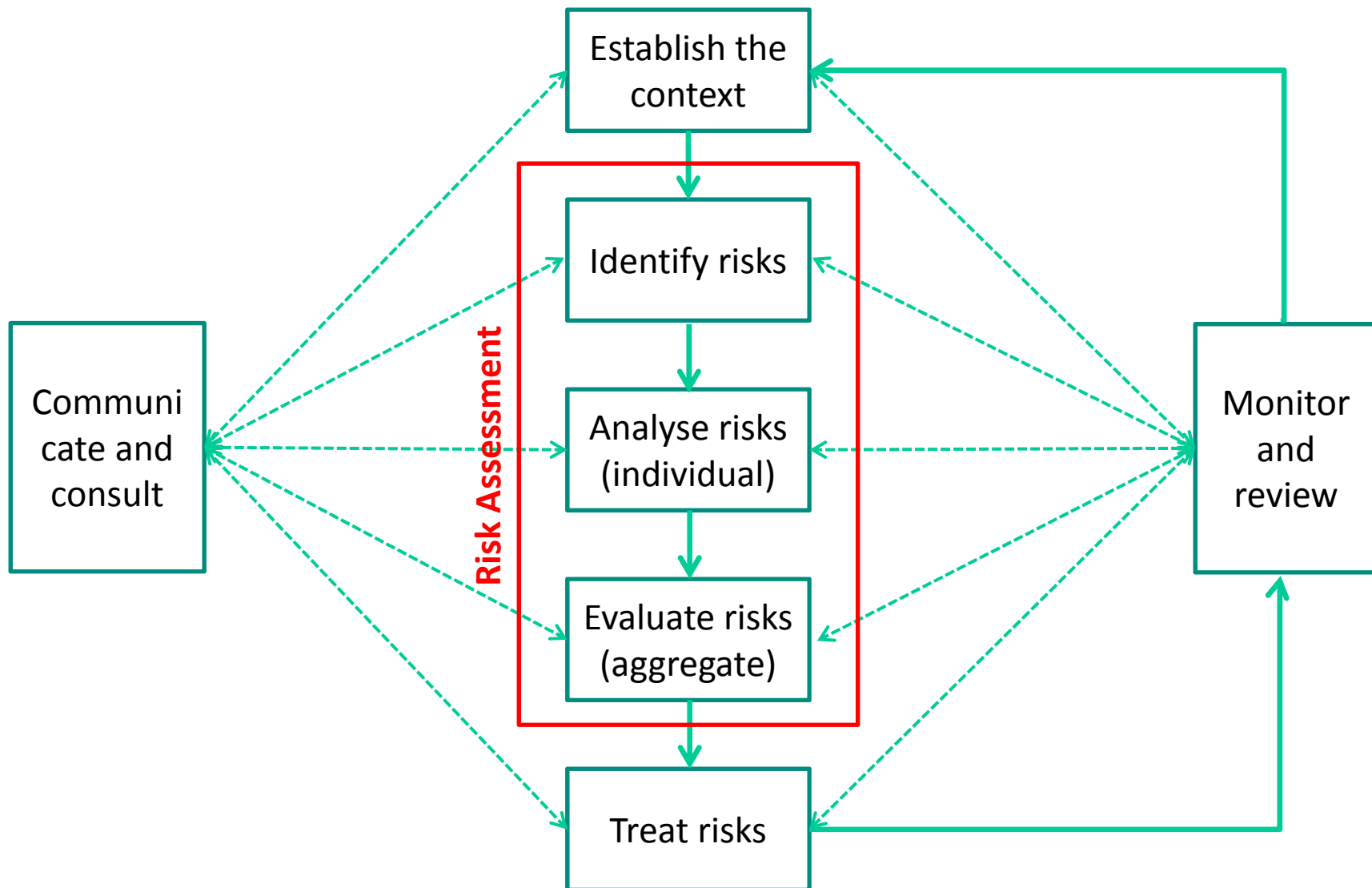
**Digital Continuity**

– e.g. TNA Risk Assessment Handbook

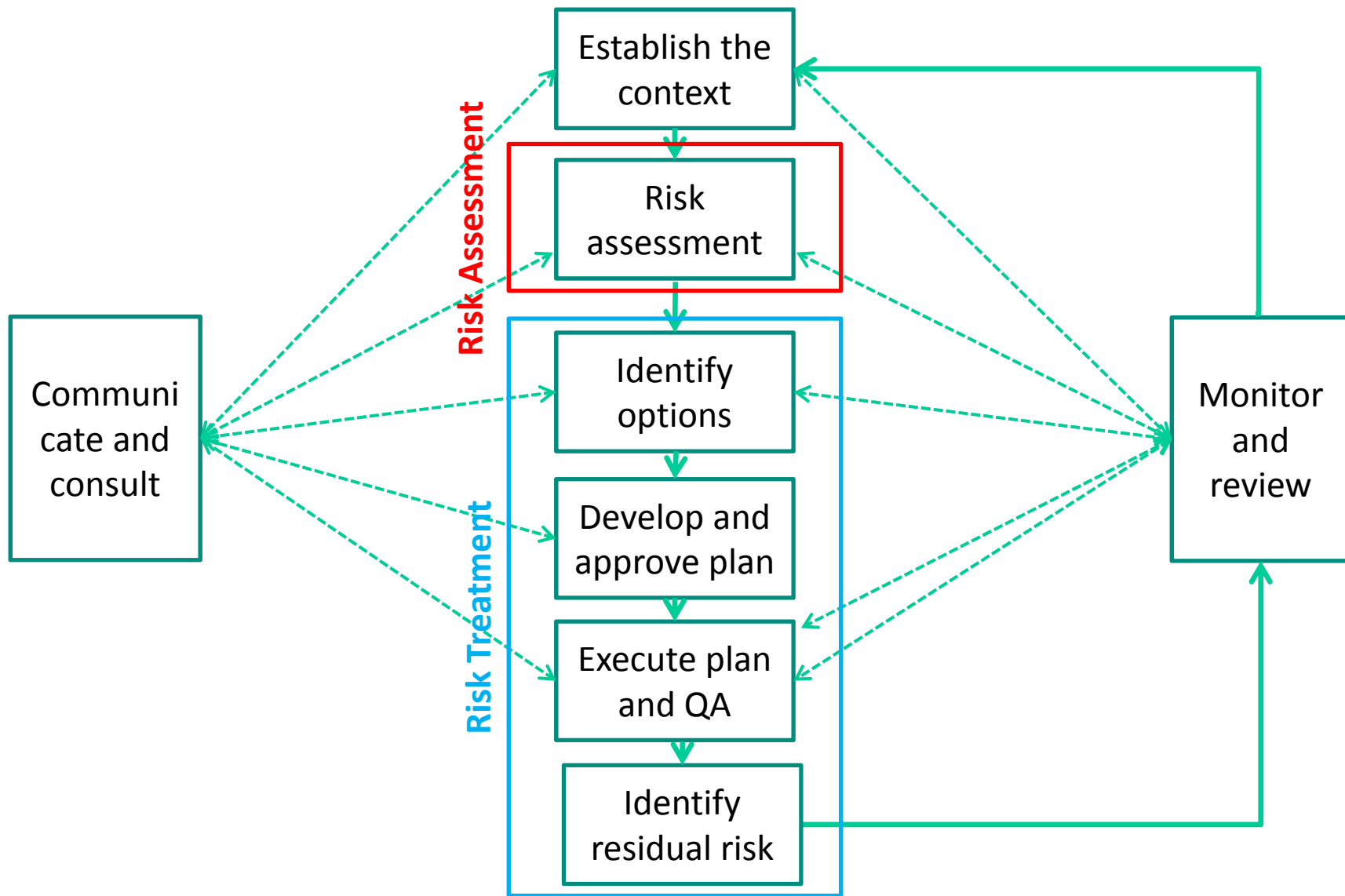




# Risk Management Process

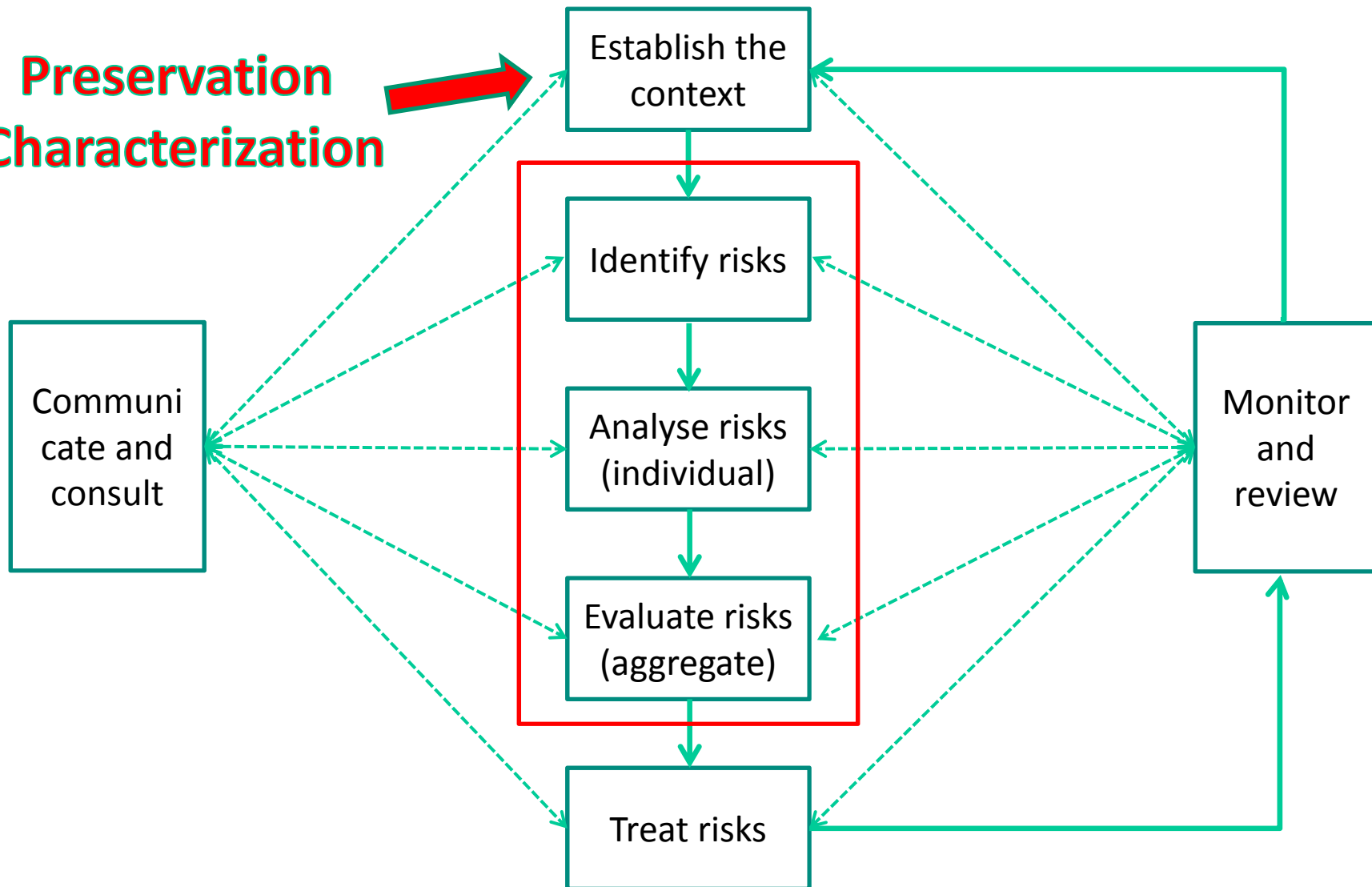


# Risk Management Process

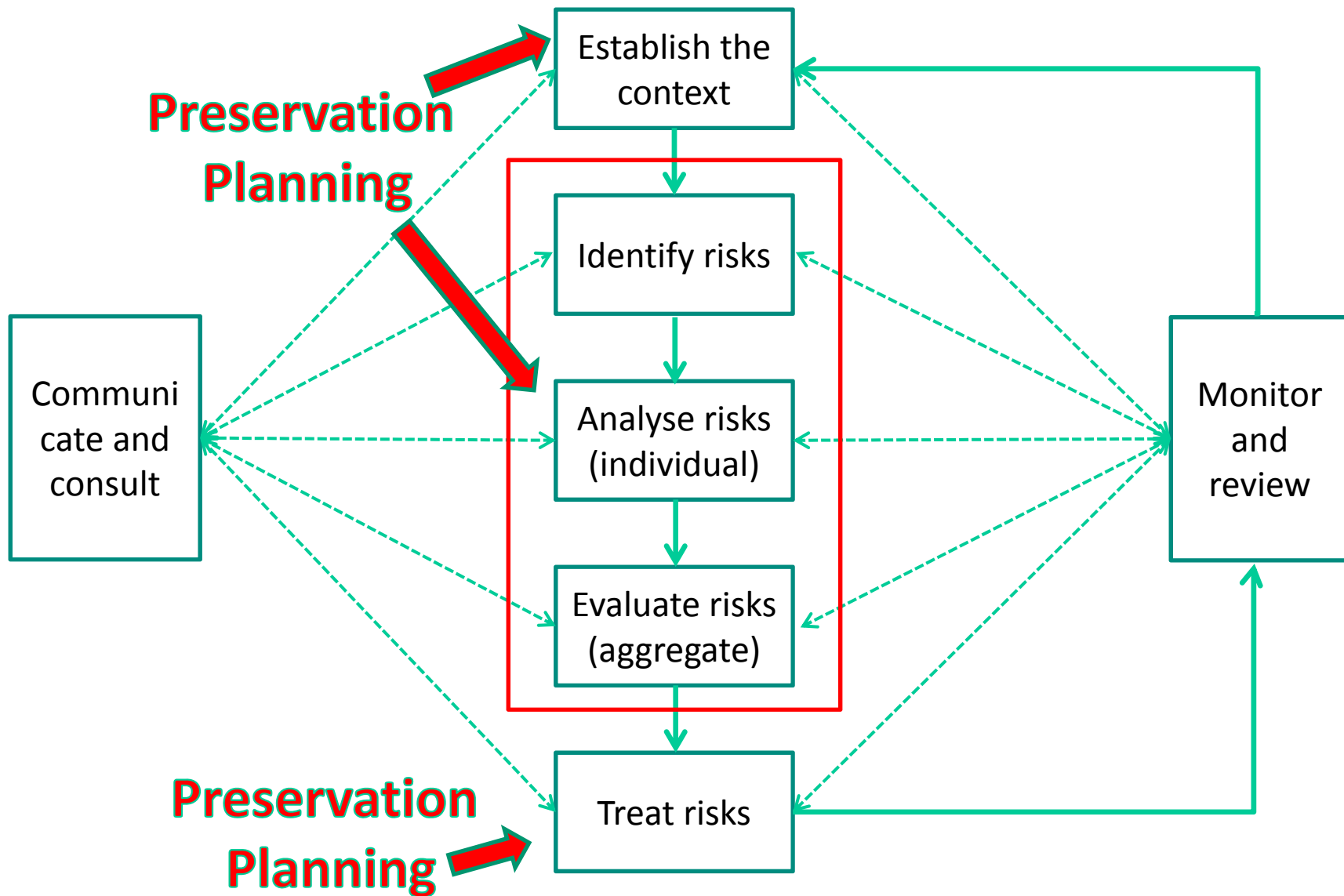


