



Center *for* Research Libraries

.....
GLOBAL RESOURCES NETWORK

Certification Report on Chronopolis

Executive Summary

The Center for Research Libraries (CRL) conducted a preservation audit of Chronopolis (chronopolis.sdsc.edu) between November 2010 and December 2011, and on the basis of that audit certifies Chronopolis as a trustworthy digital repository. *The CRL Certification Advisory Panel has concluded that the practices and services described in Chronopolis' public communications and published documentation are generally sound, and appropriate to the content being archived and to the general needs of the designated Chronopolis community.* Moreover the Panel expects that in the future, Chronopolis will continue to be able to deliver content that is manageable and usable by that community.

CRL certification applies to the repository's ability to preserve and manage digital files and data deposited by the Inter-university Consortium for Political and Social Research; and diverse sets of data files from the California Digital Library, North Carolina Geospatial Data Archiving Project, and the Scripps Institution of Oceanography.

This certification is based upon review by CRL and the members of its Certification Advisory Panel of extensive documentation gathered by CRL independently from open and third-party sources as well as data and documentation provided by Chronopolis. The review also included a site visit by CRL audit personnel to the San Diego Supercomputer Center on October 26, 2010.

The overall conclusion of the panel was that Chronopolis can be recognized by the designated community as a trustworthy repository with certain important considerations. One consideration is the level of preservation service Chronopolis provides. The repository's mission statement identifies Chronopolis' goal as providing "a preservation data grid and its supporting human, policy, and technological infrastructure." However, Chronopolis does not commit to services beyond preserving intact the bits deposited in the repository. This limitation is clearly expressed in the Chronopolis subscriber license agreement, which disclaims responsibility for performing specific "preservation actions" that some other repositories provide, such as format migration, file normalization, file type verification, and creation of descriptive metadata. Because this limitation is clearly communicated to Chronopolis stakeholders, Chronopolis can be said to provide preservation services adequate to its community.

The second consideration is Chronopolis' relatively nascent and untested administrative infrastructure and business plan. At the time of this review Chronopolis was transitioning from being a largely grant-supported project to a university-based service that is expected to provide digital storage for research data in a variety of fields on a fee-for-service basis. Chronopolis had also established strong partnerships with the organizations hosting its three modules, and has secured commitments from those organizations to cover their own respective costs for the next several years.

Enlarging the Chronopolis client base, however, will be critical to sustaining the service and its infrastructure in the coming years. The repository's operating costs will have to be shifted from federal, university and partner subsidies to the Chronopolis "designated community," i.e., the data producers, owners and users. Therefore, periodic review and perhaps recalibration of Chronopolis policies, procedures, costs, and fee structure in the future will be essential to ensuring the success of the next phase in the repository's growth.

About the Audit Participants

CHRONOPOLIS

At the time of the audit Chronopolis (chronopolis.sdsc.edu) comprised four partner organizations that provide digital preservation services: the San Diego Supercomputer Center, the University of California San Diego Libraries, the National Center for Atmospheric Research, and the University of Maryland Institute for Advanced Computer Studies. Originally funded by the Library of Congress National Digital Information Infrastructure Preservation Program (NDIIPP), Chronopolis now depends upon the expertise and resources of the four partners to provide a geographically distributed, heterogeneous and highly redundant archive system.

At the time of the audit Chronopolis managed three geographically distributed copies of four collections from within the NDIIPP community: the extensive archive of public opinion data of the Inter-university Consortium for Political and Social Research; and diverse sets of data files from the California Digital Library, North Carolina Geospatial Data Archiving Project, and the Scripps Institution of Oceanography.

CENTER FOR RESEARCH LIBRARIES

The Center for Research Libraries (CRL - www.crl.edu) is an international consortium of university, college, and independent research libraries. CRL supports advanced research and learning in the humanities, sciences, and social sciences by ensuring the survival and accessibility of source materials vital to those disciplines. In order to enable its community to accelerate the shift to electronic-only resources in a careful and responsible manner, CRL both preserves and maintains shared physical collections of materials and evaluates digital repositories of interest to its community.

CRL analysis of Chronopolis documentation and operations was undertaken by Marie Waltz and other CRL staff. Additional technical support for the site visit and the assessment of Chronopolis repository systems and architecture was provided by Ann Green of Digital Life Cycle Research & Consulting.

