ICA
Africa Programme
Programme pour l'Afrique

ICA
Training Programme
Programme de Formation

Digital Records Curation Programme

Week 6:
Digital Preservation

Week 5 Recap

What did you learn?

- Class on Email Management
- Class on Cloud Computing

Learning outcomes:

At the end of this class you will be able to:

- identify the threats to the survival of digital records
- understand the strategies that can be used to mitigate those threats
- create digital asset registers
- know where to look for digital preservation tools
- apply the DPCMM to an organisation

Group Work – Legacy Hardware

- How would you retrieve and migrate any files on the hardware or storage media so that they could be viewed on a modern device?
- As a group, think through the issues, inspecting the objects and discussing the challenges
- Be ready to report back at the end

Terminology

Bit Preservation A term used to denote a very basic level of preservation of digital resource as it was submitted(literally preservation of the bits forming a digital resource). It may include maintaining onsite and offsite backup copies, virus checking, fixity-checking, and periodic refreshment to new storage media. Bit preservation is not digital preservation but it does provide one building block for the more complete set of digital preservation practices and processes that ensure the survival of digital content and also its usability, display, context and interpretation over time.

Digital Preservation Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation… refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological and organisational change.

Digital Preservation Handbook

Terminology

- Characterisation
- Normalisation (migration on ingest)
- Migration
- Emulation
- Refreshing
- Light, dim and dark archives
- TDR

Threats to Digital Objects

- Technology obsolescence
- File format obsolescence
- Human error
- Malicious activity
- Media decay, damage or loss
- Bit rot (changes to bits in the bitstream) / data degradation
- Hardware failure
- Network or service failure
- Software failure

Digital Asset Registers

- Digital Asset Registers should:
  - Identify digital assets requiring long-term access
  - Identify the threats to future accessibility
  - Quantify the risks of those threats materializing
  - Quanitify the costs and other impacts that would be incurred if the threats materialized
  - Quantify the benefits to be derived from continued access
  - Determine a priority for action

Digital Asset Registers

- Digital Asset Registers should include:
  - Name and description of the asset
  - Who is responsible for it
  - Type (database / website, etc)
  - Volume
  - Vulnerabilities
  - Risks of losing / benefits of keeping

Digital Asset Registers

- Working in groups, identify some digital assets you’ve encountered (minimum three).
- Complete the Digital Asset Register template, assigning risk values

Selecting File Formats

The National Archives (UK) advises consideration of:

- Ubiquity (subjective but widely known)
- Support (number of compatible programmes and their ubiquity)
- Disclosure (openness of technical specs)
- Documentation quality (detailed enough to recreate?)
- Stability (rarely changing, with new versions backwards compatible)
- Ease of identification and validation (availability of validation tools and preference for formats with file signatures and version information within the file structure)

http://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf

Selecting File Formats

- Intellectual Property Rights (over technologies used by the format, such as image compression algorithms)
- Metadata Support (does the format allow inclusion of metadata)
- Complexity (the more complex the format, the more difficult and expensive to preserve)
- Interoperability (platform independent and used across programmes)
- Viability (formats with error-detection facilities are preferred)
- Reusability (can the original functionality be maintained?)

http://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf

File Management

- Make ‘read only’ / use write-blockers
- Encryption / access control
- Digital forensics – BitCurator
- Persistent identifiers (vital metadata)

Management Activities

- Technology watch
- Understand the designated community
- Work with IT to understand and improve storage and backup arrangements
- Regular programme of fixity checking
- Policies and procedures
- Training and advocacy

COPTR (individual work)

- http://coptr.digipres.org/Main_Page
- Explore the COPTR digital preservation tool registry to get a sense of the range of tools that are available