# Risk Management Process

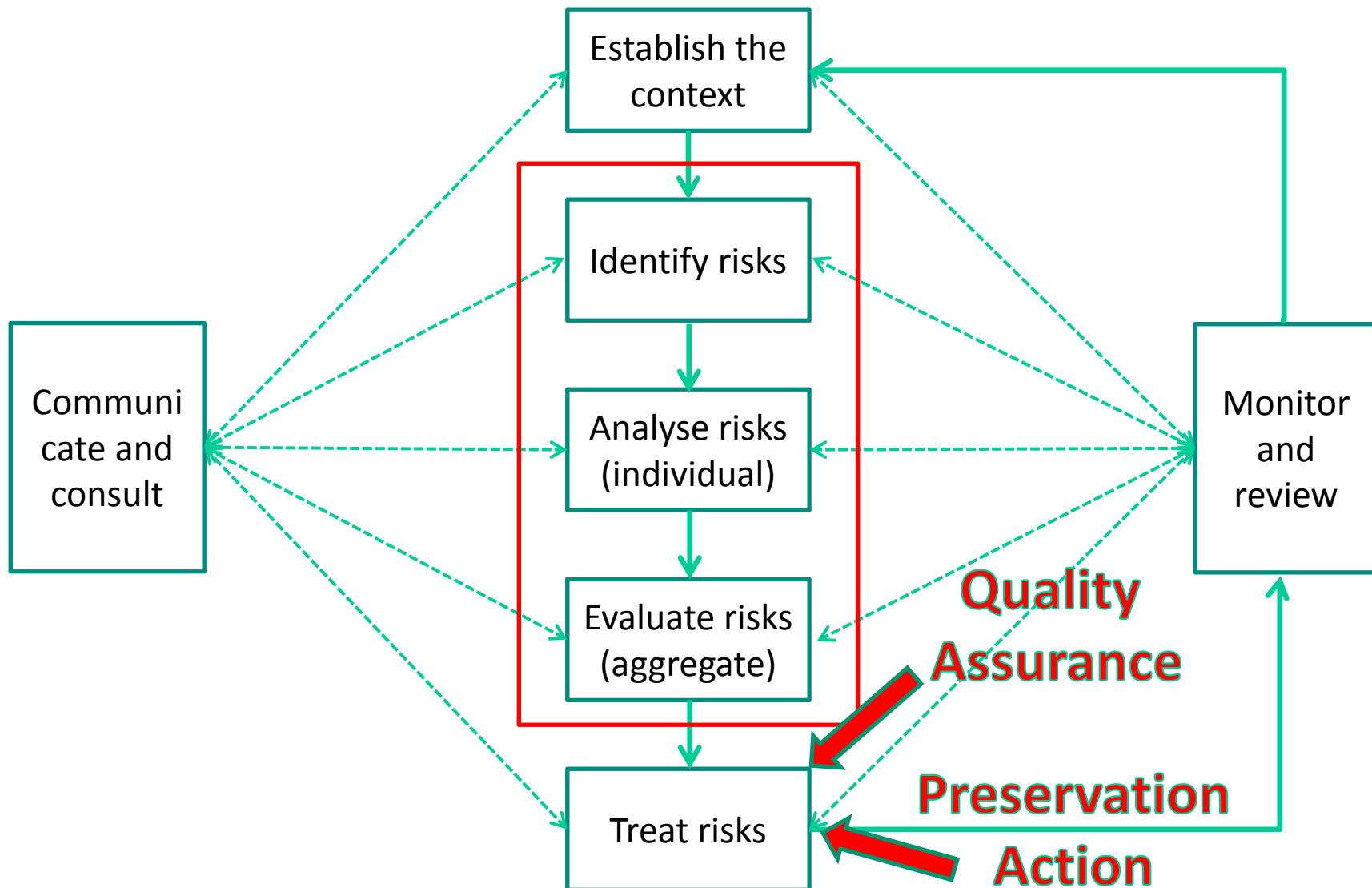
**Preservation  
Characterization**



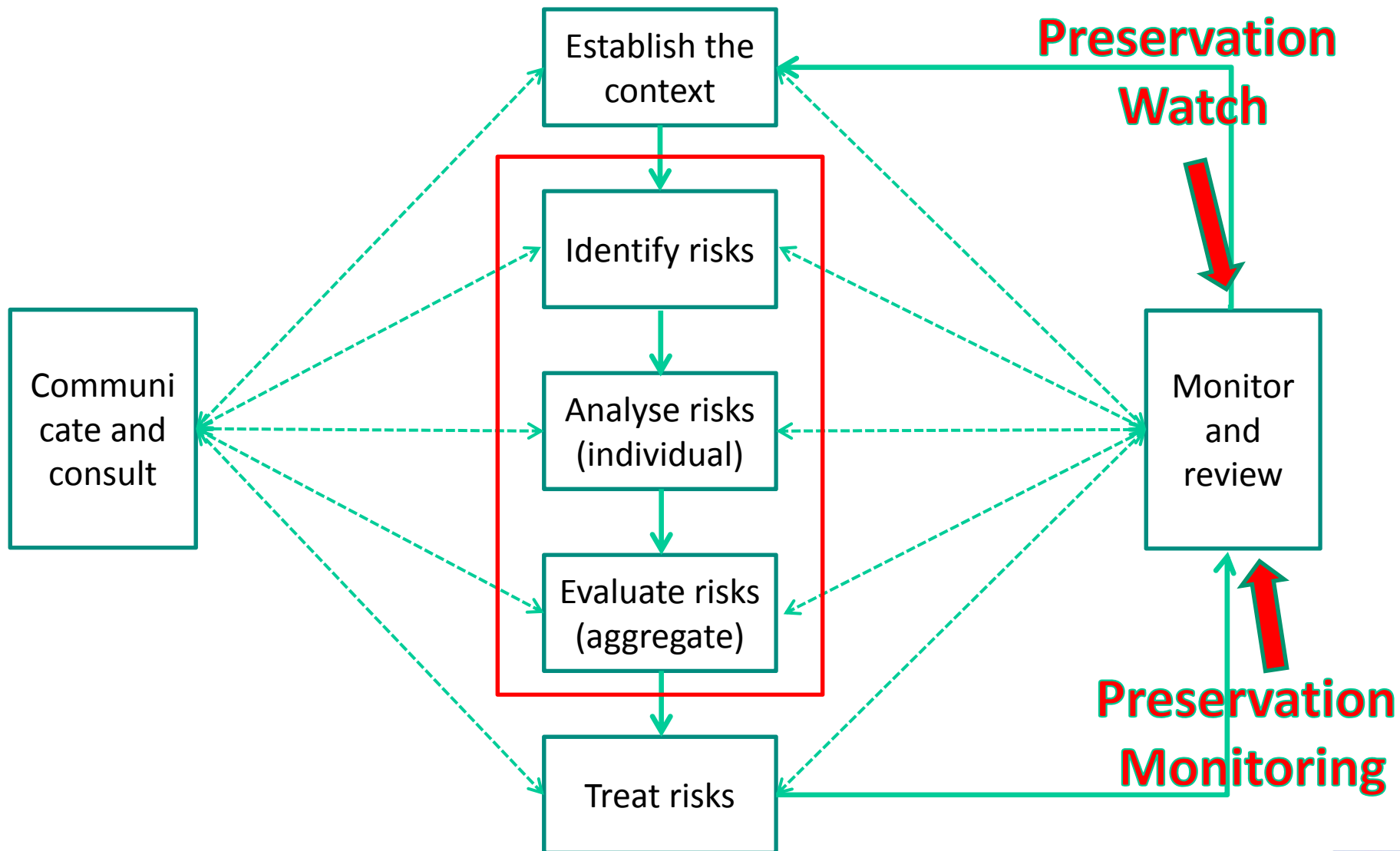
# Risk Management Process



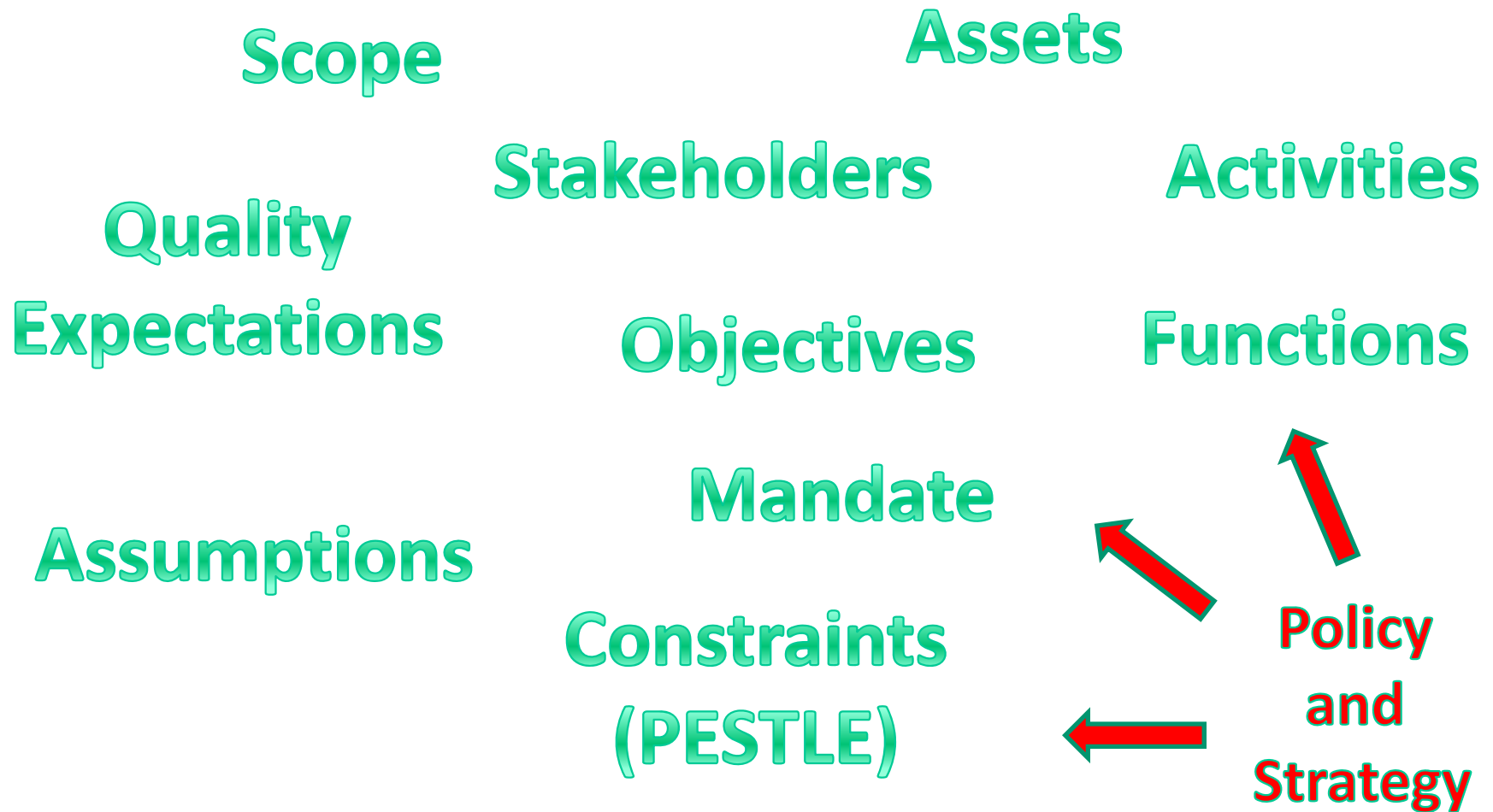
