

A workflow is being devised for e-journals that come to notice without having been assigned an ISSN, and to cater for use of the print ISSN using the new linking identifier, the ISSN-L.<sup>9</sup> At present, the intention is to follow ISSN rules and only include as e-journals those serials that were issued in digital format (i.e. “born” digital), and not “digitised journals” which were originally issued in print format, although this is actively being reviewed for the purpose of this project.

Title-level metadata on serials is essential but it is the *article* that is the information object of desire. Libraries will want to know the extent of preserved content for a given title, in order therefore to know which articles are preserved. This is more complex and, as such, has been relegated to the second phase of the PEPRS project. Provisional thinking is to create four date fields for each e-journal: earliest and latest known date of issue in digital format; earliest and latest known date of issue archived.

## (2) Metadata on preservation agencies and archiving action on each e-journal

Another key question is which archiving agencies to include in PEPRS project activity and over the longer term in the registry. The term “archiving” signals a potential widening of scope beyond that of digital preservation alone, to include “access continuity”: continuity of access to back content. This is triggered by a more recent UK report commissioned by JISC in which Morrow et al (2008)<sup>10</sup> reviewed the policies and practices of six digital preservation agencies.<sup>11</sup> It noted that some agencies focused primarily upon long-term preservation of the scholarly record, while the main emphasis for others was on “perpetual access” — the latter phrase used to refer to “continuity of access” to back content in an e-journal after the cancellation of a current subscription (“post-cancellation”) or as back-up for short-term failure.

Dependence upon leased access to content hosted at remote servers beyond the academy threatens continuity of access for researchers and students via their library. Challenging the very reasons for a library, this has become acute in the near term as financial pressures upon budgets for library materials lead to cancellations of subscriptions.

The main areas of policy interest need to be resolved into agreed, standard fields of information. Examples include title identifiers such as ISSN and title, date ranges, status of preservation, and access conditions. The next step would be the development of a common vocabulary for entries to assist users of the registry service who will want to compare attributes of preservation actions and summary descriptions of the agencies themselves.

Initially the plan was for the initial phase of the PEPRS project was to limit activity to three types of digital preservation agency: organizations operating at the international level (e.g., CLOCKSS and Portico); national

Director, EDINA National Academic Data Centre  
Causewayside House, 160 Causewayside, University of Edinburgh  
Edinburgh, EH 9 1PR, Scotland, UK  
Phone: +44 131 650 3301/2 • Fax: +44 131 650 3308  
<p.burnhill@ed.ac.uk> edina.ac.uk

Peter Burnhill

**BORN & LIVED:** United Kingdom.

**EARLY LIFE:** Economics/statistician, research into schooling, university lecturer.

**PROFESSIONAL CAREER AND ACTIVITIES:** Edinburgh University Data Library (1984), EDINA (1995/96); Director & Methodologist.

**HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS:** Looking back on how it changed/survived those years, as alternative paths of scholarly communication, informally but regularly trod, become increasingly recognised as the main highway. 🐘

libraries (e.g., British Library); and library consortia (e.g., UK LOCKSS Alliance). Were the scope of the registry to widen then the list might have to be revisited.

This and the diversity of use communities for the registry imply need and opportunity for cooperative inter-working, via interoperability, with third-party services providing information subscription status, likely organized on a territorial/nation-state basis rather than a global basis.

## (3) Data model and architecture

The registry service needs to support machine-to-machine use as well as a web-based user interface. Responsibility for specific fields of information is placed with the source best placed to deliver up-to-date information. A key feature of the data model is to establish dependence upon information sourced from the ISSN Register and from self-statement by the digital preservation (and archiving) agencies. The intention is to this exploits the “always on” presumption about the Internet in order to ensure up-to-date report by the preservation agencies, and also to keep an historic record of the statements made.

The use communities for direct access to a registry service are many and varied, especially if the international dimension is to be accounted for. One way to address this is to give equal priority to indirect access: that is, to the provision of a programming interface (API) that would provide interoperability to third-party facilities geared to serve specific-use communities across multiple locales and languages.

## (4) Business model and sustainability for the registry and its services

This registry and provision of its basic services must be for the long run, like its subject matter, digital preservation. An important part of the PEPRS project therefore is to identify costs and propose a business model for the registry service. It may also be necessary to

propose a form of governance.

Not surprisingly, the JISC-funded *Scoping Report* for this registry service touched on the matter of sustainability: “The archives themselves have to be sustainable over the long-term and to be of any use whatever, the registry must be equally long-lived.”<sup>12</sup> Discussion of this recommendation may seem premature, but the PEPRS project will be reviewed in 2009/2010 to assess whether the results of the project activity thus far and its business plan would justify the transition into service.

That might seem an appropriate open issue on which to end but perhaps this conclusion from the *Scoping Report* is more upbeat: “It seems to us that in order to gain the co-operation of the archiving organisations based around the world, which would be vital to its utility, the registry would have to be conceived as something which would serve the whole international scholarly community.”<sup>13</sup> The *Report* continues that the registry should be managed and governed “in such a way as to secure and maintain trust of both the library community and publishers.”<sup>14</sup>

## Request for Comment

PEPRS is a UK-funded project being carried out by a national academic data centre in partnership with an international standards body. In light of its potential to be international in scope and operation, and that any resultant registry service needs to exist over the long-run and to be of benefit across many sectors of the scholarly community, comments on issues raised, including governance and sustainability, are gratefully requested.<sup>15</sup> 🐘

endnotes on page 36

# MetaArchive: A Cooperative Approach to Distributed Digital Preservation

## Endnotes

1. Jones, Maggie. "Archiving E-Journals Consultancy (Final Report)," October 2003, accessed January 16, 2009; available from [http://www.jisc.ac.uk/uploaded\\_documents/ejournalsfinal.pdf](http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf); Internet.
2. Kenney, Anne R., Richard Entlich, Peter B. Hirtle, Nancy Y. McGovern, and Ellie L. Buckley. "E-Journal Archiving Metes and Bounds: A Survey of the Landscape," September, 2006, accessed January 16, 2009; available from <http://www.clir.org/PUBS/abstract/pub138abst.html>; Internet.
3. Sparks, Sue, Hugh Look, Adriene Muir and Mark Bide. "Scoping study for a registry of electronic journals that indicates where they are archived," **Rightscom and Loughborough University**, 2008; accessed January 2009), 33; available from <http://www.jisc.ac.uk/media/documents/programmes/preservation/ejournalregstudy.pdf>; Internet.
4. *ibid.*, 13.
5. SUNCAT includes serials' information from over 60 UK libraries, including the **British Library**, the **National Libraries of Scotland and Wales**, some of the largest Higher Education institutions in the UK and a number of specialist libraries. <http://www.suncat.ac.uk/>.
6. <http://edina.ac.uk>. In addition to SUNCAT, EDINA also provides organisational support for the **UK LOCKSS Alliance**, <http://edina.ac.uk/lockss/> following project work at the **Digital Curation Centre**. The **University of Edinburgh** is a **CLOCKSS Archive Node**.
7. <http://www.issn.org/>; Internet.
8. The **Directory of Open Access Journals (DOAJ)** requires ISSN assignment.
9. Accessed on January 19, 2009; available from <http://www.issn.org/2-22637-What-is-an-ISSN-L.php>; Internet.
10. Terry Morrow, Neil Beagrie, Maggie Jones and Julia Chruszcz, Tee Em Consulting, "A Comparative Study of e-Journal Archiving Solutions," May 2008. [http://www.jisc-collections.ac.uk/news\\_and\\_events/news\\_articles/ejournals\\_pr.aspx](http://www.jisc-collections.ac.uk/news_and_events/news_articles/ejournals_pr.aspx).
11. These six, **LOCKSS (Lots Of Copies Keep Stuff Safe)**, **CLOCKSS (Controlled LOCKSS)**, **Portico**, **e-Depot** (from the Dutch national library, Koninklijke Bibliotheek, KB), **OCLC Electronic Collections Online (ECO)**, and **British Library e-Journal Archiving Programme**, were selected on relevance to UK, with some relation to the 12 e-journal archiving initiatives included within the *CLIR Report* (op cit) as having met seven indicators of viability.
12. Kenney et al, op. cit., 3.
13. *ibid.*, 32.
14. *ibid.*, 16.
15. Comments should be sent to the author at <p.burnhill@ed.ac.uk>. A full version of this article is to appear in *Serials* in time for the UK Serials Group Meeting in April 2009.