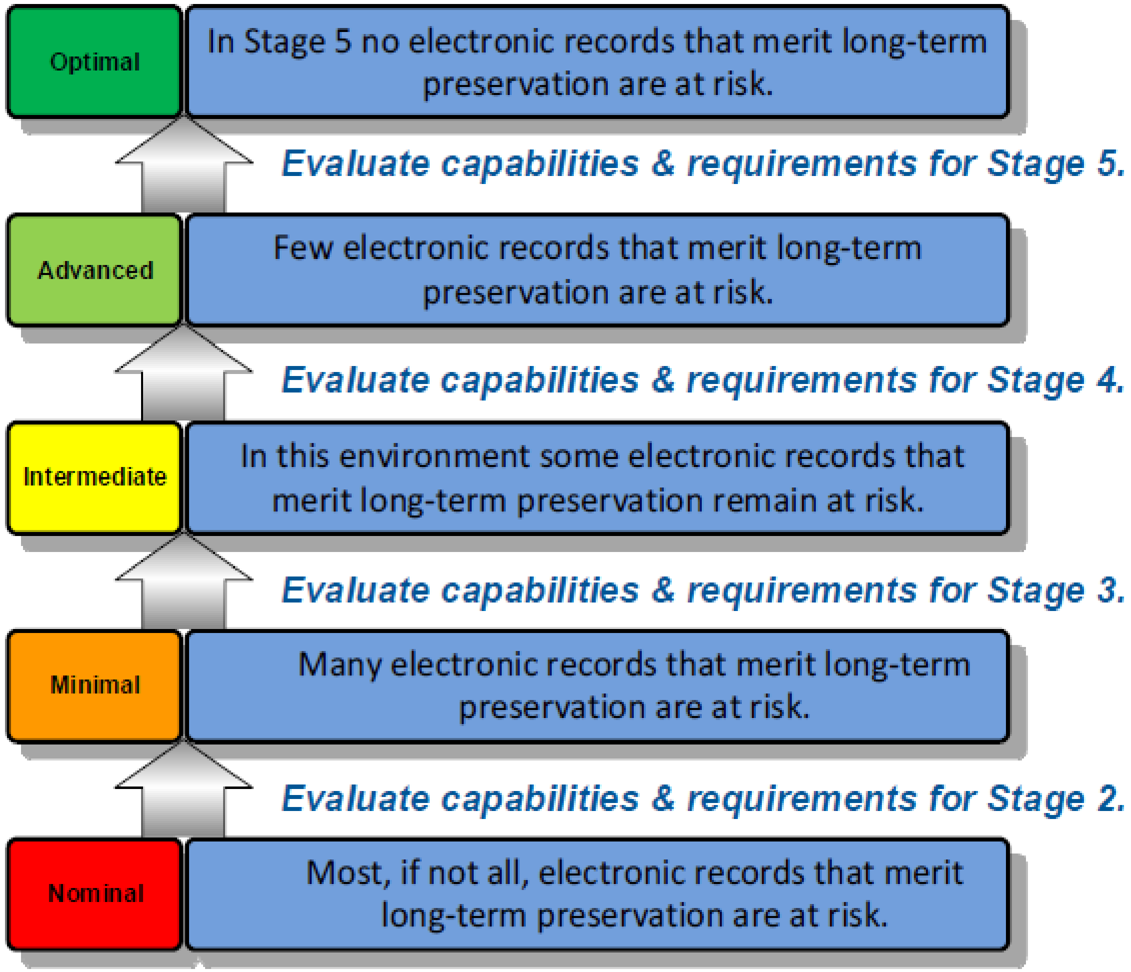
Assessing Digital Preservation Maturity

- Digital Preservation Capability Maturity Model (DPCMM) 2015
- A five level (or stage) maturity continuum.
- Based on the functional specifications of ISO 14721 (OAIS), the auditing criteria of TRAC and ISO 16363 (Audit and Certification) and accepted best practices in operational digital preservation repositories.
- Can be used for charting an evolutionary path from disorganised and undisciplined management of electronic records, or the lack of a systematic digital continuity approach, into increasingly mature stages of digital preservation capability.