# Risk Management Process



# Risk Management Process



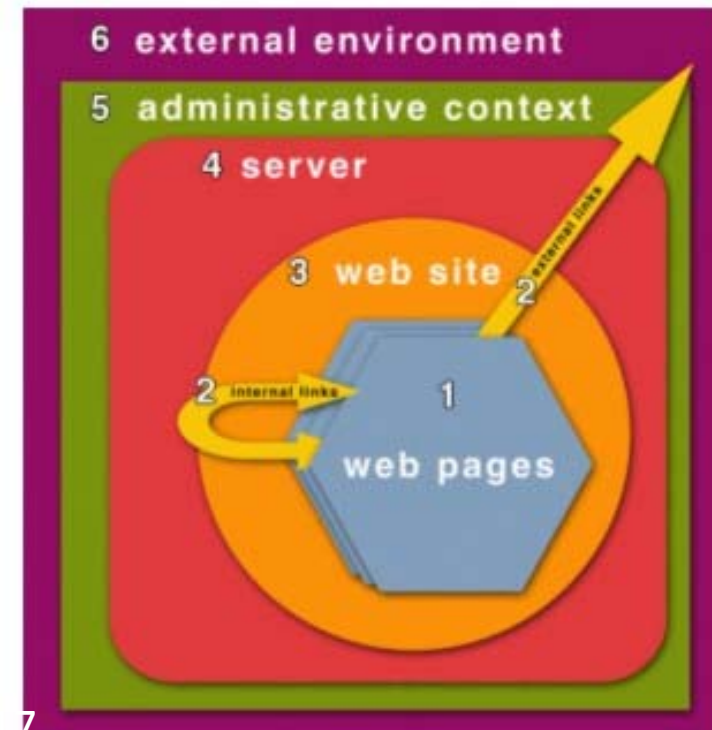
# Risk Context - Dimensions



# Risk Context: Scope

- A web page as a stand-alone object
- Considering the links into it and out from it
- A semantically coherent set of linked web pages