by **Katherine Skinner** (Digital Projects Librarian, Emory University, 540 Asbury Circle, Atlanta, GA 30322; Phone: 404-783-2534; Fax: 404-727-0827) <katherine.skinner@emory.edu> <http://metaarchive.org>

and **Martin Halbert** (Director for Digital Innovations, Emory University, 540 Asbury Circle, Atlanta, GA 30322; Phone: 404-727-2204; Fax: 404-727-0827) <martin.halbert@emory.edu> <http://metaarchive.org>

*What role will the Library take in digital preservation?* On first glance, the question seems relatively easy to answer. As the library continues to transition from its centuries-long focus on print assets to a combination of print and digital resources, it will take an active role in the preservation of our digital cultural resources that is similar to that which it has long undertaken in the print realm.

Or will it?

Of late, many of us in the library field have become preoccupied with the concept of digital preservation — and rightly so. We wonder aloud about the forms that digital preservation will take, the amount it will cost, the rigor demanded in its implementation, and the feasibility of different organizational approaches to digital preservation.

But what does it mean to participate? How do we want to be involved? And what role(s) should we, as librarians and archivists, aspire to take in the realm of digital preservation?

Questions such as these led to the founding of the **MetaArchive Cooperative**, a collaborative network of institutions that have banded together to communally approach the challenges of preserving digital assets. The original six members founded this **Cooperative** due to their strong belief that libraries both could be and should be *actively engaged* in the creation and maintenance of their own digital preservation solution. They knew that alone, none of these institutions were likely to create and maintain — much less sustain — a robust digital preservation solution. However, they believed that if they approached the issue as a group and built a *shared* infrastructure, they could accomplish together what no one institution had the resources to achieve in isolation.

## The MetaArchive Cooperative: A Shared Digital Preservation Infrastructure

The **MetaArchive Cooperative** (<http://MetaArchive.org>) formed to enable cultural memory organizations to effectively and mutually preserve their archival digital assets *for themselves*. **MetaArchive** began in 2004 as one of the original eight initiatives contracted by the **Library of Congress** under the **National Digital Information Infrastructure and Preservation Program (NDIIPP)**.<sup>1</sup> The venture was led by **Emory University** in collaboration with **Georgia Tech**, **University of Louisville**, **Virginia Tech**, **Auburn University**,

**Florida State University**, and the **Library of Congress**. The network established by this group was the first major effort to build and operate a private implementation of the open source **LOCKSS (for Lots of Copies Keep Stuff Safe)** software for digital preservation (<http://www.lockss.org>), an approach that has since been termed a **Private LOCKSS Network**, or **PLN**. The **MetaArchive PLN** is a distributed preservation infrastructure that meets the **OAIS Reference Model** standards for repositories.<sup>2</sup>

Technically speaking, the foundation of the network is the open source **LOCKSS** software developed at **Stanford University**, which enables a group of **LOCKSS** caches, or node servers, to work together across geographical space to replicate and preserve content.<sup>3</sup> **MetaArchive** is the only **PLN** in operation thus far that does not depend on the **LOCKSS** team to administer the network; we run a separate cache manager (coded in collaboration with the **LOCKSS** team) to monitor our network. The **MetaArchive Cooperative** has created and layered additional modules on top of the **LOCKSS** framework to provide our members with administrative tools, including a conspectus database and the cache manager. The conspectus database enables members to capture collection-level metadata for preservation decisions and actions, and the cache manager serves as a monitoring tool for network-wide tracking and troubleshooting activities. We are in the process of packaging these open source software components for use by other **PLNs**, and plan to release this software through **SourceForge** next year.

The organizational framework that we have constructed has been as integral to our success as the technological platform upon which we have built our preservation services. After running the network for three years, we transitioned from a sponsored-funding-supported project to an independent, membership association in 2007. As part of this work, we founded a 501c3, the **MetaArchive Services Group**, to administer the **Cooperative**. All of the components of the network we run are owned and maintained by our member institution. This decentralized apparatus enables the **Cooperative** and its services to be *independent of each member* — our members learn how to run and operate their own preservation node for the network, building their internal

*continued on page 38*

knowledge of the preservation process. They also are given opportunities to contribute to the software development efforts undertaken by the **Cooperative**.

The mission of the **MetaArchive Cooperative** is to support, promote, and extend our collaborative approach to distributed digital preservation practices. We have made our organizational model available to others as an example of how to create shared digital infrastructure. To this end, we not only run our own network, but also provide training and consulting assistance to other groups that wish to found similar preservation networks. We host workshops and make all of our documentation freely available to other collaborative projects and programs.

Unlike the public **LOCKSS** network, where participant libraries preserve journal content in which they all have a vested interest, the collections in the **MetaArchive** network are the unique holdings of each participant library and archive. In other words, **MetaArchive's** members cannot rely on the incentive that drives participation in the public **LOCKSS** network — a shared body of content to which all subscribe and upon which all rely. Instead, the **MetaArchive** network requires a strong commitment between constituent institutions — each participates in order to preserve their own data in exchange for preserving other institutions' data.

So what *are* the drivers in this **PLN** scenario? Topping the list are a strong sense of community engagement and a strong belief in the library's cultural stewardship role. Our members share the conviction that libraries have a vested interest in preserving their own digital assets. Each has determined that they do not want to cede all of their digital preservation activities to external groups, and do want to participate in creating their own preservation solution. Building alone is a costly proposition, so these institutions have coupled their resources in order to achieve their preservation goals in a community-based effort.<sup>4</sup>

To enable this, **MetaArchive** formed as a *cooperative*, not a vendor. **MetaArchive's** members do not pay for services, but rather make an investment to create and sustain their own preservation infrastructure. The **Cooperative** is more than a technical solution for preservation. It also functions as a learning environment in which members gain experience in developing and enacting a full preservation plan for their assets. Each member both contributes to and benefits from the expertise and the technical infrastructure developed by the overall community. In keeping with these principles, membership fees are kept at the absolute minimum required for the operation of the **Cooperative**, and range from \$300 to \$5K per year, together with a fee of \$2 every 3 years per 1 GB of content contributed. These minimal storage fees cover the expense of replicated storage space for the network at cost.

