Overview

Chronopolis is a digital preservation network of data storage centers managed by the San Diego Supercomputer Center and University of California, San Diego in partnership with the National Center for Atmospheric Research and the University of Maryland Institute for Advanced Computer Studies. Originally funded by the Library of Congress's NDIIPP initiative, the project began offering services to the public in 2011. Chronopolis utilizes a grid of storage centers at each partner site to provide long-term preservation-oriented storage. **Chronopolis is among the few storage service providers cited by the Center for Research Libraries as compliant with the Trustworthy Repositories Audit and Certification (TRAC) criteria.** Chronopolis is also one of the initial nodes of the Digital Preservation Network and will soon begin ingesting data into this network of preservation centers. Storage management via the DuraCloud service was made available in early 2014.

Chronopolis

SERVICE PROVIDER: San Diego Supercomputer Center/ UC San Diego

WEBSITE: chronopolis.sdsc.edu

PRODUCT RELEASED: 2011

COMPLIANCE: TRAC audited

SERVICE: Deep storage

INFRASTRUCTURE: Partnership

COST: High

Data Management

Chronopolis supports the BagIt specification for packaging submissions. The contents of the bags—the submission formats and associated metadata—are determined in collaboration with Chronopolis. Transfer of files has taken various forms in the past but is typically completed via network transfer protocols (ssh and wget). Ingest procedures validate the contents of the bag and its associated checksums. **The service performs regular audit of SHA-256 checksums every 30 days or at customer specified intervals.**

Reporting / Metadata

Chronopolis has implemented its own model for preservation metadata and **regularly captures metadata at eight points throughout a file's lifecycle.** Users can review storage contents and process metadata via Chronopolis's data portal, including overall collection status, item level browsing, log retrieval, error report retrieval, activity report viewing, download collection digests, duplicate detection, and token downloading.

Redundancy

Data in the Chronopolis network is **replicated at three geographically dispersed partner data centers** at the San Diego Supercomputer Center, National Center for Atmospheric Research (Colorado), and the University of Maryland, using the iRODS system, an open source data system for managing across the tape and disk- based storage at each site. Chronopolis also has defined internal protocols for disaster scenarios according to their TRAC audit.

Accessibility

Chronopolis is intended for long-term bit-level preservation of data; **the system does not allow active client-side management and does not support automated protocols for retrieving data out of storage.** Users must contact system administrators to arrange delivery of data from the system. Tools for improved management functionality for client-side use are currently in development.

Security

Chronopolis system administrators actively monitor the security of storage nodes within the larger data centers as specified in the service's documented security protocol. **Chronopolis is defined as a "dark" archive, limiting the amount of direct access to data available to clients and staff.** Staff roles and responsibilities also limit the number of people who are permitted to make changes to the system. Chronopolis does not state any support of encryption or other object-level security measures.

End of Service

Chronopolis' most recent documentation states that a formal contingency plan is in development and tests have been completed to export data to non-Chronopolis systems in case the service ends. For cases in which the client fails to pay required fees results in an account becoming read only after 30 days, locked after 60 days, and deleted after 90 days. Customers may opt out of the service in writing 60 days prior to the annual renewal of the agreement and must retrieve data by renewal date. Chronopolis does not specify a standard method for return of data from their service back to the client.

Levels of Preservation

	Level 1 (Protect)	Level 2 (Know)	Level 3 (Monitor)	Level 4 (Repair)
Storage	• **Two complete copies that are not collocated**	• **At least three complete copies** • **At least one copy in a different geographic location**	• **At least one copy in a geographic location with a different disaster threat** • **Obsolescence monitoring for storage system(s)**	• **At least three copies in geographic locations with different disaster threats**
Data Integrity	• **Check file fixity on ingest if it has been provided with content** • **Create fixity info if it wasn't provided with the content**	• **Check fixity on all ingests** • Virus-check high risk content	• **Check fixity of content at fixed intervals** • **Maintain logs of fixity info; supply audit on demand** • **Ability to detect corrupt data** • Virus-check all content	• **Check fixity of all content in response to specific events or activities** • **Ability to replace/ repair corrupted data** • Ensure no one person has write access to all copies
Security	• **Identify who has read, write, move, and delete authorization to individual files** • **Restrict who has those authorization to individual files**	• **Document access restrictions for contents**	• **Maintain logs of who performed what actions on files, including deletions and preservation actions**	• Perform audit of logs

Have more questions? We can help.
Contact AVPreserve at info@avpreserve.com