

---

# RISK ASSESSMENT

## NDIIPP MetaArchive Project



2009-02-16

### Summary

The following is an assessment of the key elements of the MetaArchive Cooperative (and its administrative entity the MetaArchive Services Group) that impact the Network's performance, sustainability and accountability. This document draws on the recommendations made by RLG/OCLC's *Trusted Digital Repositories: Attributes and Responsibilities* report (2002) and the training workshop "Digital Preservation Management" at Cornell University. In the following pages, the MetaArchive Cooperative documents its organizational infrastructure.

### About the Cooperative

The MetaArchive Cooperative is an independent, international membership association that provides collaborative and distributed digital preservation solutions for cultural memory organizations, including university libraries, government and historical research archives, museums, and other similar institutions (<http://www.metaarchive.org>). The central missions of the MetaArchive Cooperative are to support, promote, and extend the practice of distributed digital preservation; to serve as a catalyst and guide for other networks that seek to implement the distributed digital preservation methods it has developed; and to educate cultural memory organizations about distributed digital preservation.

### Organizational Infrastructure Attributes

**Governance & Organizational Viability:** By maintaining a collective governance structure and distributing knowledge and responsibility throughout its membership, the Cooperative ensures that its sustainability is not dependent on a centralized staff or on any one of its member organizations. The Cooperative has also created and documented policies and procedures to protect the Cooperative in the case of technical, legal, or management failures by any of its members.

**Organizational Structure and Staffing:** The Cooperative's intentionally decentralized staffing structure ensures that knowledge and skills are distributed and shared across its membership. The Cooperative's shared knowledge base, combined with its attention to quality staffing, ensures the competency of network staffing for the organization.

**Procedural Accountability and Policy Framework:** The Cooperative regularly incorporates feedback from its membership via its committee-based structure and its weekly communications. It is responsive to technological developments and is accountable to stakeholders from the extended community of cultural memory organizations.

**Financial Sustainability:** The Cooperative is supported by a funding model with a mixture of revenue streams that includes contract and grant funding, membership revenues, and consulting revenues. Capital risks are mitigated through a combination of legal agreements and the Cooperative's decentralized ownership model, in which the Cooperative does not own equipment or other assets. The

Cooperative holds a firm commitment to short and long-term planning, focused first on maintaining its core operations and second on exploring possibilities for technical and organizational growth.

**Contracts, Licenses and Liabilities:** Individual institutions that become members agree to bear responsibility for resolving copyright and intellectual property issues locally in the Membership Agreement.

### **Governance & Organizational Viability**

The MetaArchive Cooperative is an independent, international membership association hosted by the MetaArchive Services Group, based in the U.S. in Atlanta, Georgia. The Cooperative's purpose is to support, promote, and extend a community-based approach to distributed digital preservation practices for cultural memory organizations (libraries, museums, archives, centers, societies, etc.). The MetaArchive Cooperative is responsible for preserving member institutions' content in a decentralized, distributed preservation network comprised of subject- and genre-based archives (e.g., Southern Digital Culture, Electronic Theses and Dissertations, etc.), as well as maintaining and extending its methodology and approach to distributed digital preservation. The MetaArchive Cooperative is committed to the long-term retention and management of digital information as exhibited by its mission statement and operating principles. It is equally committed to helping cultural memory organizations to work together to preserve their own digital collections in house through community-based activities.

#### **The Mission of the MetaArchive Cooperative**

The mission of the MetaArchive Cooperative is to create enduring and stable, geographically dispersed archives of digital materials that can, if necessary, be drawn upon to restore collections at the contributing institutions, and to foster better understanding of distributed digital preservation methods. The Cooperative is guided in its work by the following operating principles:

- Encourage and support the long-term preservation of corpora of "at-risk" cultural memory content;
- Promote a cooperative, robust, and decentralized approach to digital preservation;
- Encourage cultural memory organizations to build their own preservation infrastructures and knowledge rather than outsourcing this core service to external vendors;
- Encourage the growth of Private LOCKSS networks (PLNs) as the best current technical option for distributing copies of replicated content in secure, distributed locations over time;
- Maintain a minimal overhead and straightforward mechanisms for collaboration;
- Host services that have wide applicability to a range of institutions and digital content;
- Utilize and create open standards and systems;
- Ensure that digital materials are stored and maintained in migratable formats and data structures;
- Promote and support high standards for preservation metadata capture;
- Undertake research and development projects to advance digital preservation.

