

Significance is in the Eye of the Stakeholder

Angela Dappert, Adam Farquhar

The British Library

30 September 2009

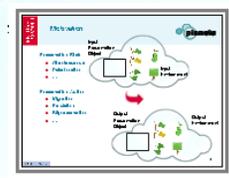
Context



Input
Preservation
Object

Preservation Risk

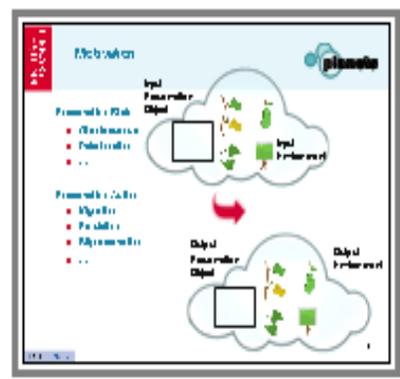
- Obsolescence
- Deterioration
- ...



Input
Environment

Preservation

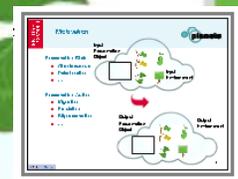
- Migration
- Emulation
- Bit-p...
- ...



Output
Preservation
Object



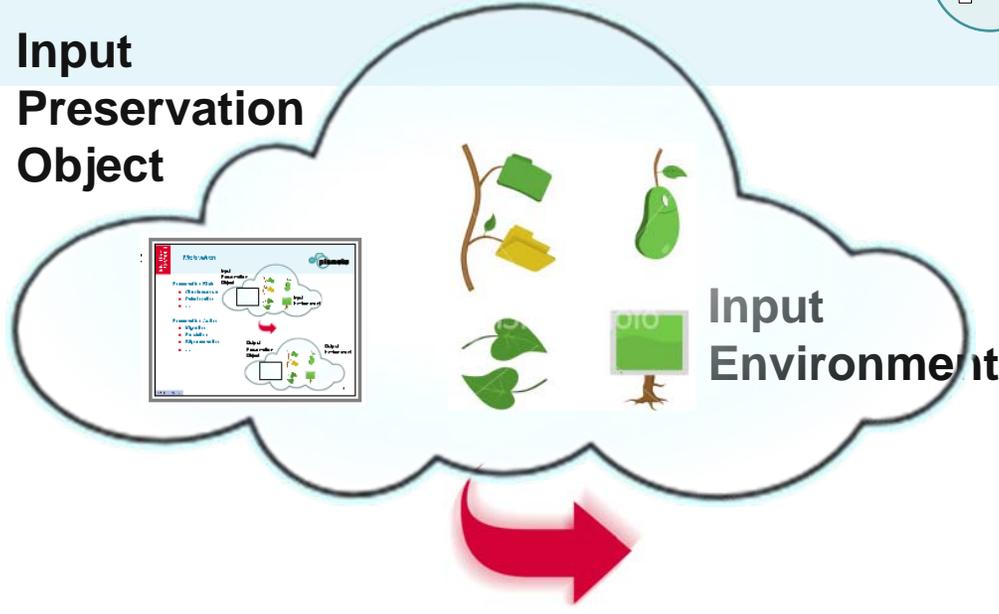
Output
Environment



Context



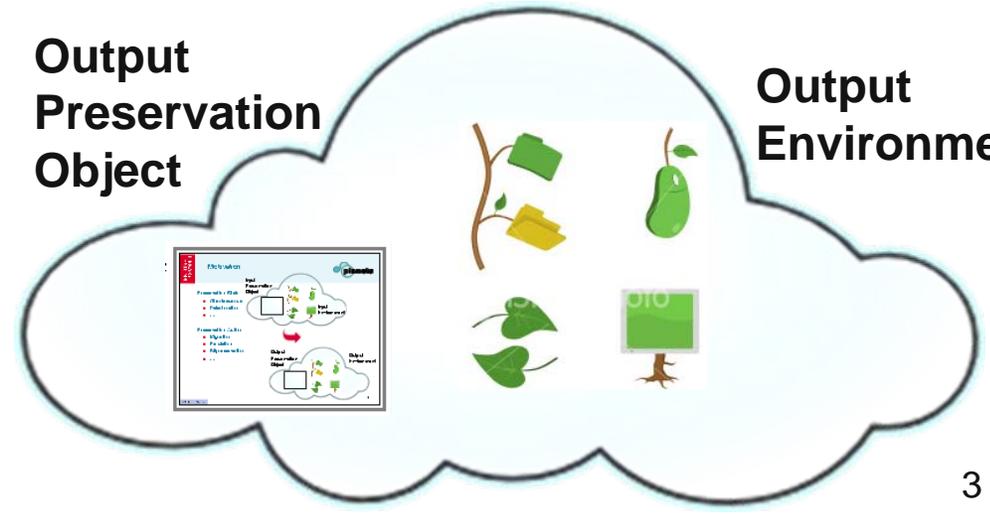
Input
Preservation
Object



Input
Environment

Can they be
the same?
What should
be the same?

