



a centre of expertise in data curation and preservation

## Significant Properties – what are they and why should I care?

Stephen Grace & Gareth Knight  
Centre for e-Research  
King's College London

DCC101, 7 October 2008



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 UK: Scotland License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/2.5/scotland/>; or, (b) send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Funded by:



Digital Curation 101, October 6<sup>th</sup>-10<sup>th</sup>, 2008, NeSC, Edinburgh

## Overview of exercise

- Outline of Significant Properties (SPs)
- SPs for emails
- Exercise: identify some SPs for images
- Discussion
- Characterisation using JHOVE
- Summary



a centre of expertise in data curation and preservation

## Many definitions...

**“The characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability and meaning of the objects”**

Wilson, 2007

“Significant properties are those properties of digital objects that affect their quality, usability, rendering, and behaviour.”

Hedstrom & Lee, 2002

“Those characteristics (both technical, intellectual, and aesthetic) agreed by the archive or by the collection manager to be the most important features to preserve over time.”

The Cedars Project Report, 2001

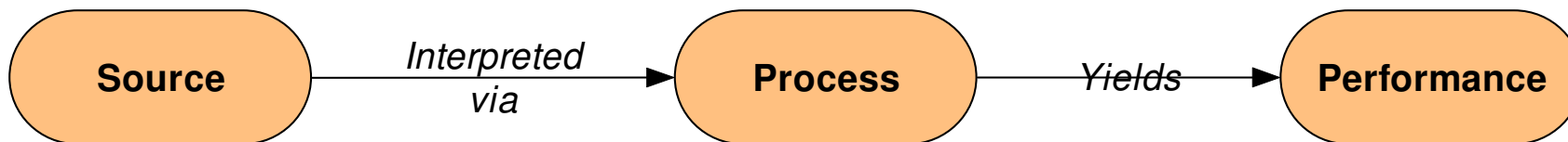
Closely tied to notions of *authenticity* and *integrity*

## Five categories of SPs

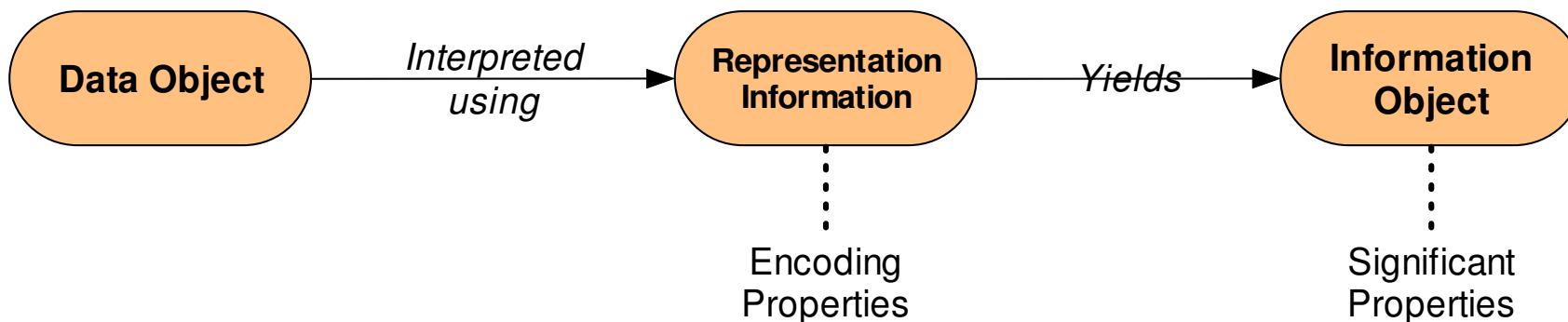
- Content
- Context
- Rendering
- Structure
- Behaviour