- A digital entity residing on a server
- A website as an entity within an administrative setting
- A website as part of an external environment

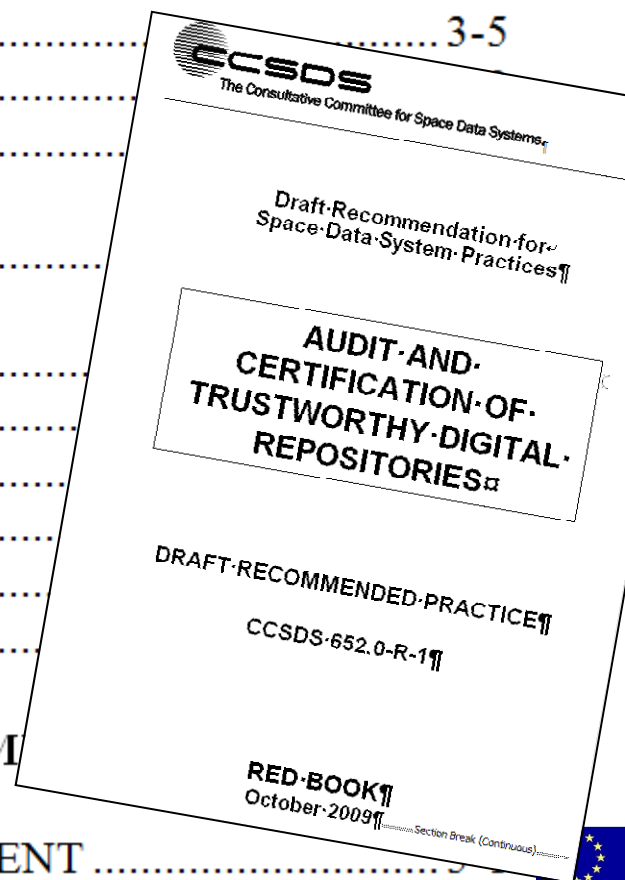
Virtual Remote  
Control for web  
archiving