Output
Preservation
Object



Output
Environment

Significant Characteristics – Working Definition 1



“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.”

Andrew Wilson: Significant Characteristic

The term significant characteristic has become over-loaded and remains ill-defined.

⇒ **Communication is hampered**

⇒ **The current interpretations do not meet the needs of content custodians**

➔ **Investigate real-life significant characteristics.**

What or Who Determines Significance?



The essence of the digital object needs to be preserved.

Merely circumstantial characteristics need not be preserved.

Is essence inherent or not?

=> Colour is not inherently essential

Example:

• ~~circumstantial~~ **circumstantial**

Account Holder	Account Balance
Anna Zamfir	300.00
Bart Yusef	700.00
Chloe Xavier	400.00
Total	1400.00

What or Who Determines Significance?



- **Stakeholders** (producer, custodian, consumer) involved in the preservation process attribute significance in a **context**.
- Captured as **requirements** (“business rules”)
- Requirements influence preservation processes.

Significant Characteristics – Working Definition 2



“Requirements in a specific context, expressing ~~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.”

Whose Characteristics Need to be Preserved?



Action	Digital Object		Characteristic	
bit-preservation	Preservation Object	files		file size
		representations		number of files

Whose Characteristics Need to be Preserved?



Action	Digital Object		Characteristic	
migration preserving semantics, possibly the look and feel	Preservation Object	files	file size	
		representations	number of files	
		intellectual entities	journal article	page count
			video	frame rate
		components	text component	alignment
			any component	semantic interpretation

Whose Characteristics Need to be Preserved?



Action	Digital Object		Characteristic	
emulation and migration preserving the look and feel	Preservation Object	files	file size	
		representations	number of files	
		intellectual entities	journal article	page count
			video	frame rate
		components	text component	alignment
			any component	semantic interpretation
	Env't	environment	software	memory usage
			hardware	screen resolution

Significant Characteristics – Working Definition 3



“Requirements in a specific context, expressing **a combination of characteristics of digital objects** ~~of preservation objects or environments~~ 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.”

What Form do Significant Characteristics Take?



Examples:

- If the component type is “text” then the font size of the component needs to be preserved.
 - The relative size of all images in the output document needs to be the same as in the input document.
-
- Significant characteristics are **not simple properties or property/value pairs.**
 - They are requirements which take the form of rich **constraints.**

Are All Significant Characteristics Created Equal?



Examples:

- The output line thickness has to be +/-1 point of the input line thickness.
 - The character sequence must remain unaltered, but introducing hyphenation can be tolerated.
-
- Specify **which characteristics** are significant.
 - Specify **to what degree** deviation from the required value can be tolerated.
 - Specify a measure of the **relative significance** of the characteristic for the stakeholders (importance factor).

Significant Characteristics – Working Definition 4



“Requirements in a specific context, **represented as constraints**, expressing a combination of characteristics of preservation objects or environments 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.”

Can we Introduce Improvements?



Examples:

- **Normalization**
 - To reduce the variety of formats held, or
 - To improve characteristics in the original.
 - Migrate files which are in formats that are susceptible to degradation to files in a more robust, resilient format
 - Move static tables to spreadsheets which enable pivot tables.
- **Manual restoration of a file by a curator to the state it was presumed to have had before a corruption.**

Significant Characteristics – Working Definition 5



“Requirements in a specific context, represented as constraints, expressing a combination of characteristics of preservation objects or environments that must be preserved **or attained** 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.”