*continued on page 40*

## against the grain people profile

### Katherine Skinner & Martin Halbert

Skinner is Digital Projects Librarian  
Halbert is Director for Digital Innovations  
Emory University, 540 Asbury Circle, Atlanta, GA 30322  
Skinner Phone: (404) 783-2534 • <katherine.skinner@emory.edu>  
Halbert Phone: (404) 727-2204 • <martin.halbert@emory.edu>  
Fax for both: (404) 727-0827 • <http://metaarchive.org>

**PROFESSIONAL CAREER AND ACTIVITIES:** Dr. Katherine 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. 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 a Co-Principal Investigator on the **SouthComb Project (Andrew W. Mellon Foundation)**, the **MetaArchive Project (Library of Congress)**, and the **MetaArchive: A Sustainable Digital Preservation Service Project (NHPRC)**. She's a founder and an editorial board member of the peer-reviewed Internet journal, *Southern Spaces* and a consultant on the **TransAtlantic Slave Trade Database Projects at Emory (NEH)**. She recently co-edited a monograph entitled *Strategies for Sustaining Digital Libraries* with **Martin Halbert**.

**Katherine** has a Ph.D. from **Emory University** and a B.A. from the **University of North Carolina at Chapel Hill**. **Katherine** worked in academic/non-profit marketing at **Mindpower, Inc.**, prior to starting graduate school in 1999. She is married and has a son.



**PROFESSIONAL CAREER AND ACTIVITIES:** Dr. Martin Halbert is Director for Digital Innovations at **Emory University**. **Martin** directs a variety of digital library research projects and services 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://www.MetaScholar.org>). The staff of Digital Innovations includes librarians, scholarly communication analysts, and systems programmers/support staff. He is the principal investigator for research projects totaling \$5M. He recently established the **MetaArchive Digital Preservation Network** (<http://www.MetaArchive.org>), a cooperative of ten university libraries and other cultural memory organizations acting in concert with the **Library of Congress** as part of the **National Digital Preservation Program**.

**Martin** has a Ph.D. from **Emory University**, an M.L.I.S. from the **University of Texas**, and a B.A. from **Rice University**. **Martin** has worked in library administration and systems positions at **Emory University** and **Rice University**. He served as an **ALA/USIA Library Fellow** in Estonia assisting with the automation of the **Tartu University Library**. He has also worked as a consultant for the **IBM Corporation**, and as a programmer for the **University of Texas**. He is married and has three children.



**HOW/WHERE DO WE SEE THE INDUSTRY IN FIVE YEARS:** We believe that academic libraries are poised at a fairly serious crossroads, particularly now that we are in the throes of a major recession. If librarians choose to function primarily as intermediaries seeking to coordinate the access and preservation functions for our campuses through outsourcing contracts to external entities, the value we add to the academic enterprise will greatly diminish and we will ultimately become a study hall / museum of aging physical media relics. If we alternatively choose to meet the digital information needs of our campuses ourselves, the value of research libraries will grow as vital and experimental arenas of scholarly inquiry and engagement with knowledge. The latter path is less charted and requires more experimentation, but is (we think) clearly preferable as a course for the future. 🌱

We believe that it is unlikely that any similar replicated digital preservation service can be established at lower costs.

The **Cooperative** membership structure is comprised of three tiers: Contributing Members, Preservation Members, and Sustaining Members. Contributing Member sites are smaller institutions interested in using the shared network infrastructure to preserve digital content but lacking the capacity to operate any technical infrastructure of their own. Preservation Member sites are responsible for the basic ongoing network activity of preserving digital content. At a minimum, every preservation site must include responsible staff and a minimally configured node server. Sustaining Member sites are responsible for hosting a preservation node and also for leading the **Cooperative** through steering committee participation and through developing the technical systems that enable the preservation network.

### Decentralized Preservation Practices

A key strength in our approach to preservation is the distributed nature of both our technical and organizational infrastructures. **MetaArchive** centers on the principle, “lots of copies keep stuff safe.” We believe this to be true, not just in terms of replicating content and distributing it across a geographically dispersed network, but also in terms of replicating *knowledge* and distributing it among our members. To this end, major systems knowledge is not simply held by a central staff, but is deliberately spread out across our member institutions’ technical staff. Our sustainability is increased through this distributed knowledge — we are not dependent on central staff members, but rather have shared expertise to draw on across all member institutions; 2) we have a built-in system of checks and balances, as network monitoring is conducted by a committed core of Preservation and Sustaining members; and 3) the Cooperative does not need to incur the costs associated with employing and hosting central staff — which allows the **Cooperative** to keep its costs low and provides a major savings for our member institutions.

In keeping with this philosophy, we also do not rely on the **LOCKSS** team to administer our network. This is a major difference between the **MetaArchive Cooperative** and other PLNs, which have largely opted to have **LOCKSS** manage and maintain their networks. We do *benefit* greatly from the **LOCKSS** team,

both in terms of the regular updates they provide to the **LOCKSS** daemon and also in terms of the technical expertise they share with our central and distributed staff members, but we chose to build on an open source framework specifically because we believe this model offers the best odds for long-term sustainability. The overall **LOCKSS** community (including myriad PLNs) is already strong and it’s growing stronger. We believe that a solid open source development community could sustain and maintain the **LOCKSS** software if called upon to do so, and we gave intentionally built a framework that relies only on this community, not on any one group within it.

### Preservation and Institutionalization

Institutions form in order to address specific needs that are not already being met within the existing environment. This is to say that when dominant and traditional business practices (and libraries are a business, whether we think of ourselves as such or not) fail to meet community or market needs, it opens a space within which new institutions with new approaches may flourish. Witness **Google**, **Elsevier’s** journal services, and myriad other examples and exemplars that have already emerged to serve the information management and access needs of the digital age.

The library as an institution continues to serve many of the needs of its constituents — it is not in danger of perishing outright. However, it has not yet proven itself a serious contender in the digital realm. Scholars as well as the public are increasingly turning to companies such as **Google** to “to organize the world’s information and make it universally accessible and useful,” an access role that for centuries belonged primarily to the library field. To whom will these groups turn when they seek to preserve their digital assets, another core mission of the information science field? Will they turn to cultural memory organizations such as libraries and archives, or to corporations such as **Amazon** and **Google**? And should we, as cultural stewards, care so long as the preservation channel adequately provides for the needs of our institutions and our constituents? Is there a difference between commercially driven solutions and those created in the not-for-profit environment?

As libraries, our work is driven by the desire to maximize our stakeholders’ long-term access to materials, not by a desire to maximize profit for stockholders. This is a highly significant distinguishing factor and one that we cannot afford to take for granted. Cultural memory organizations are not, on the whole, profit-

making enterprises. They are funded by tax dollars, foundations, and parent institutions whose constituents we serve. Our reason for being is to serve a public good — making our cultural assets, from books to datasets, accessible to the public over the long haul. If we cease to perform that function, instead outsourcing it to external parties, we are putting both our field and our cultural resources in precarious positions.