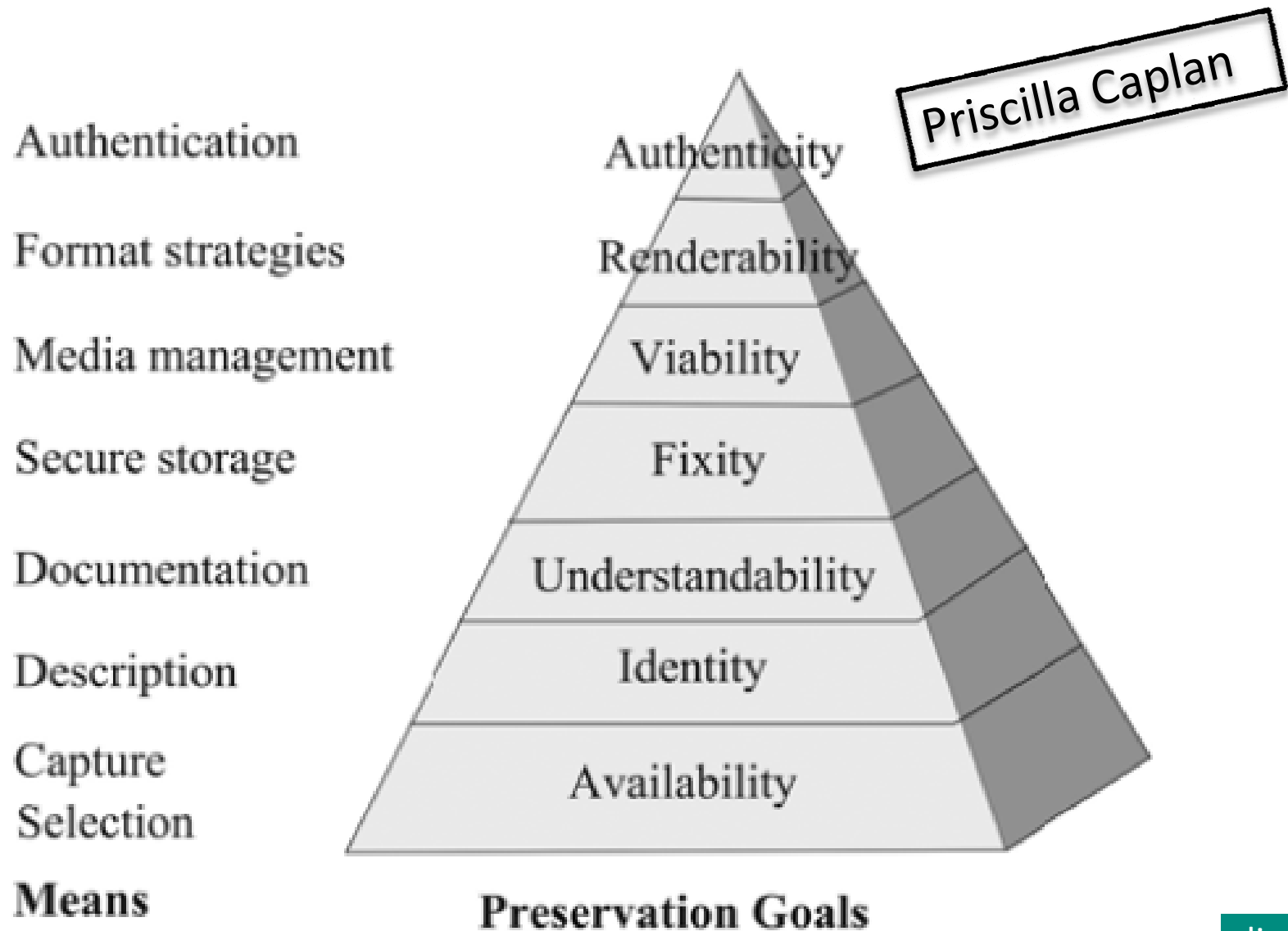
# The Context: The Bigger Scope

<b>3</b>	<b>ORGANIZATIONAL INFRASTRUCTURE .....</b>	<b>3-1</b>
3.1	GOVERNANCE & ORGANIZATIONAL VIABILITY .....	3-1
3.2	ORGANIZATIONAL STRUCTURE & STAFFING .....	3-3
3.3	PROCEDURAL ACCOUNTABILITY & PRESERVATION POLICY FRAMEWORK .....	3-5
3.4	FINANCIAL SUSTAINABILITY .....	
3.5	CONTRACTS, LICENSES, & LIABILITIES .....	
<b>4</b>	<b>DIGITAL OBJECT MANAGEMENT .....</b>	
4.1	INGEST: ACQUISITION OF CONTENT .....	
4.2	INGEST: CREATION OF THE AIP .....	
4.3	PRESERVATION PLANNING .....	
4.4	AIP PRESERVATION .....	
4.5	INFORMATION MANAGEMENT .....	
4.6	ACCESS MANAGEMENT .....	
<b>5</b>	<b>INFRASTRUCTURE AND SECURITY RISK MANAGEMENT .....</b>	
5.1	TECHNICAL INFRASTRUCTURE RISK MANAGEMENT .....	
5.2	SECURITY RISK MANAGEMENT .....	5-12



# The Context:

## Preservation Goals => Objectives



# The Context: Preservation Functions

DRAMBORA

An intellectual context for the work:

Commitment to digital object maintenance

Organisational fitness

Legal & regulatory legitimacy

Effective & efficient policies

Acquisition & ingest criteria

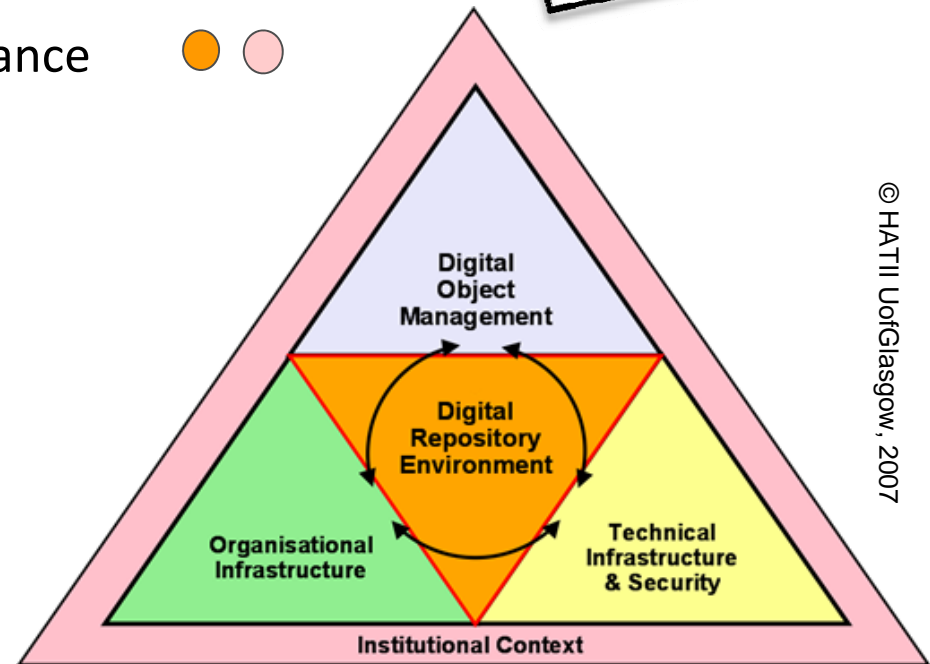
Integrity, authenticity & usability

Provenance

Dissemination

Preservation planning & action

Adequate technical infrastructure



*(CRL/OCLC/NESTOR/DCC/DPE meeting, January 2007)*



# Risk Identification: Breakdown Structures and Prompt Lists

- Technological
- Physical
- Organisational
- Socio-cultural
- Legal
- Economic
- Financial
- Political
- Contractual
- Environmental

DRAMBORA



# Risk Identification: Sources

Planets Project

New Version

Obsolescence

No Support

Unmanaged Growth

Deterioration

Loss

Access Inhibitors

Defects

New Requirements

**Risk sources**

File System

Operating System

Hardware

File Format

Software

Content

Representation Information

Data Carrier

Users

Legal or Statutory System

Budgets

**Digital  
Environment**



# Risk Identification: Vulnerabilities and Sources

José Barateiro,  
et al.

