



Significant Properties of Digital Objects

Andrew Wilson
National Archives of Australia

Topics

- ◆ background
- ◆ digital preservation
- ◆ authenticity
- ◆ significant properties
- ◆ initiatives
- ◆ closing thoughts

Background

- ◆ the nature of digital objects
 - mediated by technology
 - rapid technological obsolescence
 - media fragility
 - active intervention needed
- ◆ what does digital preservation mean?

An observation

"traditionally, preserving things meant keeping them unchanged; however our digital environment has fundamentally changed our concept of preservation requirements. If we hold on to digital information without modifications, accessing the information will become increasingly difficult, if not impossible"

Su-Sing Chen, "The Paradox of Digital Preservation", *Computer*, March 2001, 2-6.

Digital preservation

- ◆ fundamental challenge is to preserve the accessibility and authenticity of digital objects over time and domains, and across changing technical environments
- ◆ we must accept:
 - inevitability of change
 - separation of logical information object from its physical environment

Digital preservation approaches

- ◆ techno-centric: keeping and maintaining the object's original hardware and software
- ◆ data-centric: maintaining objects in current data formats
- ◆ process-centric: creating new processes to render the original object
- ◆ post-hoc: digital archaeology or forensics

Data-centric approaches

- ◆ data centric approaches aim at keeping the data accessible and useable (performance based)
- ◆ migration
 - migration at obsolescence
 - format normalisation
 - migration on request

What is authenticity?

- ◆ authenticity requires and is derived from:
 - integrity/accuracy (no unauthorised changes)
 - reliability (it is what it says it is)
 - usability (it can be retrieved and rendered)

Why is authenticity important?

- ◆ users must trust that the record is what it purports to be
- ◆ preservation approaches that are data-centric change the original object
- ◆ Q: how are users persuaded to accept the changed object as authentic?
- ◆ A: establish its identity and integrity

Another quote

“This does not mean that a record must be precisely the same as it was when first created for its integrity to exist and be demonstrated. A record is considered to be essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose is unaltered”

The National Archives (UK), *Defining the characteristics for authentic records*, 2006

[http://www.nationalarchives.gov.uk/documents/generic_reqs1.pdf]

Assuring authenticity

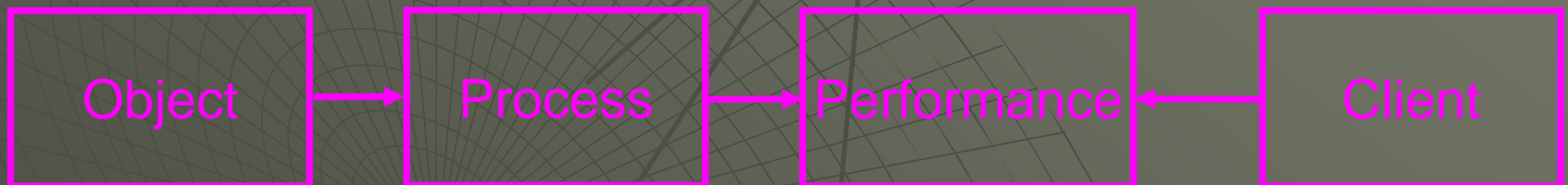
- ◆ what is sufficient to persuade users that a changed object is 'authentic':
 - appearance ('look and feel')?
 - content?
 - context?
 - structure?
 - behaviour?

Conceptual model

- ◆ National Archives of Australia developed, in 2002, a conceptual model that allows us to understand how changed digital objects can remain authentic
- ◆ the 'performance model'

Digital performance

- ◆ object is experienced as a performance:
 - data + (software + hardware) = digital object



- ◆ preserve the performance to preserve the object

Essence=Significant Properties

- ◆ in the NAA model the performance is preserved by ensuring that the 'essence' of the digital record is preserved
- ◆ 'essence' is the same thing we talk about as 'significant properties' (or 'characteristics')

Definition: significant properties

the characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record.

Significant properties

- ◆ not every attribute of the original object
- ◆ different rules for different genres and contexts
- ◆ much work needed to establish the 'significant properties' of objects in a wide range of digital data formats

Typology of significant properties

- ◆ **content:**
that which actually conveys information, not necessarily human readable, eg. text, image, slides, programming code, etc.
- ◆ **context:**
background information that enhances understanding of technical and business environments to which the digital objects relate, and the provenance (creator and subsequent changes in custody and ownership) of the object, eg. who, when, why.
- ◆ **appearance:**
how the content of the object appears to an agent interacting with it, eg. font and size, colour, layout, etc.
- ◆ **structure:**
the arrangement of component parts of the content of the object and how they relate to each other, eg embedded files, pagination, headings, etc.
- ◆ **behaviour:**
functionality that is intrinsic to an object, eg. hypertext links, updating calculations, active links, etc.

Closing thoughts

- ◆ is the concept of 'significant properties' useful and viable?
- ◆ how can we actually ascertain the significant properties of a multitude of format categories?
- ◆ does the view of what forms the significant properties of an object depend on end use?
 - if so what are the consequences?