#### **Governance/Administration**

The MetaArchive Cooperative is hosted by the MetaArchive Services Group, based in Atlanta, Georgia. The MetaArchive Services Group is incorporated under the laws of Georgia and is a 501(c)3 non-profit institution that serves as a mediating and coordinating entity for leaders from academic and cultural institutions such as university libraries, government and historical research archives, museums, and other similar institutions. The MetaArchive Services Group is governed

by a Board of Directors and administered by the President. The MetaArchive Cooperative is overseen by the Steering Committee, the Steering Committee Chair, and the Program Manager.

As a Cooperative, all members share in the ownership of the Cooperative and, as such, have a shared interest in developing and sustaining its technical and organizational infrastructures. Membership is comprised of three levels: Sustaining Members, Preservation Members, and Contributing Members. The Steering Committee is the governing body of the MetaArchive Cooperative, consisting of one representative from each Sustaining Member.

The Steering Committee is responsible for overall management, coordination, communication, and reporting efforts. Decisions made by the Steering Committee are made by simple majority vote, either documented via email or at the annual meeting. In addition, a Content Committee, Preservation Committee, and Technical Committee inform the decision-making of the Cooperative. The Cooperative hosts an annual business meeting in Atlanta, Georgia, during October of each year. The venue and agenda is distributed by e-mail to the member organizations at least one month in advance. This business meeting is used for discussion and to make decisions for the following academic year, including the election of Steering Committee Members to assume office, election of leadership for all Committees, annual review of membership fees, annual review of the Charter document, annual review of the Technical Specifications, and other needed actions.

### **The MetaArchive Cooperative Charter and Membership Agreement**

The MetaArchive Charter outlines the mission and operating principles; membership eligibility, benefits, responsibilities, and costs; organization and governance; and services and operations of the Cooperative. The Membership Agreement formalizes the legal relationship between the Cooperative and its member organizations.

The Cooperative Charter documents the practices and roles and responsibilities for both members and the organization. The Charter addresses issues that may arise in a cooperative governance structure, and outlines policies that protect the existence of the Cooperative in the case of technical, legal, or management failure by any of its members. These agreements promote transparency and accountability in all actions supporting the operation and management of the Cooperative.

The Membership Agreement articulates and informs potential members of the rights and responsibilities of joining the MetaArchive Cooperative and establishes the framework for ordering legal relationships among the members such as payment of fees, sharing of resources, duration of membership, and other facets of the relationship. The Agreement binds members, the MetaArchive Cooperative, and any successors in interest to certain conditions specified within it. It also gives some certainty to the Cooperative relationships while remaining sufficiently flexible to address evolving needs of the Cooperative and members in a changing technological and legal landscape.

In addition to the Charter and Membership Agreements, the Cooperative has formalized the following documentation: Management Plan, Operations Plan, Marketing Plan, and the Outreach Program Implementation Plan. These documents are openly available on its website and are updated at least annually to reflect changes in requirements and practices of the network. The updates are recommended by the Cooperative's committees and administration and approved by the Steering Committee.

### **Contingency Planning**

The MetaArchive Cooperative relies on its organizational members to provide the MetaArchive-LOCKSS caches that comprise the technical network. In order to ensure the quality and reliability of its preservation services, the Cooperative demands that every file stored in the network is

replicated across six distinct caches. This means that the network ideally must include at least six preservation caches. In order to ensure that the network stays stable, the Cooperative accepts members in three-year terms and requests that they provide a one-year notice of renewal or resignation at the end of their membership terms. This provides the Cooperative with a measure of insurance that it will be able to recruit a replacement member for any outgoing member prior to that outgoing member's departure from the network. In the unlikely case that the membership ever drops below this threshold, one or more organizations may install a second server as a temporary solution in order to satisfy the minimum number of caches required until a new member is recruited and in place.

If the MetaArchive Cooperative were to dissolve, its members could replicate the structure in a PLN of their own making or move their content easily to another PLN. This is a more affordable option within the Cooperative than it would be in most preservation scenarios since MetaArchive members a) own the equipment upon which their collections are preserved; b) have also attained the necessary in house technical and organizational knowledge needed to run a network; and c) have all of the network's open access software, plans, and documentation from which to work. In addition, plans are currently underway to create a bridge to the Chronopolis preservation network, which would allow members to move their collections to a Storage Resource Broker (SRB)-based solution.

