Preservation in the Cloud
Preservation in the Cloud
Preservation in the Cloud
Preservation in the Cloud = Authentic Copy
Traditional preservation practice is hands on.
To the extent possible
In the preservation of digital information, there is inevitably some element of the unknown.
In house, we may have some knowledge and possibly some control over the hardware and software used for preservation. Probably do not know everything that does or can happen to the bits.

In the cloud, we may not know what technologies are used, where they are, or when they are changed.

In any case, we cannot know what technologies will be used for preservation in the future.
Preservation as a Service for Trust (PaaST)
Preservation as a Service for Trust (PaaST)

Preservation Service Specification
2013-2018

a multi-national, interdisciplinary research project exploring issues concerning digital records and data entrusted to the Internet. Its goal is to generate theoretical and methodological frameworks to develop local, national and international policies, procedures, regulations, standards and legislation, in order to ensure public trust grounded on evidence of good governance, a strong digital economy, and a persistent digital memory.

http://interparestrust.org/trust
ITrust

Research teams from North America, Latin America, Europe, Africa, Asia, Australasia

Transnational Team: NATO, UNESCO, World Bank, International Monetary Fund, Red Cross, International Centre for the Preservation and Restoration of Cultural Property, Alliance for Permanent Access to the Records of Science, Barrick Gold Corporation

Researchers from universities, archives, libraries, think tanks, national, regional and local government, private companies, independent consultants

62 research projects in the areas of law, policy, control, access, infrastructure, security, resources, social/societal issues, education, and terminology

Experts from archives and records management, information science, information technology, computer security, digital forensics, law, social sciences, engineering, physics and art.
OMG is an international, open membership, not-for-profit technology standards consortium. Founded in 1989, OMG standards are driven by vendors, end-users, academic institutions and government agencies. OMG Task Forces develop enterprise integration standards for a wide range of technologies and an even wider range of industries.


In June 2014, ITrust proposed to the OMG the development of an OMG standard on digital preservation based on the functional requirements being defined in PaaST

The proposal was accepted and is going forward under the OMG Government Domain Task Force

The preservation standard will be aligned with the existing OMG Records Management Services Specification
<table>
<thead>
<tr>
<th>ITrust PaaST Project</th>
<th>Object Management Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articulating requirements for preservation of authentic</td>
<td>Translating PaaST requirements into a form that supports production of software to</td>
</tr>
<tr>
<td>records in the Cloud</td>
<td>implement them.</td>
</tr>
<tr>
<td>Applying the knowledge gained from archival science</td>
<td>Generalization to all types of digital information, but including requirements for</td>
</tr>
<tr>
<td>and records management</td>
<td>authentic records</td>
</tr>
<tr>
<td>Coordinated with other ITrust Projects</td>
<td>Formulation in a platform-independent model</td>
</tr>
</tbody>
</table>
PaaST Scope: What

- Digitally encoded information
  - Primarily records, but potentially all types.
  - Preservation only of information objects that are in ‘definitive form,’ i.e., whose documentary form and content are not subject to any intentional change except for
    - Annotations for the purpose of preservation
    - Changes in the way they are materialized in order to enable their continued existence and usability in the face of technological changes.
PaaST Scope: Capabilities

- Does not include the management of current records.
  - Will be compatible and interoperable with the OMG Records Management Services Specification

OMG Records Management Services:
  - Capture
  - Categorization
  - Provenance
  - Authenticity
  - Disposition
  - Reference

OMG RMS v. DoD 5015.2 RMA:
  - RMS manages records in business systems
  - RMA removes records from business systems and stores them in virtual filing stations
  - RMS implementations could comply with 5015.2
Preservation requirements are being articulated as ‘services;’ i.e., sets of related capabilities such that each service could be executed using different and potentially unrelated technologies and under separate and independent operational control.

Hence, the overall context in which PaaST preservation activities are carried out is called the “Preservation Environment” rather than a preservation system.
Preservation in the Cloud
Preservation in the Cloud
Contextual Decisions

- What will be preserved, under what terms and conditions
- Who will perform what service(s)
- Decisions are articulated in business rules.
  - Business rules are expressed in appropriate documents
    - In house: records schedules identify permanent records
    - Service provider: contract specifies what permanent records will be sent to the provider for what purpose.
- and implemented in PaaS applications
Know what is preserved
Specify the properties that must not change
Document how it is stored as binary data objects
Verify that the data objects do not suffer corruption or loss
Assess whether the data objects preserve the properties that must not change
Change the data objects or the technology used to process them
Being able to output **authentic copies**.
PaaST
Preservation
Ontology

PreservationObject
  Content
  DocumentaryForm
  Creation : date
  LastModification : date

ArchivalObject
  RecordsCreator

Document
  0..*
  has
  governs
  Instantiates

PermanentDocumentProperties

FixedDocumentaryForm
  StableContent
  DeterminedPresentationFeatures
  ConsistentContentDifferentiation

Record
  Document ID
  RecordAssociation
  Disposition
  1..*
  has
  governs

RecordAggregate
  1
  has
  Instantiates
  0..*

PermanentAggregateProperties

Members
  MemberArrangement

DocumentVersion

PermanentRecordProperties

RecordVersion

RecordAggregateVersion
Services designed to facilitate orchestration as needed
Orchestration Scenarios

Comprehensive preservation system

Some services performed in house, others under contract

Services acquired as needed
Implementation of Trust: Preservation Services

- **Submission**: transfer of records into a Preservation Environment
- **Characterization**: inspection of the technical, archival and representational properties of the records
- **Authenticity**: capture and reporting of information about authenticity and the application of authentication methods.
- **Preservation Storage**: control of storage to maintain identity, prevent corruption and satisfy other preservation requirements.
- **Preservation Change**: technological changes to ensure survival and usability
- **Access**: focused on delivery of copies of PreservationObjects and support for assessing their authenticity

Does not address search, discovery, or use after delivery.
Thank you

For more information:

3rd ITrust International Symposium
World Bank, May 15:
https://tinyurl.com/k75y6dq

InterPARES Trust:
https://interparestrust.org/

OR

kthibodeau@fordham.edu