Implication – Different Preservation Actions are Supported

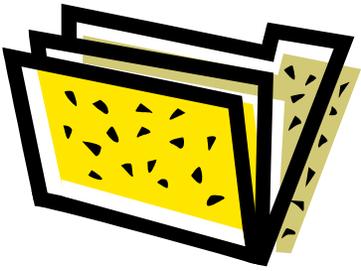
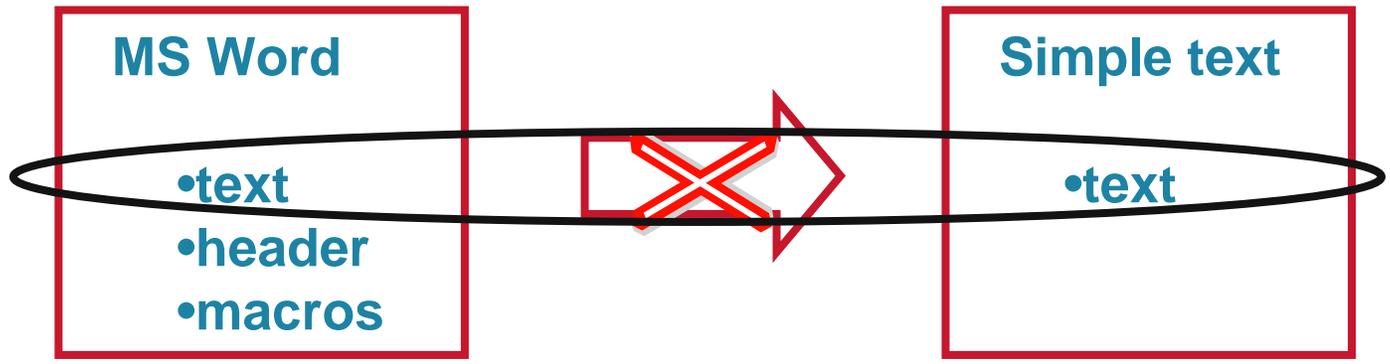


Action	Digital Object		Characteristic	
bit-preservation	Preservation Object	files	file size	
		representations	number of files	
intellectual entities		journal article	page count	
		video	frame rate	
migration preserving semantics, possibly the look and feel		components	text component	alignment
			any component	semantic interpretation
emulation and migration preserving the look and feel	Env't	software	memory usage	
		hardware	screen resolution	

Implication – Significance is not Inherent to a File Format



Registries of file formats and **applicable** properties



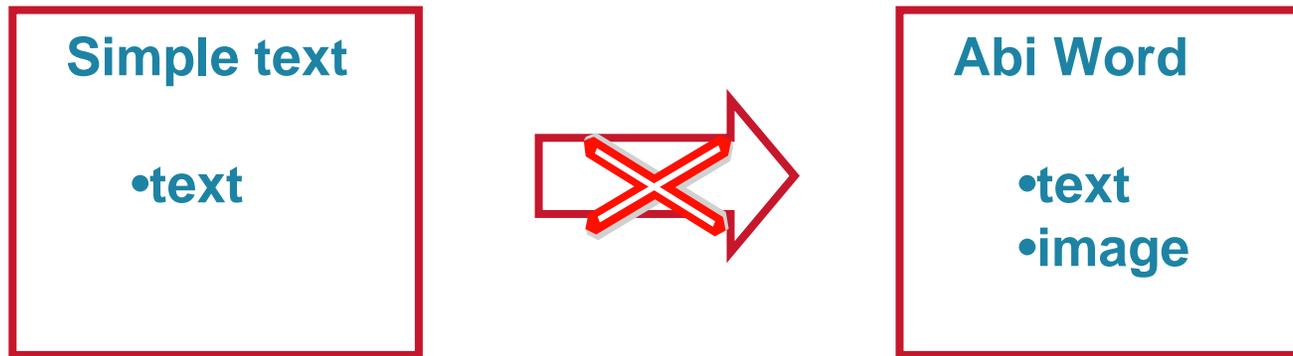
This file contains only text:
Only text characteristics are
significant

Increase the set of appropriate actions

Implication – Significance is not Inherent to a File Format



Registries of file formats and **applicable** properties



Significant characteristic:
Abi Word is not supported for this institution.

Decrease the set of appropriate actions

Representation Information

- is the set of **important characteristics** of a **data object**, including its **environment**.
- is needed to make sense of it for a given designated community at a given time (**context**).
- Its purpose is to **guarantee the accessibility, usability, and meaning** of preservation objects.

Representation Information does **not** specify

- constraints for transformations over time.
- characteristics of an acceptable derived data object.
- requirements for preservation actions.

Example:

- **The software package for a data object's proper rendering must be known to be able to use it,**

**but need not be used
by the corresponding information object
after a migration**

Conclusion



- **Significant characteristics can be considered one specific form of preservation guiding requirement.**
- **Significance is in the hands of stakeholders**
 - not inherent in a preservation object or file format.
- **The domain of significant characteristics extends beyond preservation objects to include environments.**
- **Significant characteristics require a rich constraint language.**
- **Consequences for implementation of**
 - **preservation metadata dictionaries,**
 - **property registries, and**
 - **preservation services.**