### **Organizational Structure and Staffing**

The MetaArchive Cooperative's ongoing programmatic duties are not centralized; they have been intentionally structured such that they may be distributed across the MetaArchive Cooperative's member organizations' staff and faculty. Key to the Cooperative model is the belief that not only must MetaArchive's technical infrastructure be distributed across organizational members, but so must the knowledge and ability to run the infrastructure. The distributed nature of the services is intended to keep the organizational infrastructure of the MetaArchive Cooperative as lightweight as possible. Sustainability then depends on the distributed framework and the continued commitment of members, not on MetaArchive's employment of a central project staff.

There are three categories of membership for the Cooperative: Sustaining, Preservation, and Contributing. All members contribute content to the network for preservation purposes. To facilitate this work, the Cooperative encourages all members to assign the following roles to individuals within their institutions:

- Program Coordinator, who oversees the work of their institution within the MetaArchive Cooperative.
- Content Selector, who identifies and prioritizes content from their institution that should be preserved. The content selector is knowledgeable about the content of an institution's digital archives.

Sustaining and Preservation members additionally bear responsibility for a) running the Cooperative's technical network through establishing and monitoring a MetaArchive-LOCKSS cache for the overall network, b) processing their own collections for ingest into the network, c) processing Contributing members' collections for ingest into the network, and d) ingesting designated content into their cache for preservation.

To these ends, the Cooperative encourages each Sustaining and Preservation member to assign the following additional roles to individuals within their institutions:

- Data Wrangler, who is a programmer and/or is technically adept and can help prepare local digital archives for ingestion into a preservation network.

- Systems Administrator, who maintains the institution's cache server(s) for the preservation network.

Each Sustaining Member contributes technical support staff to the development and maintenance of the network and is responsible for the ongoing activity of preserving digital content. If at any point this donated staff time from our Sustaining Members does not satisfy the needs of the Cooperative, the Steering Committee has the option to allocate funds from the Cooperative's membership fees to pay for staff or consultants to assist with technical duties.

To facilitate this work, the Cooperative encourages each Sustaining Member to assign the following additional role to an individual(s) within their institutions:

- Programmer, who contributes to the technical development efforts undertaken by the Cooperative.

Each Sustaining Member of the Cooperative also names one person to the MetaArchive Steering Committee, the leadership group for the Cooperative's planning, development, and sustainability work.

The Cooperative's internal staff consists of three positions:

- A Program Manager (half-time), who provides leadership for the Cooperative and accepts responsibility for coordinating the activities of the digital preservation network.
- A Software Engineer (half-time), who installs software components and performs testing and programming on an ongoing basis.
- A Systems Engineer (half-time), who monitors the overall network, assists with the setup and maintenance of all preservation caches, conducts network testing, contributes to programming tasks, and researches and recommends equipment for purchase.

These three positions are funded through contracts and grants through 2010. By 2011, the Cooperative anticipates funding these internal, part-time staff positions through the MetaArchive Services Group to continue supporting the work of the Cooperative. Although these positions will be funded by the MetaArchive Cooperative, the staff will be distributed among our member organizations via contracts. Coupled with the Cooperative's Director, these positions will hold primary responsibility for the development and maintenance of the MetaArchive Cooperative.

Additional staffing for the administration of the network is conducted on an as-needed contract basis with consultants hired by the President and Executive Director of the MetaArchive Services Group. Services include legal, accounting, and business consulting.

### **Staff Training**

As a Cooperative, each Sustaining and Preservation Member institution is responsible for the general training of its staff members and by signing the Membership Agreement, commits to staffing the network with qualified and well-trained people. The Cooperative provides direct training opportunities at its annual meeting and also offers ongoing online and telephone support. In 2009, the Cooperative will publish a *Guide to Distributed Digital Preservation*. This book (which will be made available both as a freely downloadable PDF and as a printed book) will provide open documentation about running Private LOCKSS networks to serve as another training resource for our network.

### **Procedural Accountability and Policy Framework**

Membership in the MetaArchive Cooperative is open to cultural memory organizations engaged in digital preservation whose activities and objectives are consistent with the MetaArchive Cooperative's mission and principles. Cultural memory organizations are defined as:

- universities and colleges
- government entities
- research libraries, archives, and library consortia,
- museums and historical societies,
- collaborative digitization efforts,
- electronic text centers,
- science, social science, and humanities research computing centers,
- academic and other publishers; publishing associations;
- scholarly societies; and
- others that have significant or unique collections of at-risk digital content.