This is not to say that we should not outsource *any* of the digital preservation work we undertake. Just as is true in the print and physical artifacts domain, there will doubtless be portions of this work that are well suited to the work of external parties and too expensive to build in house. But, in order to know which parts are suited to in-house work and which parts should be outsourced, we must begin to explicitly engage in our own digital preservation solutions.

So from where will the successful approaches to digital preservation emerge? Sociologists tell us to watch the fringes, not the center, for seismic changes.<sup>5</sup> Today, one of these fringes might well be the library — or in this case, many libraries, banding together in collaborative, cooperative ways to accomplish the preservation of their unique resources in a communally owned network environment that they run for themselves. 🐜

### Endnotes

1. **Library of Congress**. 2002. “Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program.” Accessed December 29, 2008, from [http://www.digitalpreservation.gov/library/resources/pubs/docs/ndiipp\\_plan.pdf](http://www.digitalpreservation.gov/library/resources/pubs/docs/ndiipp_plan.pdf).
2. Consultative Committee for Space Data Systems. 2002. Reference Model for an Open Archival Information System (OAIS).” CCSDS 650.0-R-1–Blue Book. Accessed December 28, 2008 from <http://public.ccsds.org/publications/archive/650x0b1.pdf>.
3. **David S. H. Rosenthal, Thomas Robertson, Tom Lipkis, Vicky Reich, and Seth Morabito**. 2005. “Requirements for Digital Preservation Systems: A Bottom-Up Approach.” *D-Lib Magazine* 11, no. 11. Accessed December 29, 2008 from <http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html>.
4. For more on the use of collaboration to achieve library goals, see **Michael Day**. 2008. “Toward Distributed Infrastructures for Digital Preservation: The Roles of Collaboration and Trust the Roles of Collaboration and Trust.” *The International Journal of Digital Curation*, 1, Vol. 3: 15-28. Accessed December 28, 2008, from <http://www.ijdc.net/ijdc/article/view/60/61>.
5. **Huseyin Leblebici, Gerald R. Salancik, Anne Copay, and Tom King**. 1991. “Institutional Change and the Transformation of Interorganizational Fields: An Organizational History of the U.S. Broadcasting Industry.” *Administrative Science Quarterly* 36: 333-363.

# Distributed Globally, Collected Locally: LOCKSS for Digital Government Information

by **Daniel Cornwall** (Head of Information Services, Alaska State Library, P.O. Box 110571, Juneau, AK 99811-0571; Phone: 907-465-1315; Fax: 907-465-2665) <dan.cornwall@alaska.gov> <http://library.state.ak.us>

and **James R. Jacobs** (International Documents Librarian, Stanford University Library, 123B Green Library, Stanford University, Stanford, CA 94305; Phone: 650-725-1030; Fax: 650-723-9348) <jrjacobs@stanford.edu> <http://jonssonlib.stanford.edu>

## Introduction

Ever since the **Government Printing Office (GPO)** brought **GPO Access** online in 1993 in order to make government information accessible on the Web, some librarians and others have dreamt of a system that would extend the mostly successful, 150 year old geographically distributed **Federal Depository Library Program (FDLP)** model to the digital world.<sup>1</sup> Today, thanks to the efforts of **Carl Malamud**, the **Stanford-based Lots of Copies Keep Stuff Safe (LOCKSS)** team, and 15 libraries around the country, a successful model for the digital **FDLP** has been launched.<sup>2</sup>

This paper will describe the **LOCKSS** model of digital preservation and why that model is beneficial to apply to the realm of digital government information. Next, we will illuminate **Carl Malamud's** herculean efforts toward better access to government information. We will then discuss how we've built the **USDocs Private LOCKSS Network (USDocsPLN)** using those documents harvested by **Malamud**. The paper concludes with a call to action.

## Benefits of a Distributed Collection

The subject of digital preservation is of vital concern to libraries and other cultural institutions; Organizations like the **Library of Congress**, **U.S. National Archives (NARA)**, **Internet Archive** and many others have been working on solutions to preservation and long-term access to digital information.<sup>3</sup> Within the government documents library community, there is one school of thought that local digital collections of government documents are wasteful duplication of resources.<sup>4</sup> In this view, **GPO's** assumption of storage and preservation duties has freed libraries from the burden of being documents storehouses to let them focus solely on public services. As the **LOCKSS** model demonstrates, this school of thought is mistaken and in fact will endanger long-term access to and preservation of government information.

The **LOCKSS** model<sup>5</sup> is a proven distributed preservation model based on a peer-to-peer (P2P) architecture<sup>6</sup> in which each node in the **LOCKSS** network locally hosts an exact replica of the content being preserved.

The open-source **LOCKSS** software then compares content on each host and repairs any differences, thus assuring preservation and authenticity. Approximately 200 libraries in the global public **LOCKSS** network have successfully preserved e-journals and publisher content for over ten years.<sup>7</sup> Fifteen **LOCKSS** libraries have now embarked on a project to apply this successful model to government documents.

There are myriad reasons why a distributed digital preservation system for government information is necessary. Among them are: protection from natural disaster, server outage, etc.; assurance of authenticity; prevention of surreptitious withdrawal or tampering of information; and building local services for local collections.

A system of geographically disbursed digital collections provides resiliency in the aftermath of a disaster. After Hurricane Rita, the **McNeese University Library** in Lake Charles, LA, lost a large amount of their physical collection, including many Louisiana state documents.<sup>8</sup> Imagine that instead of physical documents, **McNeese** had held the ONLY copy of digital documents and that other LA libraries had relied on **McNeese** rather than building their own digital collections. When the hurricane hit and washed away **McNeese's** servers, all libraries in Louisiana would have lost access. Even if **McNeese** followed best practices and kept an offsite backup of their materials, libraries might still be without access for weeks or months while waiting for **McNeese** to come back online.

While this imaginary wipeout of LA state documents did not happen, we face that very real situation with digital federal documents. **GPO** has been tasked since 2001 to provide a mirror server for **GPO Access**.<sup>9</sup> As of this writing, **GPO** has still not done so. If anything happens to **GPO's** servers, we'll lose access to hundreds of thousands of born-digital federal documents.

Local digital collections also insulate against Internet outages and server downtimes. According to the **FDLP-L** archives, **GPO** servers were taken offline seven times in 2007. During those periods, no one could access **GPO's** documents. With a **USDocsPLN**

in place, users would not notice down times because they would be automatically rerouted to their nearest collection.

Authenticity, a critical feature to have in any trusted government information infrastructure, is enhanced with a distributed collection. Digital government information has been altered without notice.<sup>10</sup> While there are no documented instances of this happening to **GPO** content, the potential is there as long as **GPO's** servers continue to be the exclusive source for government information. Multiple copies on geographically disparate servers allow possible alterations to be inspected and corrected, thus protecting against deliberate tampering. The **USDocsPLN** explicitly does this. Research suggests that only a large-scale network attack lasting months could successfully change content stored in a **LOCKSS** network.<sup>11</sup>

Related to the problem of alteration is that of outright withdrawal. In the **FDLP** world of distributed physical collections, there are processes in place to protect against this. In order to withdraw a publication from depository collections, **GPO** must notify the holding libraries of the item to be withdrawn and order them to either return the publication to **GPO** or destroy it. Sometimes withdrawal is appropriate and libraries comply.<sup>12</sup> But in some instances, publications are withdrawn needlessly or explicitly to protect the government's reputation. In these instances, depository librarians have been known to create a loud hue and cry that usually results in the withdrawal order being cancelled.<sup>13</sup> In the current centralized digital model, this protection does not exist. No public process need be followed. A simple delete command is all it takes. A cached copy can sometimes be found in **Google** or the **Internet Archive's Wayback Machine**, but often not.