# Interpreting SPs in abstract

## NAA Performance Model



## OAIS Reference Model

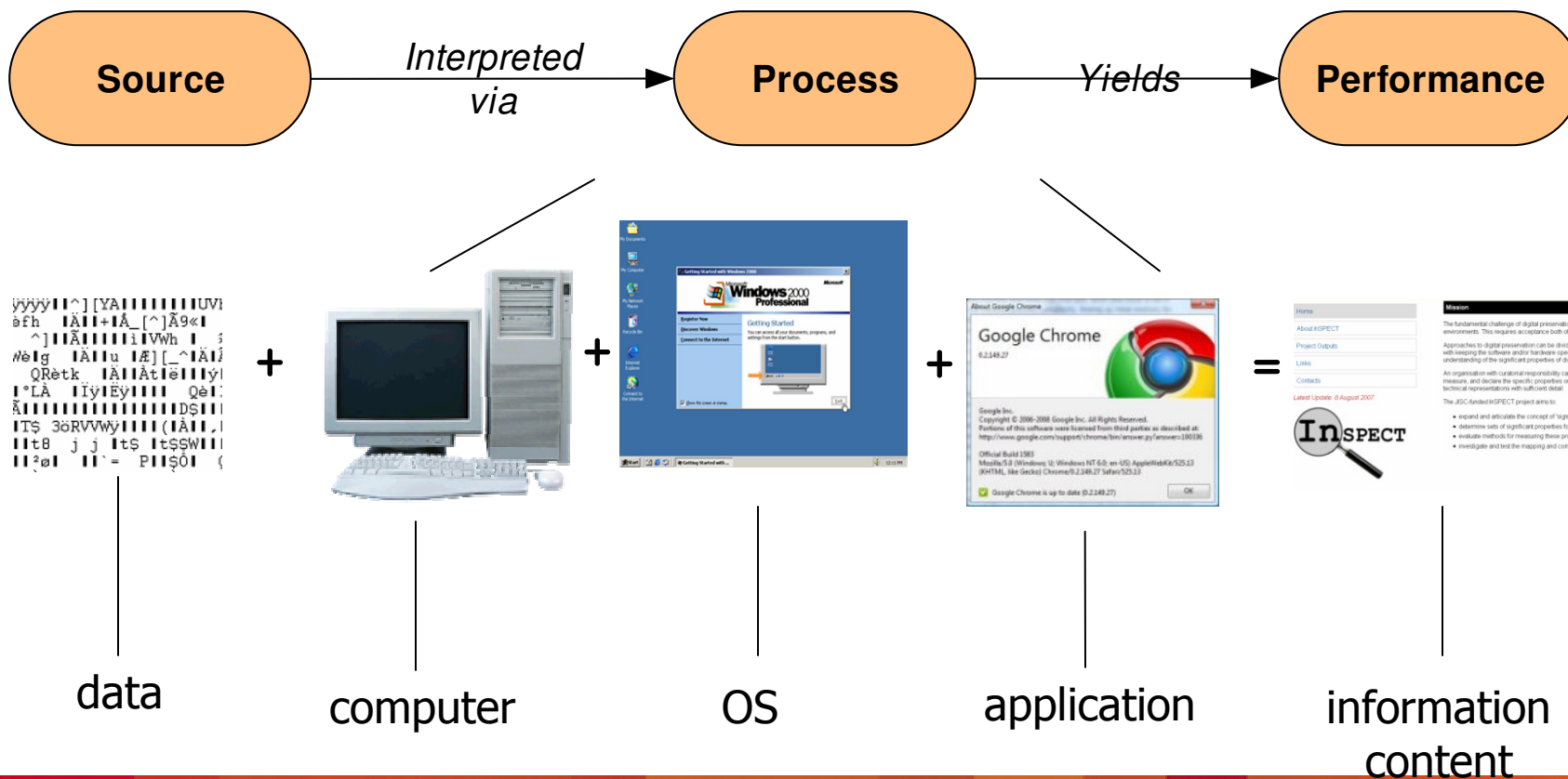


# Significant Properties and Representation Information

- Representation properties describe the data object
- Significant properties describe the information object

# Interpreting SPs in practice

## NAA Performance Model





a centre of expertise in data curation and preservation

## Why should I care?

“traditionally, preserving things meant keeping them unchanged; however ... if we hold on to digital information without modifications, accessing the information will become increasingly more difficult, if not impossible.”