The Cooperative's governance structure, as previously described, ensures that the Cooperative incorporates feedback from its membership and is responsive to technological developments. The Steering Committee keeps abreast of changes through weekly conference calls and listservs that address both technical and administrative issues. Changes to policies and procedures are documented in the Charter, Operations Plan or other documents, and are reviewed and approved by the Steering Committee. The committee structure ensures transparency and accountability in all actions supporting the operation and management of the Cooperative. In addition, the MetaArchive Cooperative is accountable to stakeholders from the broader community of cultural memory organizations through its web site [www.metaarchive.org](http://www.metaarchive.org), which provides open access to the Cooperative's documentation of policies and procedures.

### **Information Integrity Protections**

The MetaArchive Cooperative's network is comprised of cache servers acting in concert to preserve content owned and selected by institutions participating in the network. These MetaArchive-LOCKSS caches are administered by the MetaArchive's member organizations that contract with each other to preserve this content in a collaborative, distributed manner. The technical framework provides a systematic way for these MetaArchive-LOCKSS caches to constantly check in with one another for signs of file degradation or "bit rot," as well as for a cache's corruption or disappearance due to technical or physical catastrophe. The MetaArchive-LOCKSS caches conduct "polls" to determine if any copy has been compromised. If they discover an inconsistency, they seek a quorum to ensure that they know which copy is no longer accurate. The software can then repair that damaged copy. Similarly, if an entire MetaArchive-LOCKSS cache experiences technical failure and withdraws/disappears from the network for any reason, the cache can be safely recreated via the other caches in the network.

Each of the network's MetaArchive-LOCKSS caches is under separate administrative control, guarding both against natural/technical disaster and the failure or withdrawal of any one cache. Any cache in the system can drop out — whether intentionally or due to technical or organizational failures — without jeopardizing either the network or its contents. Myriad testing scenarios have consistently demonstrated the stability of this preservation system's technical infrastructure.

## **Financial Sustainability**

### **Business Planning**

The Cooperative has engaged in extensive business planning in its contractual and grant work with the National Digital Information Infrastructure and Preservation Program (NDIIPP) and the National Historical Publication and Records Commission (NHPRC). Cooperative documentation includes a Charter (described previously) and a Preservation Business Model that is currently comprised of a Market Analysis, a Marketing Plan, a Management Plan, and an Operations Plan. The MetaArchive Cooperative and the Steering Committee have a clear understanding of short-term goals and

objectives and a framework upon which to build for long-term planning. Plans for the Cooperative are approved each year at the Annual Meeting.

A long-term planning process is being conducted in 2009 and will build upon the work of the Cooperative's committees. The Committees will review and discuss issues relevant to the long-term viability and future success of the network and begin to systematically plan for the next five to ten years of work. This planning effort feeds into an ongoing strategic planning process for the MetaArchive Services Group. A strategic plan for the MetaArchive Services Group will be finalized and approved by the Board of Directors in 2009.

The Board of Directors of the MetaArchive Services Group has established a Strategic Development Sub-Committee to explore grant-funding opportunities and advise the board on new directions that would enhance the Cooperative's core functions. This Sub-Committee is responsible for drafting a fund development plan based on the strategic objectives outlined in the strategic plan to address both short- and long-term funding strategies, identify potential funders, and establish a resource development timeline.

### **Financial Processes and Procedures**

The MetaArchive Cooperative budget is approved by the Steering Committee, with oversight from the MetaArchive Services Group Board of Directors. As a 501(c) 3 nonprofit organization, the MetaArchive Services Group is required to file an annual tax return to the Internal Revenue Service. MetaArchive Services Group has contracted with an accountant to maintain financial records and to produce reports for review by the MetaArchive Cooperative Steering Committee and the MetaArchive Services Group Board of Directors. To date, the budget amount is still sufficiently small that it does not require an annual audit. Grants and contracts with federal agencies have been managed to date through the member institutions that lead and subcontract with those entities. Expenditures made from the MetaArchive Services Group budget require the approval and signature of the Board Chair, the Treasurer, and the Executive Director. The Treasurer of the MetaArchive Services Group informs the Board on risk, benefit, investment, and expenditures related to the Cooperative's finances.

### **Funding Sustainability**

The Cooperative structure is designed to provide for the basic expenses related to the function of the network and to maintain a low overhead for the services it provides to its members by distributing technical and organizational work across its membership whenever this is possible. Membership fees are intentionally set at an affordable rate for member institutions and are used to insure that the operational and administrative costs of operating the network are met, including hosting Steering Committee meetings, workshops, trainings, and equipment. In addition, a capital reserve for catastrophic losses is maintained. It is projected that the network will be self-sustaining by 2012, but that there will be an effort to keep membership fees low through supplementing this work with grant- and contract-based funding. To date, the Cooperative has been highly successful at such fundraising and the current grants and contracts (through Emory University) will fund operations until 2010. New funding from federal and private sources is projected to be in place to continue supporting new development efforts for the Cooperative well before 2010. The Steering Committee will continue to work with the Board of Directors to monitor and bridge gaps in funding and pursue opportunities to enhance current activities.