Besides the preservation aspects, building local digital collections can serve to provide unique services for local communities. For instance, text mining is becoming a useful way of analyzing documents either one at a time or in large collections. It could be as simple as a tag cloud of a speech<sup>14</sup> or as complex as analyzing patent applications.<sup>15</sup> Local digital collections could provide researchers with a full or selected amount of **GPO Access** to analyze without requiring access to **GPO** servers that could potentially impact security or performance. Those collections could also be repurposed and remixed to facilitate new ways of analyzing information and creating new bodies of knowledge.

*continued on page 43*

### Libraries Need a Little Help From Their Friends

Libraries have traditionally taken an active role in collecting content to meet the needs of their local user base. This was a straightforward process in the print world, with vendors galore and, in the case of U.S. government documents, the **FDLP**. In the digital world, things are much murkier, the process a little more convoluted. The responsibility to collect and preserve content remains but the process is more challenging; on the open Web, there are no vendors to pull together disparate publishing streams or depository systems for easy inclusion into local library collections. On the Internet, libraries need to implement a more aggressive approach toward collecting Web-based materials as well as identifying new partners in their efforts — libraries must rely on the kindness of strangers and library fellow travelers.

One such fellow traveler to the government documents community is **Carl Malamud**.<sup>16</sup> **Malamud** is an Internet- and open government activist who runs the Website, *public.resource.org*. Since the U.S. government has been producing digital public domain government information, **Malamud** has been successfully shaking it free from government control and onerous access fee structures and making it more accessible to citizens. **Malamud's** overarching goal is to release government information into the open so that others can build more advanced interfaces and facilitate better access to the workings of our governments.<sup>17</sup>

His first campaign led to the creation of the **Securities and Exchange Commission's** **EDGAR** database of SEC filings and corporate disclosure documents (which has recently had a name change to **IDEA**).<sup>18</sup> He has since, in his efforts to "open source America's operating system,"<sup>19</sup> set his sights on Federal and State Courts and case law, State and municipal codes, **U.S. Copyright Office**, **National Technical Information Service (NTIS)** videos, **Government Accountability Office (GAO)** legislative histories, and, of most interest to Government Documents Librarians participating in the **FDLP**, documents from the **GPO**.

The **GPO** is the official publisher of the U.S. Government and manages the **FDLP**. They publish and distribute to libraries publications from 21 federal agencies as well as such integral publications as the Federal Register, Congressional Record, Congressional Reports, Bills, documents and Hearings, Public Laws, Papers of U.S. Presidents and much more. **GPO Access** is built on an older technology called WAIS with a very primitive user interface and limited search capabilities. For that reason, **Malamud**, with the assistance and cooperation of the **GPO**, harvested **GPO Access** documents from **GPO** servers in late 2007 and made them accessible/downloadable via BitTorrent, Rsync, HTTP and FTP. Those documents comprise 200+ gigabytes of data from 1991-2007 amounting to 5,177,003 PDF

## against the grain people profile

Head of Information Services, Alaska State Library  
P.O. Box 110571, Juneau, AK 99811-0571  
Phone: (907) 465-1315 • Fax: (907) 465-2665  
<dan.cornwall@alaska.gov>  
<http://library.state.ak.us> <http://freegovinfo.info>

### Daniel Cornwall

**BORN & LIVED:** Born in Los Angeles, California. Lived in San Antonio, Texas and Panama City, Florida. For past ten years I've lived in Juneau, Alaska.

**FAMILY:** Wife, **Louise**, father in California, four brothers around the country and one sister in Canada.

**PROFESSIONAL CAREER AND ACTIVITIES:** Worked at **Alaska State Library** for ten years, promoted twice. Member of **Alaska Library Association** and **American Library Association**. Frequent presenter at state conferences. Heavy user and promoter of the **ALA** government documents wiki. Project coordinator for state agency databases across the 50 States. Co-blogger/administrator of free government information.

**IN MY SPARE TIME I LIKE:** I enjoy reading, blogging, hiking, cooking and outdoor astronomy when Juneau's weather permits it.

**FAVORITE BOOKS:** *The Mage Storms* series by **Mercedes Lackey**, *The Post-American World* by **Fareed Zakaria**, *A Vow of Conversion* by **Thomas Merton**, *Physics for Future Presidents* by **Richard A. Muller**.

**PET PEEVES/WHAT MAKES ME MAD:** Accusations without evidence (by governments), statements that suggest a lack of basic knowledge, lack of self-consistency, and yes, librarians who buy into the "travel agent" theory of librarianship w/o control of critical resources.

**PHILOSOPHY:** The less secrecy, the better for everyone. Do as you would have others do to you. Ask nothing of others that you are not prepared to do yourself.

**MOST MEANINGFUL CAREER ACHIEVEMENT:** Getting Alaska state agency monographs into the **LOCKSS** system back in 2005.

**GOAL I HOPE TO ACHIEVE FIVE YEARS FROM NOW:** Working with others in my state, I hope that most Alaskans will become aware of the rich feast of state and federal government information that is available for their taking.

**HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS:** I see librarianship as a thriving, user-centered profession. Librarians will learn how to bankroll the trust they are given by their brick and mortar patrons and become trusted online guides. They will also be curators of specialized local collections in all media for the education and convenience of their patrons. I also see public and special librarians following their academic cousins and giving much more instruction about resources and research than they do today. This instruction will be where the patron is, whether in-person or online. This will happen because libraries will gain fund as they are seen as transformational places. 🌱



pages, 54,600 GAO Reports, 448,496 Congressional Reports and more. It's these **GPO** documents upon which the **USDocsPLN** has so far focused.

#### Current Status

The **USDocsPLN** is now up and running. The 200+ gigabytes of digital documents have been downloaded from **Malamud's** site (<http://bulk.resource.org/gpo.gov>) and distributed

among the 15 partners in the project, where they will be preserved within the **LOCKSS** network. This was an extremely cost-effective project as 1 terabyte (which equals 1,000 gigabytes) of storage is now below \$200, hardware is typically less than \$1,000, and there is only minimal administrative cost once the **LOCKSS** box has been configured. The group will continue to evaluate and add to the network other

*continued on page 44*

collections of digital government documents, including, but not limited to, other collections on [public.resource.org](http://public.resource.org).<sup>20</sup>

Participating libraries in the LOCKSS-USDocs private network include:

- Alaska State Library
- Amherst College
- Georgia Institute of Technology
- Library of Congress
- Michigan State University
- North Carolina State University
- Northeastern University
- Rice University
- Stanford University
- University of Alabama
- University of Illinois/Chicago
- University of Kentucky
- University of Wisconsin-Madison
- Virginia Tech
- Yale University

While it's exciting to have this large group of research libraries participating in the US-DocsPLN, we realize that the cost of being a LOCKSS Alliance member may be a barrier for some libraries — fees range from \$1,000 to \$10,800 per year, depending on institution size. We are working to increase the number of LOCKSS Alliance members in order to distribute software and other development costs across a larger network. More members mean less cost per institution.