Stages of the Digital Preservation Capability Maturity Model

A maturity model is a set of structured levels that describe how well the practices, processes and behavior of an organization can reliability and sustainably produce desired outcomes. The Digital Preservation Capability Maturity Model© (DPCMM)[4] is a five level (or stage) maturity continuum.

In Stage 1 a systematic electronic records management and/or digital preservation program has not yet been undertaken or a digital preservation program exists only on paper. The highest level (Stage 5) represents sustainable digital preservation capabilities systematically managed by process optimization and continuous process improvement. A high level description of key characteristics of each stage is provided on the next few pages.



Optimal

In Stage 5 no electronic records that merit long-term preservation are at risk.

Evaluate capabilities & requirements for Stage 5.

Advanced

Few electronic records that merit long-term preservation are at risk.

Evaluate capabilities & requirements for Stage 4.

Intermediate

In this environment some electronic records that merit long-term preservation remain at risk.

Evaluate capabilities & requirements for Stage 3.

Minimal

Many electronic records that merit long-term preservation are at risk.

Evaluate capabilities & requirements for Stage 2.

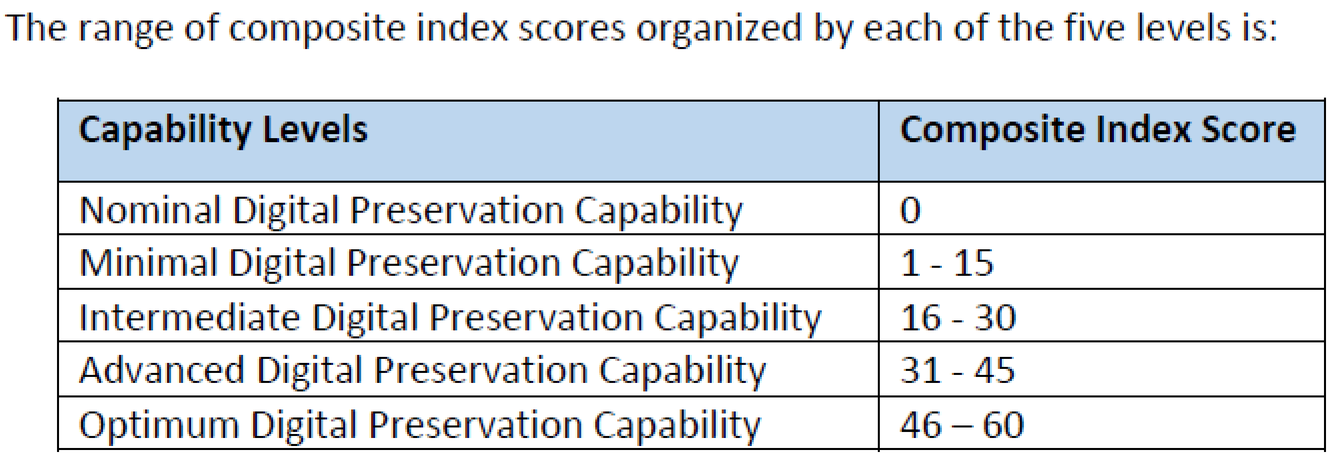
Nominal

Most, if not all, electronic records that merit long-term preservation are at risk.

Image © *Digital Preservation Capability Maturity Model* 2014

DPCMM

- In groups, apply the DPCMM to an organisation that one of you has worked for.
- Generate a score.
- How does this provide the basis for planning a digital preservation strategy?



The range of composite index scores organized by each of the five levels is:

Capability Levels	Composite Index Score
Nominal Digital Preservation Capability	0
Minimal Digital Preservation Capability	1 - 15
Intermediate Digital Preservation Capability	16 - 30
Advanced Digital Preservation Capability	31 - 45
Optimum Digital Preservation Capability	46 – 60

Image © *Digital Preservation Capability Maturity Model* 2014

Conclusion

- Threats to digital preservation
- Role of digital asset registers in digital preservation
- Terminology used and actions required
- File formats and managing them over time
- Digital preservation capacity and maturity models

Any questions?

PCOM ICA

Programme Commission

Commission du Programme