<b>Vulnerabilities</b>	Process	Software faults Software obsolescence
	Data	Media faults Media obsolescence
	Infrastructure	Hardware faults Hardware obsolescence Communication faults Network service failures
<b>Threats</b>	Disasters	Natural disasters Human operational errors
	Attacks	Internal attacks External attacks
	Management	Economic failures Organizational failures
	Legislation	Legislative changes Legal requirements

Table 1. Taxonomy of vulnerabilities and threats to digital preservation.

# Risk identification

Blake, TNA

## ORGANISATIONAL RISK

### R1. UNDERSTANDING & ACTION

Continuity risk is neither understood nor addressed cohesively at either the right levels or across the organisation (esp. IM, IT and IA responsibilities)

### R2. RISK GOVERNANCE

Continuity risk is not reflected in the risk management and information governance processes at either the right levels or across the organisation

### R3. INFORMATION VALUE

The Organisation does not understand the nature and value of its Information Assets enough to be able to apply Continuity risk management

## PROCESS RISK

### R4. IM SYSTEMS & PROCESSES

Existing, legacy or future IM systems and processes do not maintain Continuity to Information Assets over time or through change

### R5. IT SYSTEMS & PROCESSES

Existing, legacy or future IT systems and processes do not maintain Continuity to Information Assets over time or through technological change

### R6. BUSINESS SYSTEMS, STRUCTURES & PROCESSES

Existing, legacy or future organisational business systems, structures and processes do not maintain Continuity to Information Assets over time or through organisational change

## OPERATIONAL RISK

### R7. CONTEXT ABSENT

Required Information Context / Metadata is absent at creation / capture

### R8. CONTEXT NOT MAINTAINED

Information Context / Metadata is not maintained over time or through change

### R9. CONTENT & CONTEXT SEPARATED

Information Context / Metadata and Information Content (Data) are separated over time or through change

### R10. CONTEXT LOST

Information Context / Metadata is lost over time or through change

### R11. CONTENT LOST

Information Content (Data) is lost over time or through change

### R12. PROVENANCE NOT MAINTAINED

Provenance / Audit data about the Information Asset is not maintained over time or through change

### R13. OBSOLETE TECHNOLOGY

The Information Asset format cannot be accessed by available technology (infrastructure, platforms, applications etc)

### R14. TECHNOLOGY LOCK-IN

The Information Asset is locked in to a specific technology / vendor

### R15. ACCESS RESTRICTIONS

Management of encryption and file-level passwords for the Information Asset are not maintained

### R16. PREVENTING DISCOVERY

The Information Asset is hidden in a structurally complex digital format or an inaccessible location

### R17. INCOMPATIBLE TECHNOLOGY

Information Asset format cannot be used with the required functionality by available technology (infrastructure, platforms, applications)

### R18. INSUFFICIENT CONTEXT

There is insufficient Information Context / Metadata to understand the Information Asset

### R19. INSUFFICIENT PROVENANCE

There is insufficient audit / provenance data to trust the authenticity of the Information Asset

## CONTINUITY FAILURE

### R20. FAILURE IN THE INTEGRITY OF THE INFORMATION ASSET

Information is partial: missing crucial metadata, content or context

## CONTINUITY FAILURE

### R21. FAILURE IN THE AVAILABILITY OF THE INFORMATION ASSET

Information cannot be located or cannot be opened with available technology

## CONTINUITY FAILURE

### R22. FAILURE IN THE USABILITY OF THE INFORMATION ASSET

Information cannot be used as needed with the available technology, cannot be understood without its context or cannot be trusted as authentic

# Risk Analysis

Determine

Probability

Impact

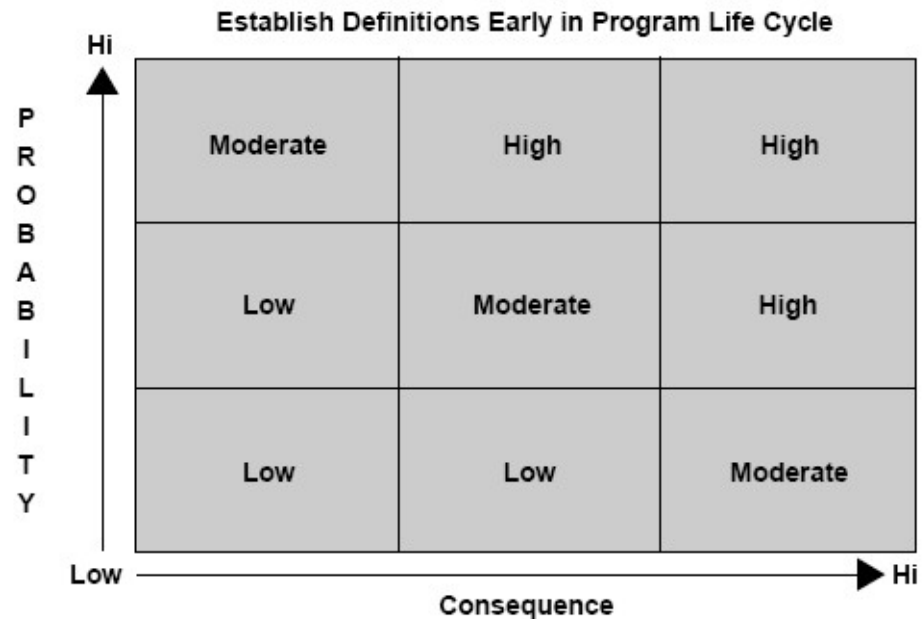
(Proximity)



Calculate severity



of the identified risks







# Factors Influencing Risk Impact

## Risk of loss

- Future rarity
- Alternative storage provision
- Heritage value

## Mandatory requirement

- Legal deposit obligation
- Existing external commitment

## Strategic considerations

## Opportunity & timing

- Size & rate of growth

## Opportunities for access

- Alternative access provision
- Revenue

## User need

- User demand
- Risk to physical collections
- Remote access



## Doability

- Effort
- Freely available

## Operational improvements



# Risk Impact Influenced by

Virtual Remote  
Control for web  
archiving

- relevancy to the organization's collection(s);
- significance (essential, desirable, ephemeral);
- archival role (primary archives for resource, informal agreement for full or partial capture, other);
- maintenance (key indicators of good site management);
- redundancy (captured by more than one archive);
- risk response (time delay and action based on test notifications);
- capture requirements (complexity of site structure, update cycle, MIME types, dynamic content, and behaviour indicators);
- size (number of pages, depth of crawl required, etc.).