### How You Can Help

The preservation of federal documents is too important to be left to the federal government alone; we have the makings of a viable system to preserve digital government publications. There are several ways you can help.

- **Join our private LOCKSS Network.** Join the LOCKSS Alliance, get a server for under \$1,000, and contact us. The more servers in the USDocsPLN, the merrier.
- **Notify us of collections of electronic federal documents.** LOCKSS staff can show you how easy it is to allow LOCKSS to ingest and preserve your materials.
- **Attack the root problem.** Demand members of Congress legislate and fund a system that will ensure that GPO proactively deposits publications and data through the FDLP and other interested partners. While the USDocsPLN project is a good start and an excellent ad-hoc effort, it should be the government's responsibility to put information in the hands of taxpayers. We should not have to be prying it out of the government's hands. A distributed digital FDLP benefits everyone. 🐼

endnotes on page 45

## against the grain people profile

International Documents Librarian, Stanford University Library  
123B Green Library, Stanford University, Stanford, CA 94305  
Phone: (650) 725-1030 • Fax: (650) 723-9348  
<jrjacobs@stanford.edu> <http://jonssonlib.stanford.edu>  
<http://freegovinfo.info> <http://radicalreference.info>  
<http://questioncopyright.org>

### James R. Jacobs

**BORN & LIVED:** Englewood, NJ. See bio at <http://freegovinfo.info/about/jrjacobs> for more.

**EARLY LIFE:** Northeast states; lots of soccer, tennis, baseball etc.

**FAMILY:** Spouse, mother/father in PA, youngest of four siblings (brother in NYC, sister in Cleveland, OH, sister in Groton, NY).

**PROFESSIONAL CAREER AND ACTIVITIES:** I've worked in libraries since I was 15 when I worked in a small public library in Homer, NY. Professionally, I've been a Government Documents Librarian since 2002, first at UC San Diego and now at Stanford University. I'm active in ALA's Government Documents Roundtable (GODORT) and am a moderator for govdoc-I, the primary listserv of government information librarians.

**IN MY SPARE TIME I LIKE:** Spare time? J information activist/blogger with <http://freegovinfo.info>, <http://radicalreference> and <http://questioncopyright.org>. I also like to dabble in open-source software, hike (urban and rural), eat good food, and read.

**FAVORITE BOOKS:** *Sometimes a Great Notion*, *Lord of the Rings*, *Baroque Cycle*, *Leaves of Grass*, *Dharma Bums*, *Another Roadside Attraction*, *Tao Te Ching*, *Cat's Cradle*, *People's History of the United States*.

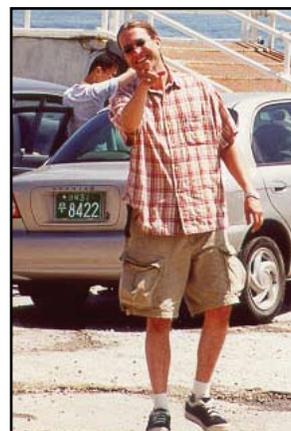
**PET PEEVES/WHAT MAKES ME MAD:** People who say, "it can't be done" instead of imagining the possibilities; people who act selfishly.

**PHILOSOPHY:** Information wants to be free; librarians to facilitate that process.

**MOST MEANINGFUL CAREER ACHIEVEMENT:** Writing "Government Information in the Digital Age: The Once and Future Federal Depository Library Program" which has had over 15,000 downloads; building Radical Reference and Free Government Information to give free reference to activists and independent journalists and advocate for access to and preservation of digital government information.

**GOAL I HOPE TO ACHIEVE FIVE YEARS FROM NOW:** That a large number of libraries have the technical and administrative wherewithal to be building local digital collections, sharing with each other and building services to increase access and shine light on government activities.

**HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS:** I'm an optimist. I see libraries continuing their vital work of preserving and giving free access to society's vital information in all formats. I also see them expanding their trusted position by leveraging the Web to make more information available to more people. 🐼



# How Do You Improve Upon The Essentials?

By adding more. More to see, more to do, more tools. Project MUSE has always been an essential online resource for faculty and students. Now, our new website offers even more, with greater functionality and more efficient search and discovery tools. And it's easier than ever to use.

With Project MUSE, you get:

- Highly respected humanities and social science journals
- The most current, full text, peer-reviewed content
- Affordable pricing and great value

So make us part of your essentials kit for online research.

FREE 45-day trial at <http://muse.jhu.edu/trialrequest>



Project  
**MUSE**<sup>®</sup>

Today's Research.  
Tomorrow's Inspiration.