To guide its Chronopolis audit CRL also enlisted a panel of advisors representing the various sectors of the academic research libraries field. The Certification Advisory Panel included leaders in collection development, preservation, library administration, and digital information technology, and is so constituted as to ensure that the certification process addresses the interests of the entire CRL community.

THE MEMBERS OF THE CRL CERTIFICATION ADVISORY PANEL ARE:

MARTHA BROGAN (CHAIR)
 Director of Collection Development & Management
 University of Pennsylvania

ANNE POTTIER
 Associate University Librarian
 McMaster University

WINSTON ATKINS
 Preservation Officer
 Duke University

OYA Y. RIEGER
 Associate University Librarian for Information
 Technologies
 Cornell University

WILLIAM PAROD
 Senior Repository Developer
 Northwestern University Libraries

PERRY WILLETT
 Digital Preservation Services Manager
 California Digital Library

MARK PHILLIPS
 Assistant Dean for Digital Libraries at
 University of North Texas Libraries

Technical analysis and support for the Chronopolis audit was provided by Ann Green, strategic analyst at Digital Life Cycle Research & Consulting, New Haven, Connecticut.

A. Audit and Assessment Methodology and Criteria

This assessment was undertaken to determine whether or not Chronopolis meets the commitments it has made in regard to the long-term preservation of digital data for the research community and whether the repository complies with established criteria for trusted digital repositories. The assessment included a site visit, a review of the information independently gathered by CRL from published and unpublished sources, and a review of documents and documentation provided by Chronopolis.

CRL conducted its audit with reference to:

- generally accepted best practices in the management of digital systems
- the interests of the CRL community of research libraries
- the practices and needs of scholarly researchers in the humanities, sciences, and social sciences in the United States and Canada
- the criteria included in Trustworthy Repositories Audit & Certification: Criteria and Checklist¹
- the Open Archive Information System reference model² (OAIS)
- other metrics developed by CRL through its analyses of digital repositories.

The primary metrics used by CRL in such assessments are those specified in the Trustworthy Repositories Audit and Certification (TRAC) checklist. TRAC was developed by a joint task force created by the Research Libraries Group (RLG) and the National Archives and Records Administration in 2003 to provide criteria for use in identifying digital repositories capable of reliably storing, migrating, and providing long-term access to digital collections. TRAC represents best current practice and thinking about the organizational and technical infrastructure required for a digital repository to be considered trustworthy and thus worthy of investment by the research and research library communities.

CRL assessed Chronopolis on each of the three categories of criteria specified in TRAC and has assigned a level of certification for each. The numeric rating (below) is based on a scale of 1 through 5, with 5 being the highest level, and 1 being the minimum certifiable level. (The minimal certification rating of 1 is assigned in instances where a repository has inconsistencies or deficiencies in areas that might lead to minor defects of a systemic or pervasive nature, but where no major flaws are evident.)

TRAC CATEGORY	CHRONOPOLIS RATING	OPTIMUM RATING
Organizational Infrastructure	3	5
Digital Object Management	4	5
Technologies, Technical Infrastructure, Security	4	5
TOTAL	11	15

The basis for assignment of these ratings is provided in Section B, Detailed Audit Findings, below.

It should be noted that CRL certification of Chronopolis applies specifically to the repository's ability to preserve and manage at the bit level digital social science, atmospheric and geospatial data from a variety of sources. *CRL did not assess Chronopolis procedures and processes for normalizing, migrating, or otherwise altering and preserving data for distribution via future platforms or devices.*

¹ TRAC - http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

² OAIS - <http://public.ccsds.org/publications/archive/650x0b1.pdf>

B. Detailed Audit Findings

On the basis of the audit, CRL identified two aspects of Chronopolis operations that should be clearly understood by current and prospective stakeholders. These aspects correspond to specific TRAC criteria or to features of the repository that members of the Certification Advisory Panel believe are important to the North American research community. They are as follows:

1. FUNDING PLAN AND FINANCIAL BENCHMARKS (TRAC CRITERIA A4.1, A4.2, A4.5)

As a reliable service, Chronopolis must be able to sustain its activities from year to year, at least for its current stakeholders. TRAC metrics A4.1 and A4.2 require that a repository have “short- and long-term business planning processes in place to sustain the repository over time” as well as “processes to review and adjust business plans at least annually.” At the time of the audit these processes were in place for Chronopolis but were as yet untested.