# Risk Evaluation

- **Look at all risk as an aggregate**

Determine

Probability

Impact

(Proximity)



**Calculate severity**



**Identify need for action**

**Cost**

**Objectives**

**Policy and Strategy**

**Organisational risk  
threshold and appetite**





# Risk Treatment Options

## **Accept**

accept the potential risk

## **Reduce**

implement controls to lower probability or impact of the risk

## **Avoid**

eliminate the risk cause and/or consequence

## **Fallback**

Put in place alternative action for when the risk materializes

## **Transfer**

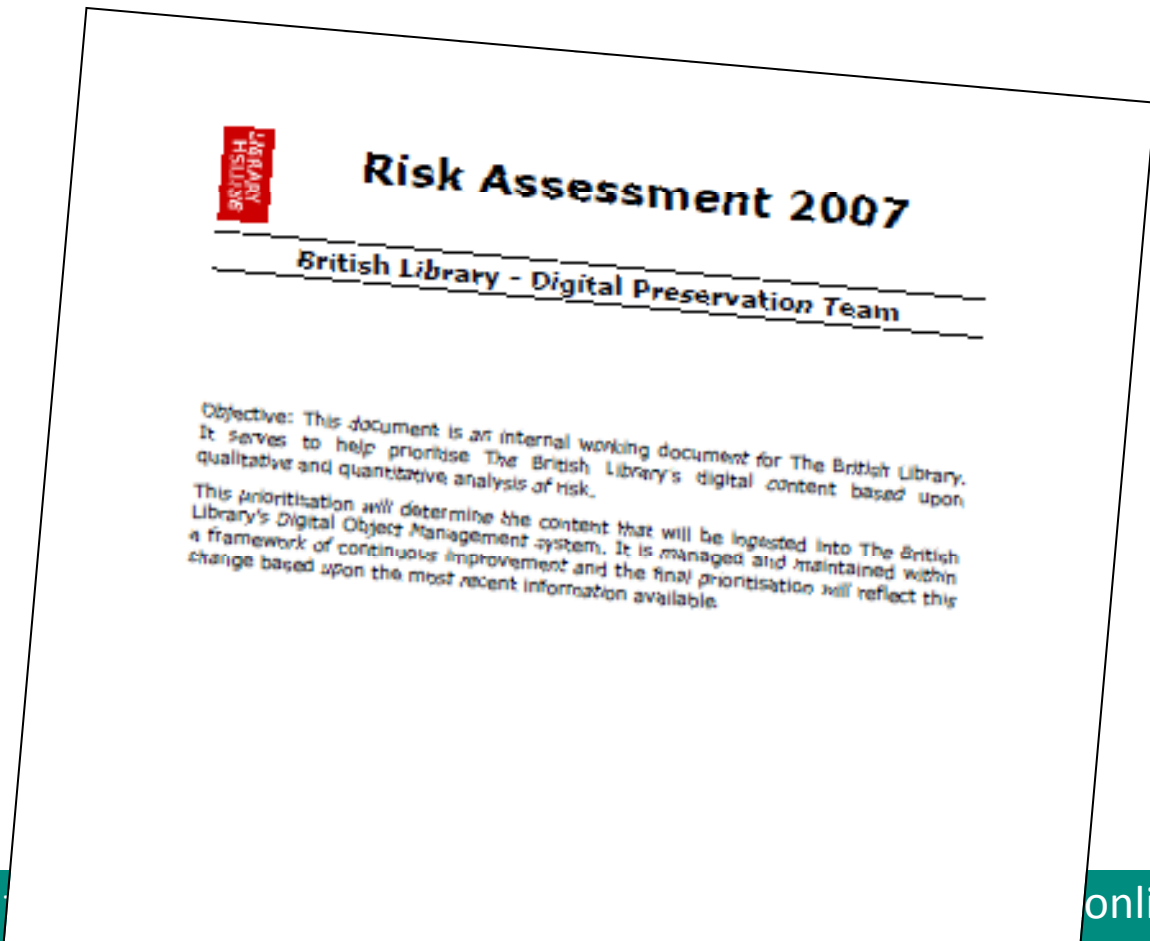
compensate for loss, such as purchasing insurance



Digital Preservation Coalition

# Example Risk Assessment: The British Library 2007

Available  
online





# How much information do we need?

Unsatisfactory storage	Bit stable storage	Content stable storage	Archival storage
Hand-held carriers	Images have been transferred on managed hard disk storage  Storage is backed up	Content has been QA'ed  Metadata has been produced and QA'ed  File formats have been identified  Representation Information has been deposited	Automatic check for corruption via checksums  Automatic replication over remote locations  Digital signatures  Integration with Primo / ILS





# Tools to Help

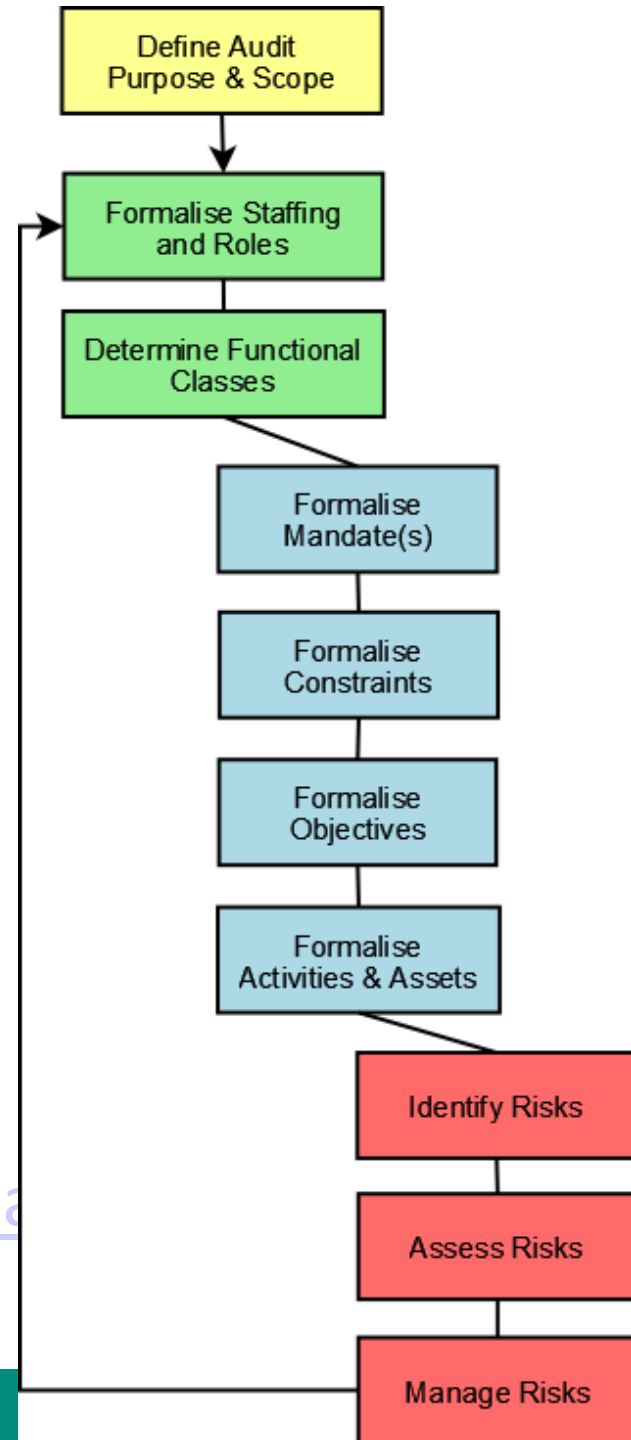
- Risk management:
  - Drambora (The Digital Repository Audit Method Based On Risk Assessment) : self-certification



# DRAMBORA

- Digital Repository Audit Method Based On Risk Assessment
- Online interactive tool
- Developed by the Digital Curation Centre (DCC) and Digital Preservation Europe (DPE)
- Identify, assess, manage, and mitigate risks
- Risk ontology

<http://blogs.ecs.soton.ac.uk/keepit/tag/drambora/>





[Register for DRAMBORA](#)

**Logged in:** Andrew McHugh

- Auditor
- Business Manager
- Data Liaison Officer

**at:** Florida Digital Archive at University of Florida

**Last Login:** 28 Nov 2008

[Log Out](#)

[Home](#)

[Online Help](#)

[User Admin](#)

[Before the Assessment](#)

[Assessment Centre](#)

[Report Results](#)

[Latest News](#)

[Get Expert Help](#)

[Download Offline Version](#)

[Submit Feedback](#)

[DRAMBORA Training](#)

[About](#)

[Objectives](#)

[Benefits](#)

[The DRAMBORA Team](#)

[Dissemination](#)

[DRAMBORA Users](#)

## DRAMBORA Online Tool :: Assessment Centre :: View Risk

[| Audit Home](#) | [| Mandate View](#) | [| Constraints View](#) | [| Objectives View](#) | [| Activities View](#) | [| Risks, Risk Assessment and Risk Management View](#) |

Use this page to navigate between the various related characteristics of this single risk. You can select alternative risks using the selection panel on the right hand side of the screen.