<http://muse.jhu.edu>

## Endnotes for Distributed Globally ... from page 44

1. The **Federal Depository Library Program (FDLP)** was established by **Congress** in 1813 to ensure that the American public has access to its Government's information. Since then, depository libraries have been collecting, organizing, maintaining, preserving, and assisting users with information from the Federal Government. Today there are nearly 1,250 **FDLP** libraries across the U.S.
2. For a more complete discussion of the digital **FDLP**, please see: **James A. Jacobs, James R. Jacobs, and Shinjoung Yeo**, "Government Information in the Digital Age: The Once and Future Federal Depository Library Program" (2005). *Journal of Academic Librarianship*. 31 (3), pp. 198-208. doi:10.1016/j.acalib.2005.02.006. Postprint available free at: <http://repositories.cdlib.org/postprints/657>. Accessed December 26, 2008.
3. **Berman, Francine**. "Got data? A Guide to Data Preservation in the Information Age." *Communications of the Association for Computing Machinery (ACM)*. 51, no. 12 (2008): 50-56. <http://portal.acm.org/citation.cfm?id=1409360.1409376&coll=portal&dl=ACM&idx=J79&part=magazine&WantType=Magazines&title=Communications>. Accessed December 28, 2008.
4. **Durant, D.M.** (2004) The Federal Depository Library Program: Anachronism or Necessity? [Journal Article (Print/Paginated)] <http://eprints.rclis.org/2313/>. Accessed November 30, 2008.
5. Much has been written about the **LOCKSS** project and its successful model of long-term preservation of digital content: <http://lockss.org/lockss/Publications>. Accessed December 26, 2008.
6. Peer-to-peer File Sharing. Wikipedia. [http://en.wikipedia.org/wiki/Peer-to-peer\\_file\\_sharing](http://en.wikipedia.org/wiki/Peer-to-peer_file_sharing). Accessed December 26, 2008.
7. A full list of journals and publishers can be found on the **LOCKSS** site: [http://lockss.org/lockss/Publishers\\_and\\_Titles](http://lockss.org/lockss/Publishers_and_Titles). Accessed December 26, 2008.
8. **Blakeley, Rebecca** (2008, October). "What Was Lost, Now is Found: Using Digital Repositories to Rebuild What Hurricane Rita Destroyed." <http://www.slideshare.net/reblakeley/dlc-presentation-2008-presentation>. Accessed November 30, 2008.
9. **Tulis, S.** (2001) Fall Federal Depository Conference/Depository Library Council Meeting HighlightsGOVDOC-L archives at <http://lists.psu.edu/cgi-bin/wa?A2=ind01&L=GOVDOC-L&P=R173102>. Accessed November 30, 2008.
10. **Althaus, S. & Letaru, K.** (2008, November). Airbrushing History, American Style. [http://www.clinecenter.uiuc.edu/airbrushing\\_history/](http://www.clinecenter.uiuc.edu/airbrushing_history/). Accessed November 30, 2008.
11. **Giuli, P.**, et al. (2005) Attrition Defenses for a Peer-to-peer Digital Preservation System. *Proceedings of the 2005 USENIX Annual Technical Conference*, April, 2005. <http://www.eecs.harvard.edu/~mema/publications/hotel.pdf>. Accessed November 30, 2008.
12. An incomplete list of government documents withdrawn from the **FDLP** 1987-2000. [http://www.thememoryhole.org/foi/yanked\\_govt\\_docs.htm](http://www.thememoryhole.org/foi/yanked_govt_docs.htm). Accessed December 26, 2008.
13. Department of Justice Rescinds Order to Have Libraries Withdraw/Destroy Publications (2004, August) Posting on *Resource Shelf*. <http://www.resourceshelf.com/2004/08/02/dept-of-justice-rescinds-order-to-have-libraries-withdrawdestroy-publications/>. Accessed December 26, 2008.
14. Volunteers at **Free Government Information** have been testing out a Web service called **TagCrowd** (<http://tagcrowd.com>) to create tag clouds of text. Please see the following for examples: <http://freegovinfo.info/taxonomy/term/364>. Accessed December 26, 2008.
15. **March Jr., R.** (2008, February). Text Mining Patent Literature. *Searcher*, 16(2), 30-41. Accessed November 30, 2008, from **MasterFILE Premier** database.
16. For more on **Carl Malamud**, see: **Singel, Ryan**. "Online Rebel Publishes Millions of Dollars in U.S. Court Records for Free." *Wired Magazine*. December 12, 2008. [http://www.wired.com/politics/onlinerights/news/2008/12/open\\_pacer?currentPage=all](http://www.wired.com/politics/onlinerights/news/2008/12/open_pacer?currentPage=all). Accessed December 14, 2008.
17. This is increasingly happening. **Witness OpenCongress** [<http://www.opencongress.org/>], **GovTrack.us** [<http://www.govtrack.us/>] and various projects from the **Sunlight Foundation** [<http://www.sunlightfoundation.com/projects/>]. These are but a few projects currently building Web2.0 mashups using government information.
18. **Interactive Data Electronic Applications (IDEA)** is the **SEC's** updated Web application of **SEC** data intended to replace **EDGAR**. <http://idea.sec.gov/>. Accessed December 26, 2008.
19. For more on **Malamud's** efforts to "open source America's operating system," please see <http://freegovinfo.info/taxonomy/term/530>. Accessed December 26, 2008.
20. See the [gpo.gov](http://bulk.resource.org/gpo.gov/Harvest_Report.html) harvest report for a complete listing of documents. [http://bulk.resource.org/gpo.gov/Harvest\\_Report.html](http://bulk.resource.org/gpo.gov/Harvest_Report.html). Accessed December 26, 2008.

# The Alabama Digital Preservation Network (ADPNet)

by Aaron Trehub (Assistant Dean for Library Technology, Auburn University Libraries) <http://adpn.org/>

The Alabama Digital Preservation Network (ADPNet) is a geographically distributed digital preservation network for the state of Alabama — the first working statewide Private LOCKSS Network (PLN) in the United States. Inspired by Auburn University's experience with another LOCKSS-based initiative, the MetaArchive Cooperative, ADPNet was designed from the outset to be a low-cost, low-maintenance digital preservation solution for libraries, archives, museums, and other cultural-heritage organizations in Alabama. It was also designed to be a model for other states and consortia that are interested in exploring a distributed digital preservation solution. Work on ADPNet began in late 2006 under a two-year National Leadership Grant from the Institute of Museum and Library Services (IMLS). That grant ended in September 2008, and ADPNet is now a self-sustaining, member-managed program operating under the auspices of the Network of Alabama Academic Libraries (NAAL), a department of the Alabama Commission on Higher Education in Montgomery. ADPNet currently has seven member institutions: the Alabama Department of Archives & History (ADAH), Auburn University, Spring Hill College, Troy University, the University of Alabama, the University of Alabama at Birmingham, and the University of North Alabama. The network hopes to recruit new member institutions this year.

## Why Alabama?

Alabama is an attractive candidate for a geographically distributed digital preservation network for several reasons. The first is the frequency of hurricanes, tornadoes, flooding, and other natural disasters, especially on and around Alabama's Gulf coast. In the past ten years, Alabama has been hit by at least four major hurricanes and many more tropical storms. In 2005, Hurricane Katrina devastated the coastal communities of Bayou la Batre and Coden and flooded downtown Mobile. The coastal communities are not the only parts of the state that have suffered from natural disasters, however. The interior of the state is vulnerable to tornadoes: in March 2007 a tornado swept through Enterprise, Alabama, destroying a high school and causing ten deaths. The second factor is Alabama's financial situation. Alabama is a relatively poor state, ranking 44th out of 50 in per capita real GDP in 2007. There isn't a lot of money to throw around, which means that technical solutions have to be simple, robust, and above all inexpensive to implement and maintain. Finally, Alabama is home to a rich and growing array of digital collections at libraries, archives, and museums. Many of these collections can be found in AlabamaMosaic, a statewide repository of digital materials on all aspects of Alabama's history, geography, and cultures. AlabamaMosaic currently contains over 20,000 digital objects from fifteen institutions around the state, and the

number continues to grow. This combination of circumstances — extreme weather, meager state financial resources, and rich digital collections — made Alabama an ideal test-case for a simple, inexpensive, but effective digital-preservation solution like LOCKSS.

Although ADPNet was originally inspired by and has some similarities with the MetaArchive Cooperative, there are important differences between the two initiatives. First and most importantly, the Alabama network is a single-state solution. This has simplified governance and allowed the network to be absorbed into an existing legal and administrative entity, one with bylaws and a committee structure already in place. Second, the Alabama network was designed to be a practical solution to a pressing statewide problem, not a research-and-development project. In order to attract participants, ADPNet had to be simple, robust, and above all inexpensive. This, and the fact that only one or two institutions in Alabama had had any prior experience with LOCKSS, meant that the members opted for the simplest, least expensive hardware and software solutions available, in the hope that these would be easier to deploy and manage and more attractive to other institutions in the state. (It should be pointed out, however, that although ADPNet's focus is not research and development, it has contributed at least one important addition to the LOCKSS toolbox: a generic LOCKSS plugin for harvesting CONTENTdm collections). Finally, unlike the MetaArchive, ADPNet is not a fee-based service organization. Rather, the preservation network is intended to complement AlabamaMosaic, another statewide initiative that has been kept going by in-kind contributions from its participating institutions. In other words, ADPNet was designed to run on relatively small expenditures and sweat equity, not on recurring infusions of grant money or annual membership fees. To some degree these differences reflect Alabama's expense-averse institutional culture. They also reflect a preference for self-sufficiency and informality where administrative arrangements are concerned.