To date Chronopolis has subsisted on grants and contracts from the Library of Congress National Digital Information Infrastructure and Preservation Program (NDIIPP) and the University of California San Diego, with substantial support from its partner institutions: the University of Maryland Institute for Advanced Computer Studies (UMIACS) and the National Center for Atmospheric Research. Over the next few years the service will begin to generate operating revenue from providing data preservation services to other university and scientific efforts.

In a three-year business plan provided to CRL by Chronopolis in April 2011, a number of then-current and prospective funding sources were identified:

1. Support from Chronopolis node partners. UCSD, NCAR and UMIACS had all committed to providing base level funding for Chronopolis during the next several years. This money was “not tied to specific grants or a customer base, but was to provide the wherewithal to support the basic technical infrastructure, including hardware, software and staffing, to assure that the current systems continue to function and that the data in the systems are fully preserved.”
2. A multi-year subvention from the University of California San Diego university administration for local use of Chronopolis. UCSD instituted a campus-wide Research Cyber-Infrastructure program, for which Chronopolis will form the preservation backbone. UCSD has committed to providing funding to Chronopolis for two years, which will cover a substantial portion of Chronopolis operating costs.
3. Fee-based digital preservation services to other clients. Chronopolis management anticipated an increase in demand from UCSD and other universities for data storage and management due to the new NSF and NIH requirement for data management plans from grant applicants.

In addition to the specific funding sources listed above, Chronopolis continues to have a strong working relationship with the national preservation community, including the Library of Congress Office of Strategic Initiatives’ National Digital Information Infrastructure and Preservation Program (NDIIPP), the National Digital Stewardship Alliance (NDSA), and the Digital Preservation Initiative (DPN).

2. PRESERVATION POLICY (TRAC CRITERIA A3.1, A3.6, B3.1, B4.1)

Implicit in several TRAC criteria is the existence of a written preservation policy, a document that expresses the repository’s commitment to preservation and outlines its chosen preservation strategies. The Chronopolis Mission Statement characterizes Chronopolis as “a preservation data grid and its supporting human, policy, and technological infrastructure.” However, unlike many other preservation repositories Chronopolis does not commit to preservation beyond the bits deposited by the client.

However, the Chronopolis Service Level Agreement (SLA) emphatically circumscribes the preservation services Chronopolis provides, as follows: “Please note that Chronopolis does NOT perform specific ‘preservation actions’ upon files during or after ingest. This includes actions such as file format migration, file normalization, file type verification, creation of descriptive metadata, and rights management. If a customer wishes to have these services, they need to be done by the customer before data is deposited into Chronopolis.”

Interviews with members of the then-current Chronopolis designated community indicated that they were aware of and satisfied with the data storage and backup services Chronopolis provided. At the time of the audit, then, Chronopolis was fulfilling its commitments to its key clients by providing fairly minimal preservation services.

C. Ongoing Requirements

The TRAC document notes that “. . . attaining trusted status is not a one-time accomplishment—achieved and forgotten. To retain trusted status, a repository will need to undertake a regular cycle of audit and/or certification.” To that end CRL expects that Chronopolis will also make certain disclosures on a regular basis. CRL and Chronopolis have agreed that ongoing certification is contingent upon Chronopolis making the following disclosures every two years:

- A detailed listing of new content added to the repository since certification;
- Description of any significant changes in repository system architecture or configuration, operating systems and/or critical software;
- New agreements and contracts with key depositors of content, content users, major funders or sources of revenue, and providers of critical repository services;
- New key policies regarding acquisition, management, and disposition of archived content and related files and metadata;
- Records of significant events (such as content migrations, system failures, loss or corruption of digital content) and significant changes in the characteristics of digital content ingested since the most recent audit; and of significant events and changes in the operations of the repository.
- The most recent three years of financial statements for the repository organization or service unit. The financial statements should indicate the categories and, where appropriate, sources of revenue and the level of same; the functional allocation of expenses; and changes in the financial position of the organization supporting the service unit.
- Revenue and expense projections by function, for the repository organization or service unit, for the next three years.

Certification is also contingent upon Chronopolis agreement to a periodic, systematic sampling and/or inspection of the repository’s archived content by CRL, or by a third party designated by CRL and Chronopolis jointly, using either a manual or an automated process, as determined by mutual agreement between CRL and Chronopolis.