**Risk Name:**

**Identified\*:**

**Potential Impact\*:**

**Probability:**

**Severity:**

**Risk Description:**

**Risk Vulnerability:**

**Risk Relationships:**

<b>Nature of Risk:</b>	Physical Environment:	✗
	Personnel, Management & Admin Procedures:	✓
	Operations & Service Delivery:	✗
	Hardware, Software or Communications Equipmt & Facilities:	✗

**Risk Owner(s):**

**Functional Class(es):**

**Linked to :**

**Management Strategy(ies):**

### identified risks

- **Budgetary reduction** (Repository's operational budget is reduced)
- **Enforced cessation of repository operations** (Repository is forced to cease its business activities.)

### defined activities

### defined objectives

### defined constraints

### defined mandate

### assessment progress

### saved snapshots



# Tools to Help



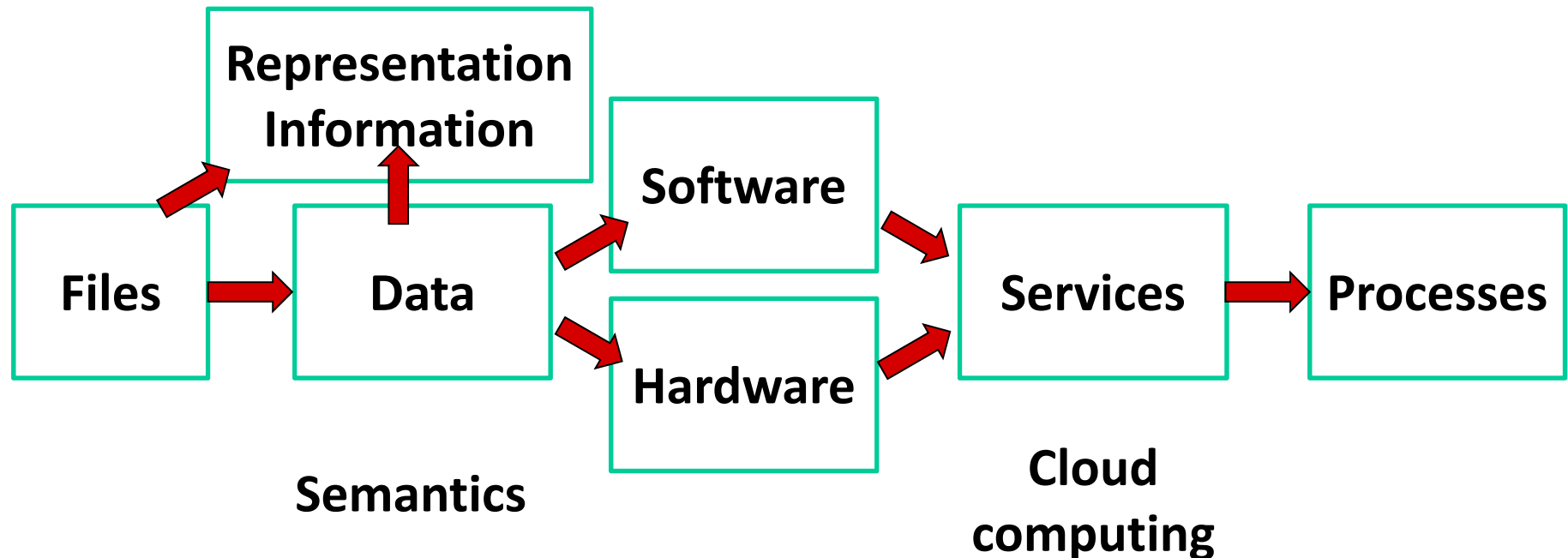
## Risk Management:

- Drambora (The Digital Repository Audit Method Based On Risk Assessment) : self-certification
- TIMBUS project: ERM (Enterprise Risk Management) tools extended to digital preservation

# TIMBUS

**Digital Preservation**

**Risk and Business Continuity  
Management**





# TIMBUS Task 4.1 ERM

- Intelligent Risk Management
  - Learning from previous situations
  - Reasoning from context
  - Automating risk detection and response
- Complete business modelling, including IT systems, legal constraints, etc.  
Rather than DP focus alone





# Tools to Help



## Risk management:

- Drambora (The Digital Repository Audit Method Based On Risk Assessment) : self-certification
- TIMBUS project: ERM (Enterprise Risk Management) tools extended to digital preservation
- TRAC/TDR: framework for establishing certified trustworthiness



#### 4.4.1 The repository shall have specifications for how the AIPs are stored down to the bit level.

##### Supporting Text

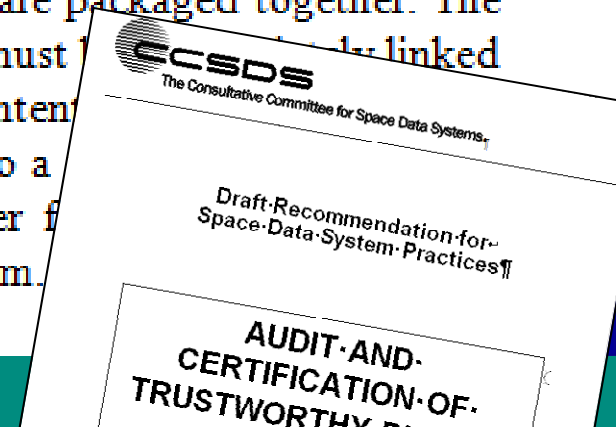
This is necessary in order to ensure that the information can be extracted from the AIP over the long-term.

##### Examples of Ways the Repository Can Demonstrate It Is Meeting This Requirement

Documentation of the format of AIPs; EAST and DEDSL descriptions of the data components (see references [B6] and [B7]).

##### Discussion

The repository should specify the Representation information down to the bit level of each AIP component and must specify how the separate components are packaged together. The Representation Information must be available for each AIP and must be linked to the AIP. Often, repositories are tempted to describe AIP content where a program will then be used to convert the information to a form usable by their Designated Communities. However, if those programs ever fail, the information would be lost in all the AIPs that relied on that program.



# Tools to Help

- Risk management:
  - Drambora (The Digital Repository Audit Method Based On Risk Assessment) : self-certification
  - TIMBUS project: ERM (Enterprise Risk Management) tools extended to digital preservation
  - TRAC/TDR: framework for establishing certified trustworthiness
- Context identification: DROID, JHOVE, FIDO, FITS, file, ...
  - Assess the characteristics of your digital assets
  - Profile your collections
- Risk treatment planning: Plato
- Risk treatment: A variety of preservation and QA tools



Digital**Preservation**Coalition



Thank you