Su-Shing Chen, “The Paradox of Preservation”, *Computer*, March, 2001, pp. 2-6.

- Emulation needs an understanding of the environment of creator and user
- Migration needs an understanding of the qualities of digital objects





a centre of expertise in data curation and preservation

## What is significant?

“A digital object’s Significant Properties are not empirical; archives will make judgments at levels appropriate to fulfil their preservation responsibilities and meet the needs of the archive’s user communities”

The Cedars Project Report, 2001

“Definitions of Significant Properties that affect the aesthetics, implied meaning, and affordances of digital objects tend to be much more subjective and tied to the context of creation and use.”

Hedstrom & Lee, 2002

Fundamental Questions of digital preservation:

1. What must you retain to ensure the integrity and authenticity of the digital object?
2. What can you lose without potential implications?



| D | C | C

a centre of expertise in data curation and preservation

## As a creator

- Clearly identify core users (Designated Communities) and involve them in production of localised significant properties “frameworks”
- Ensure decisions are appropriately documented in project/collection documentation and encoded within metadata
- Take into account the support for significant properties when choosing file formats, metadata schemas, and repositories



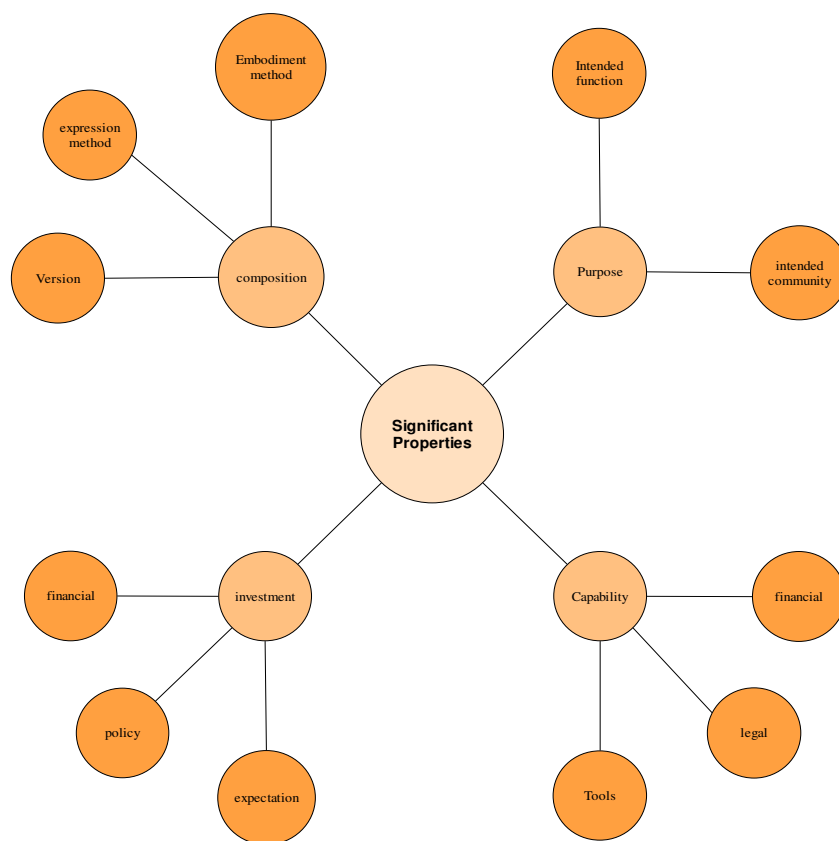
a centre of expertise in data curation and preservation

## As a curator

- Seek to capture and retain the creator's significant properties information on ingest
- Map these against their own significant properties framework (this will help in decisions about whether to normalise data and what preservation interventions are appropriate)
- Encode/document significance
- Test properties after preservation interventions



## Criteria for deciding significance



- **Composition of the digital object**
  - Form in which the idea is expressed
  - Expression method in a digital environment
- **Purpose**
  - Intended function (e.g. diplomatic analysis)
  - Type of user
- **Organisational investment**
  - Strategic
  - Financial
  - Expectation
- **Capability**
  - Tools
  - Legal
  - Financial