## Why LOCKSS?

LOCKSS was originally designed to harvest and archive e-journals. The MetaArchive project demonstrated that LOCKSS technology could also be used to harvest, archive, and preserve locally-created digital content. Our experience with MetaArchive showed us that LOCKSS was simple, robust, and easy to maintain. It also ran on inexpensive hardware — a crucial consideration in Alabama. And we were impressed with the level of technical support provided by the LOCKSS staff. In a series of conversations in late 2005, Auburn and six other Alabama institutions agreed to pool resources to build a LOCKSS-based preservation network for the state if external start-up funding could be obtained. NAAL Director Sue Medina and I drafted and submit-

ted a funding proposal to the IMLS in January 2006. The proposal was funded in September 2006; it provided support for equipment and travel to the seven participating institutions through September 2008. Crucially, it also covered those institutions' annual membership fees in the LOCKSS Alliance for the same period. For their part, the institutions split the equipment costs with IMLS and contributed staff time and other in-house resources to the project.

## Accomplishments

At its inception, ADPNet identified four specific tasks. The first was to highlight the importance of preserving digital content among libraries, academic institutions, state agencies, and other cultural heritage institutions in Alabama. The second, to demonstrate the feasibility of state-based, low-cost models for digital preservation by creating a working example of such a network in Alabama. Third, to create an administrative structure to manage the network and assure its long-term sustainability. And fourth, to demonstrate that the network can support different types of digital content from different types of institutions, from public libraries and small colleges to large state agencies.

The network has achieved or is in the process of achieving all four tasks. On the technical side, ADPNet has been up and running since 2007. The network currently consists of seven LOCKSS nodes, each with a terabyte and a half of storage capacity. All seven member institutions have contributed content to the network, and almost 40 digital collections ("archival units," in LOCKSS-speak) have been harvested to date. They contain image, text, audio, and video files and include the 1867 Alabama voter registration volumes at ADAH; historical photographs from the Alabama Cooperative Extension Service collection at Auburn University; images of book bindings published in the *Confederate States of America* at the University of Alabama; and audio files of oral history interviews with civic leaders in Birmingham at the University of Alabama at Birmingham (for a partial list of harvested content, see "Collections" at <http://adpn.org/>). More digital content is on the way.

On the administrative side, the network drafted a governance policy that was adopted by NAAL at its annual business meeting in October 2008 (the policy is available on the ADPNet Website, under "Resources"). ADPNet has a lightweight governance structure consisting of two committees: the Steering Committee and the Technical Policy Committee. Every participating institution has the right to appoint a member to the Steering Committee, which in turn solicits nominations for the Technical Policy Committee. Together, the two committees are responsible for the day-to-day management of ADPNet.

*continued on page 47*

In keeping with the network's guiding principles, the requirements for membership are as simple and affordable as we could make them. Participating institutions must agree to install and run a **LOCKSS** server in the network; contribute content to the network; and join the **LOCKSS Alliance** for an annual fee. There is no **ADPNet** membership fee.

Surveys have shown that **ADPNet** has succeeded in raising awareness of the importance of digital preservation among Alabama libraries, archives, and state agencies. The task now is to translate this increased awareness into participation in **ADPNet**.

### Going Forward

**ADPNet's** main mission is to build and sustain a robust, inexpensive distributed digital preservation network for Alabama, but it also hopes to serve as a model for similar networks in other states and other countries who may think they can't afford to preserve their local digital heritage. **Private LOCKSS Networks** offer communities a low-cost, highly customizable alternative to more expensive digital preservation solutions. If **ADPNet** had a motto, it might be "keep it simple and keep it cheap." This basic approach appears to be working well for Alabama. It remains to be seen whether it will work for other states and consortia, but the signs so far are encouraging. 🌱

## You Gotta Go to School for This? — A Rave Review

by **Jared A. Seay** (Reference Librarian and Head, Media Collections, Addlestone Library, College of Charleston, Charleston, SC 29424) <seayj@cofc.edu>  
[www.cofc.edu/~seay](http://www.cofc.edu/~seay)

**O**h for the days when the word "rave" merely referred to speaking or writing enthusiastically (or incoherently) about something. Yet, linguistic evolution marches on. The term "rave" now describes a wild dance party. Certainly this is not news to those among you (you know who you are) who are connoisseurs of modern party culture. But, the thought of such an event occurring in a library (one with real books in it) sends the mind reeling, especially if one's mind tends toward reeling as mine does. A rave in an academic library? Who knew?

I was introduced to this rave library thing when our staff recently received an urgent all points bulletin of a student rave which was to occur at 11pm in the green just outside the library. Learning that this would involve hundreds, if not thousands, of (potentially riot prone) students who would actually move INTO the library for the main event, I realized I had served my public dutifully for the previous eight hours. I thus selflessly considered that my presence would be unnecessary.

So, though I was not physically present during the event, by piecing together eyewitness accounts, police reports, pieces of shredded textbooks, and the extensive **YouTube** record of the event, I was able to conclude that my decision to leave on time after work was a wise one. Even so, it was a literally earthshaking and exciting event for the library.

Apparently it is a common thing these days for students to blow off steam during final exams. Why one would find it necessary to "blow off steam" during this time is beyond me, since during my student days I found exam times conducive to quiet reflection and relaxation. Of course, besides the fact that I may not be indicative of the norm, my memory is fading a bit, and I sometimes lie outright for dramatic effect especially when it involves my youth. In any case, this library rave thing is a growing phenomenon on campuses in which students are given access to turn the university library into a rocking, screaming, pounding, music thumping rave arena. Students fill the open spaces within the library and dance and scream (with total abandon) to unseen sources of loud music for ten minutes without pause.

For those used to the library as a place of study and sedate informational exchange, it is surreal to see the main floor packed with a seething mass of screaming, dancing, exhilarated students. I think the sight of students crowd surfing across the rotunda amid a blizzard of shredded textbooks (being precipitated from the second and third floors) was particularly memorable and striking. For those ancient ones among us who have never experienced such a thing (or don't remember it) the rave has all the noise, chaos, and mayhem of a street riot, but with much less tear gas and much happier people. It is also over as quickly as it begins, which tends to calm the police/security who just stand to the side and grin a lot.

Across the country there are also so called "silent raves" in which everyone listens and dances and wiggles (silently of course) to the same music on their iPods. It's a sort of a synchronized chaotic, silent dance thing. This removes the noise level problem and accompanying police intervention while maintaining the shared communal audio and physical experience.

Is this made for a library or what? It's perfect. We are only one step away from the "silent-read-rave." I predict it will soon sweep across librarydom. A silent-read-rave is of course where all the participants in the library listen to the same audio book in individual headphones synchronously and react (dance, sway, thoughtfully absorb, mime literary criticism) simultaneously. The more adventurous of the participants may even be compelled to branch out into actual reading (we're talking real paper here). Now there's a library innovation for you. 🌱

continued on page 00