Academic Preservation Trust
Core Preservation Service Policy, Version 2.0

Scope

This policy sets out the current preservation-service components offered by the Academic Preservation Trust [APTrust] in its core, high-assurance preservation service. The policy applies to all content that members deposit into the APTrust environment via the tools and ingest policies set down in our basic operating principles. This policy is guided by the APTrust Mission Statement.

Objectives

This policy outlines the core preservation activities that each depositor can expect to receive from APTrust. It is broken down into three parts. The Administrative section indicates the frequency and responsible party for updating the policy; the Content section outlines the parameters of responsibility; and the Service section details the services and commitments provided to APTrust users.

Policy statement

1. **Administrative**
   - 1.1. This policy is to be reviewed and updated annually by a group designated by either the APTrust Board and/or the Advisory Committee.
2. **Content**
   - 2.1. Content is preserved in its original format (guided by the these goals)
     - 2.1.1. Materials deposited will be rebagged but no other changes to the individual files will be undertaken by APTrust staff/services
     - 2.1.2. Content will be managed solely by the depositor
     - 2.1.3. Deposited content MUST NOT contain sensitive or personally identifying information unless encrypted.
     - 2.1.4. Depositors MUST deposit content at their own risk.
3. **Service Components**

   These are the particular elements (benchmarks) for preservation in APTrust's core, high-assurance computing service.
   - 3.1. Bit level object preservation
     - 3.1.1. APTrust staff are notified when a failure occurs and a copy from the secondary storage repository in Glacier in Oregon is used to replace the corrupt copy, unless a depositor is using Glacier-only storage. Fixity is not checked on Glacier-only items, and there is no second copy to restore.
       - 3.1.1.1. Depositors are notified of a failure via email when a failed fixity check occurs.
       - 3.1.1.2. Fixity checking is not performed on items stored in Glacier-Only storage or Glacier Deep Archive Storage so no failure notifications are possible for those items, and no replacement is carried out.
     - 3.1.2. Fixity checks will be performed on every file
       - 3.1.2.1. This is true of files in S3, not of those only in Glacier
       - 3.1.2.2. Fixity algorithms will be cryptographically secure

3.1.2.2.1. So that no one can translate the fixity value back into the original file
3.1.2.2.2. To make it unlikely or impossible for someone to replace a valid file with an invalid one that produces the same fixity value.

3.1.3. Frequency of Fixity
3.1.3.1. Current frequency is every 90 days per file.
3.1.3.2. Based on cost for frequency (i.e. if less than or greater than 90 days)

3.2. Geographical diversity of content storage (East and West coast)
3.2.1. Standard Storage Schema
3.2.1.1. APTrust stores one (1) copy in Amazon Web Services (AWS) in Northern Virginia and one (1) copy in Glacier in Oregon. Internally, S3 and Glacier each have 3 copies of that file, totalling 6 overall.
3.2.2. Glacier-Only Storage Schema
3.2.2.1. Files are stored only in Glacier, in a single region (Ohio, Oregon, or Virginia), as specified by the depositor. Glacier maintains 3 internal copies in that region. APTrust does not perform fixity checks on Glacier-only storage. All Glacier files are encrypted at rest and decrypted upon retrieval. AWS manages the encryption keys as well as the encryption/decryption of data.
3.2.3. Glacier Deep Archive Storage Schema
3.2.3.1. Similar to Glacier-only but restoring bags from Glacier Deep Archive can take up to twelve hours, four times longer than standard Glacier. This increases the time it takes for APTrust to reconstruct the bag and move it to S3.

3.3. Multiple storage technologies
3.3.1. Two different storage technologies
3.3.1.1. S3 (spinning disk)
3.3.1.2. Glacier [technology not announced by AWS, but not spinning disk]

3.4. Multiple Checksums
3.4.1. Using both SHA256 and MD5
3.4.1.1. Partner may submit either MD5 or SHA256 or both. APTrust calculates both on ingest.
3.4.2. AWS uses eTag for receipt of transfer
3.4.3. Upon restoration, APTrust verifies both MD5 and SHA256 checksums.

3.5. Appropriate access (authentication and authorization is documented as part of our TDR work particularly section 4.6.1)
3.5.1. This includes Two Factor Authentication for Administrators.
3.5.2. Two-factor deletion prevents destruction of data by a single bad actor.

3.6. Search and discovery of metadata

3.7. Logging of preservation actions
3.7.1. Register of PREMIS events based on the Library of Congress standard
3.7.1.1. Auditing of PREMIS events (we capture and put locally)
3.7.1.2. Lower level logs of activities for system integrity
3.7.1.3. S3 records object-level API activity logs that we capture in S3
3.7.1.4. Able to copy those logs to local storage logs
3.7.1.5. Restorations of objects/files captured as WorkItem records

3.8. Data syncing across storage layers
3.8.1. In standard storage option, files are kept in sync in S3 and Glacier
3.8.2. In all storage options, preserved files include metadata tags identifying the depositing institution, intellectual object name, file identifier, and known good MD5 and SHA256

checksums.

3.9. Standardized metadata set (the current elements agreed upon by partners)

3.10. APTrust maintains a basic relationship among objects contributed by a partner

3.10.1. All items in preservation storage are tagged with essential metadata (depositing institution, intellectual object name, file identifier, and known good MD5 and SHA256 checksums)

3.10.2. The more robust version of this is managed in Pharos.

3.11. Content Retrieval

3.11.1. Batch downloading (i.e. by the bag)

3.11.2. Individual file restoration

3.12. Ongoing Development

3.12.1. Updating the APTrust environment is ongoing. Based on user input and requests new services are added as needed. There is no formal schedule for revising and updating--it is continuous.

Related policies and subordinate documents (see most recent)

APTrust Mission Statement

Collection Development Policy

Preservation and Storage within APtrust

NOTE: Other items may be added to this section without requiring policy revision

Further information

NDSA Fixity Best Practices

NOTE: Other items may be added to this section without requiring policy revision

Definitions

Term	Definition
Fixity	Verifying that an object has not been inadvertently changed, adapted, or corrupted
PII	Personally Identifiable Information (e.g. Social Security Numbers)

Roles and stakeholders

- APTrust staff is responsible for updating the technical Services section of this document and vetting for accuracy. Each significant update will require a new version of this policy.
- APTrust Advisory Group is responsible for approving any new versions of this document.

Role/stakeholder	Responsibilities
Advisory Committee	Review entire document as needed but no less than annually
APTrust Technical Staff	Review and update Services (Section 3) as needed - at a minimum annually

Previous Versions

Version 1.0 (2018)

Review

Review frequency: Annual or as needed
Next review date: Summer 2021

Version History

Version	Status	Date	Notes
.1	Done	1/12/17	Comments from APTrust Community
.2	Done	2/9/17	Communications Group approves vote for APTrust
.7	Done	3/16/17	APTrust Advisory Committee approval
.8	Done	3/17/17	Minor edits completed based on Advisory Feedback
1.0	Done	4/2018	Board Vote and Approval - DOI uploaded
1.1	Done	11/2019	Open for comments from APTrust Community
1.8	Done	10/2020	Advisory Committee Vote
2.0	Done	10/2021	Board Vote and Approval - DOI uploaded