A Foundation for Trust in Digital Preservation

Kenneth Thibodeau,
Research Leader
Preservation as a Service for Trust (PaaST) Project

International Symposium – InterPARES Trust at 20!
Casa de las Américas, La Habana, Cuba
18 de febrero de 2019
Preservation as a Service for Trust (PaaST)

- 1,350 requirements for preservation of any type of digital information, including records.
- Technology neutral
  - Maximizes potential for automation
  - Enables managing change in technology
- Adaptable to different preservation objectives and policies
  - Supporting various divisions of responsibility
- **Verifiable** authentic preservation
The last word: TRUST

a reliable expectation that things are preserved authentic
A Service is a set of related actions that accomplish a preservation function, task, or objective.

Requirements organized in 3 groups of 13 services:

- **Preservation:** Submission, Storage, Access, Preservation Change
- **Preservation Management:** Preservation Rules, Assessment, Verification, Problem Handling
- **Information Management:** Class Management, Set Management, Data Management, Document Management, Reporting,
Digital Preservation Model

Preservation Target

Preservation Process

Reproduced Preservation Target

software

time

hardware
In House Preservation

Preservation Target

Hardware

Software

Time

Preservation Process

Reproduced Preservation Target
Basic Concepts: Things

- **Intellectual Entity** ≡ an object that is capable of communicating information
  - E.g., text, photograph, motion picture, database, et al.

- **Preservation Target** ≡ an Intellectual Entity that is (to be) preserved
A Preservation Target is distinguished by the specification of **Permanent Features**

- Feature $\equiv$ attribute or operation
- Permanent Feature $\equiv$ one that must remain invariant.

**Permanent Features** form the basis for verifying authenticity
Concepts: Digital Preservation

- Digital preservation $\equiv$ a process that enables the reproduction of a PreservationTarget
  - A Machine Readable PreservationTarget is reproduced in a RuntimeVersion; i.e., by being loaded into a computer
  - A Human Readable PreservationTarget is reproduced in a form accessible by a human
Concepts: Authentic Preservation

- Operational Definition: a reproduced PreservationTarget is authentic if all of its PermanentFeatures are identical to those declared when the PreservationTarget was designated for preservation.
Building a Foundation for Trust

Data about Preservation Target

Preservation Target

Preservation Process

Reproduced Preservation Target
The PaaST Foundation for Trust

Data about Preservation Target

| Permanent Features | Preservation Services |

Preservation Process

Verification of Authenticity

Preservation Target

Reproduced Preservation Target
Permanent Features

- Specified by the person authorized to decide what is preserved, how it is preserved, and whether preservation is successful.

- Practical constraint: it must be possible to specify how each Permanent Feature is expressed.
Specifying Permanent Features

- Permanent Attribute
  - Existence
  - Value
  - Expression

- Permanent Operation
  - Function
  - Return value or postcondition
PermanentAttribute example: Email

- PermanentAttributes: (e.g.) sender, addressee, copy addressee
- Existence
  - sender & addressee: required
  - copies: conditional
- Value
  - Sender: specified for sent, unspecified for received
  - Addressee: unspecified for sent, specified for received
  - cc: ad hoc.
- Expression: location or label of a data element in a message; MIME type, font, emphasis (e.g., italics, bold)
# PaaST Functions

## Preservation Services

<table>
<thead>
<tr>
<th>Service</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submission Processing</td>
<td>Bring PreservationTargets into a Preservation Environment</td>
</tr>
<tr>
<td>Preservation Change</td>
<td>Change the way a Preservation Target is encoded and/or reproduced</td>
</tr>
<tr>
<td>Preservation Storage</td>
<td>Store, retrieve and ensure integrity of DigitalComponents</td>
</tr>
<tr>
<td>Access</td>
<td>Instantiate Preservation Targets in RuntimeVersions, human readable output</td>
</tr>
</tbody>
</table>

## Preservation Management

<table>
<thead>
<tr>
<th>Management</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preservation Rules</td>
<td></td>
</tr>
<tr>
<td>Assessment of data quality</td>
<td></td>
</tr>
<tr>
<td>Verification of preservation</td>
<td></td>
</tr>
<tr>
<td>Problem Handling</td>
<td></td>
</tr>
</tbody>
</table>

## Information Management

<table>
<thead>
<tr>
<th>Management</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Management</td>
<td></td>
</tr>
<tr>
<td>Set Management</td>
<td></td>
</tr>
<tr>
<td>Data Management</td>
<td></td>
</tr>
<tr>
<td>Document Management</td>
<td></td>
</tr>
<tr>
<td>Reporting</td>
<td></td>
</tr>
</tbody>
</table>
Preservation Services

Start

Process Submission

Store Preservation Target

Change Preservation Technique

Access Preservation Target
Digital preservation requires expert curation

- Deciding what is to be preserved
- Assigning responsibilities
- Defining Preservation Rules
- Specifying Permanent Features
- Deciding when to implement Preservation Change
- Determining criteria for success
- Evaluating performance