### **Risk of Capital Assets**

As a decentralized Cooperative, MetaArchive ensures that all equipment used in the network belongs to its member institutions; the Cooperative does not own these assets. Any equipment purchased by the Cooperative, such as disks and/or servers added to MetaArchive-LOCKSS caches within the network to encompass new collections, is donated in-kind to the member institutions. Individual institutions may apply for technical and financial assistance with the restoration of a member's servers,

software, and collections in the event of catastrophic loss. In such circumstances, the purchased equipment would again be donated to the member institution, not held as assets of the Cooperative.

## **Contracts, Licenses, and Liabilities**

### **Data Liabilities**

The risk of data loss in the network is low because all preservation sites serve as joint custodians of ingested and preserved content. The content is distributed across multiple caches, such that each file is replicated at least six times in the network and stored in a geographically distributed manner. Members acknowledge that the MetaArchive, like other technology-based efforts, is a “fail-safe” method and cannot guarantee the preservation and migration of materials. Each member institution agrees in the Membership Agreement to hold full responsibility for protecting their own content through alternate means beyond those copies deposited in the MetaArchive Preservation network.

The MetaArchive Cooperative strives diligently to comply with copyright law. The MetaArchive network is a “Dark Archive” and therefore is not accessible outside of preservation routines and requirements. It is only available to Cooperative members for purposes of preservation and to replace an originating institution's local files when necessary. Each originating institution agrees to bear responsibility for resolving copyright and intellectual property issues locally in the Membership Agreement.

## **Conclusion**

The Cooperative has been legally independent since 2007, and prior to that time, was in operation from 2004-2006 under the auspices of one of its member institutions, Emory University. Therefore, the Cooperative's organizational infrastructure is still relatively new. As the Cooperative grows and evolves, its members will continue to systematically plan for and document the criteria that ensure the sustainability, performance, and accountability of the Cooperative.

## **Notes on this document**

This document was completed by the MetaArchive Cooperative Steering Committee on 2009-03-03, edited by Matt Schultz in 2009-10, and approved by the Steering Committee in 2009-10.

## **Appendix A: Biographical Information on Cooperative Leadership**

**Martin Halbert**, President of the MetaArchive Services Group, Emory University. Dr. Halbert is Director of Digital Innovations at the Emory University Libraries and directs all digital library services and systems functions for the Emory General Libraries. He is responsible for researching and leading library information technology initiatives, including all digital scholarly communication projects of the MetaScholar Initiative (<http://metascholar.org>). Dr. Halbert provides a leadership role within the library for computer systems operations, development, planning, and integration, and is the principal investigator for research projects totaling more than \$4M. As the founder of the MetaArchive Cooperative and the founding President of the MetaArchive Services Group, Martin provides a solid voice of leadership to our distributed digital preservation network development and sustainability efforts.

**Katherine Skinner**, Executive Director of the MetaArchive Services Group and Program Manager of the MetaArchive Cooperative, Emory University. Dr. Skinner is the Digital Projects Librarian at Emory University and provides leadership for the university's digital projects that are supported through grants or other sponsored funding sources. She is responsible for developing new funding streams and reaching out to faculty members to help them think about how digital media can enrich their scholarly pursuits. In this role, she has coordinated efforts involving interdisciplinary interest groups from more than three dozen universities worldwide, including faculty members (in the sciences, social sciences, and humanities), information technologists, librarians, curators, and campus administrators. She is also the co-principal investigator on preservation and access projects totaling more than \$2M.

**Tyler Walters**, Treasurer of the MetaArchive Services Group and Sustainability Director of the MetaArchive Cooperative, Georgia Tech. As the Associate Director, Technology and Resource Services for the Georgia Institute of Technology Library and Information Center, Tyler Walters provides leadership, vision, and expertise in digital library programs, information technologies, electronic resources management, metadata, and archives and records. He brings this expertise to the MetaArchive Cooperative in myriad ways in his advisory capacity, including assisting with technology planning, drafting the Cooperative Charter and Membership Agreement, and having a strong engagement in its sustainability planning.