

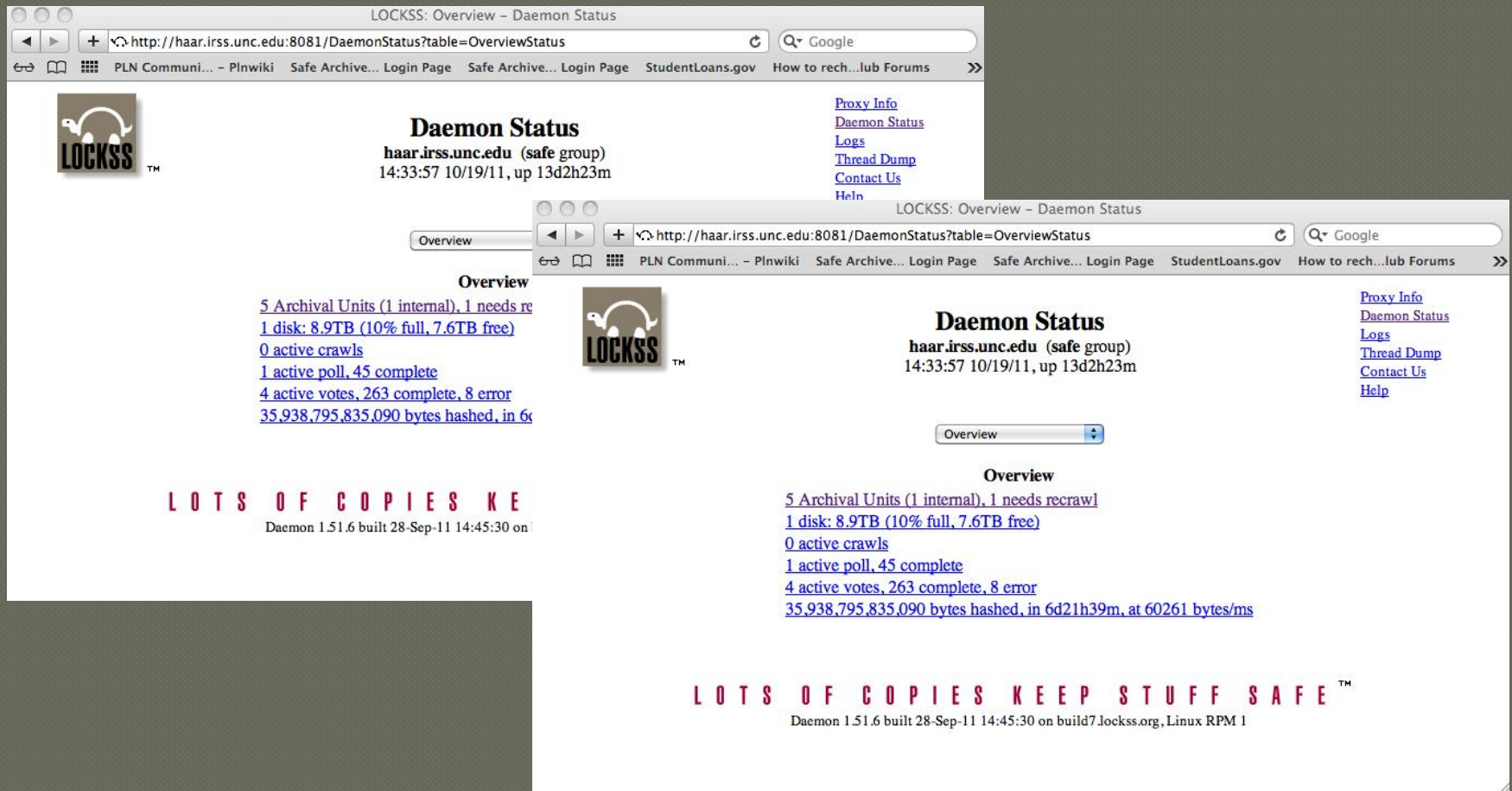


SafeArchive

Audit System Overview

Jonathan Crabtree
Odum Institute UNC Chapel Hill

Data preservation requires replication



The screenshot displays the LOCKSS Daemon Status web interface. The browser window shows the URL `http://haar.irss.unc.edu:8081/DaemonStatus?table=OverviewStatus`. The page features the LOCKSS logo and a navigation menu with links: [Proxy Info](#), [Daemon Status](#), [Logs](#), [Thread Dump](#), [Contact Us](#), and [Help](#). The main content area is titled "Daemon Status" and shows the status for `haar.irss.unc.edu (safe group)` as of 14:33:57 10/19/11. Below this, the "Overview" tab is selected, displaying the following statistics:

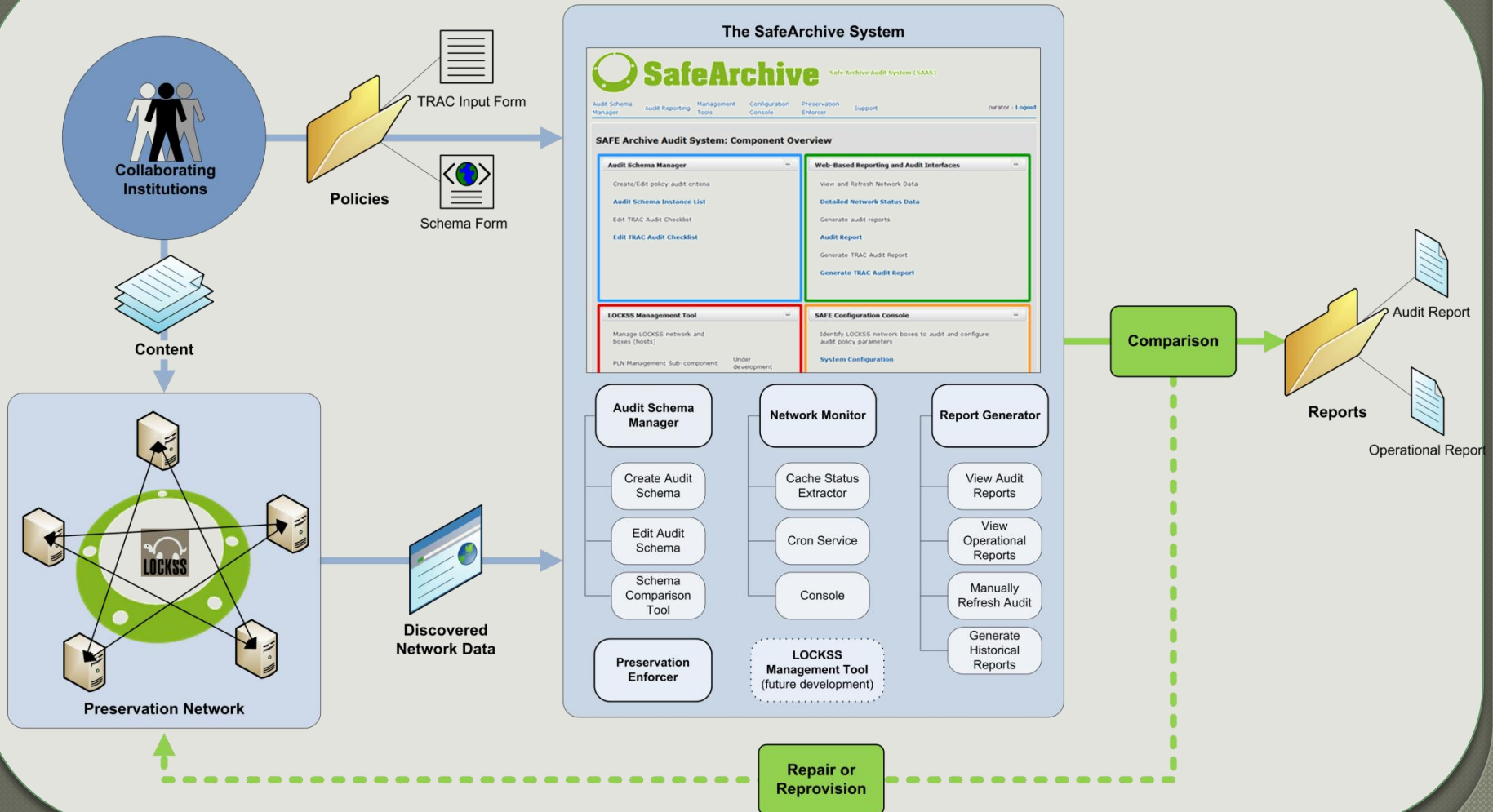
- 5 Archival Units (1 internal), 1 needs re
- 1 disk: 8.9TB (10% full, 7.6TB free)
- 0 active crawls
- 1 active poll, 45 complete
- 4 active votes, 263 complete, 8 error
- 35,938,795,835,090 bytes hashed, in 6

At the bottom of the interface, the text "LOTS OF COPIES KE" is displayed in large, red, spaced-out letters, followed by the version information: "Daemon 1.51.6 built 28-Sep-11 14:45:30 on".

Facilitating Collaborative Preservation

- *Collaborators* make resource commitments
- *Policy* records commitments, system properties, TRAC policies
- *Content is stored in* The Dataverse Network, other standard repository systems, on a web server, etc.
- *LOCKSS* provides replication, integrity, freshness, harvesting
- *SafeArchive software* provides monitoring, auditing, restoration, and provisioning

Centralize Network Auditing



SAFE PLN Statistics at a glance

| item | value |
|------------------------|----------|
| Group Name | safe |
| Institutions | 4 |
| Total Hosts | 7 |
| Total AUs | 5 |
| Total Repository Space | 23.26 TB |
| Used Space | 2.83 TB |
| Used Space | 12% |

As of 2011-10-19T00-00-23

SAFE PLN: Poll Agreement Results by AU

| AU | Latest-3 | Latest-2 | Latest-1 | Latest Poll |
|--------------------------|----------|----------|----------|-------------|
| ICPSR Dataverse | | | | |
| IQSS Dataverse | | | | |
| Index of /lockss/plugins | | | | |
| Odum Dataverse | | | | |

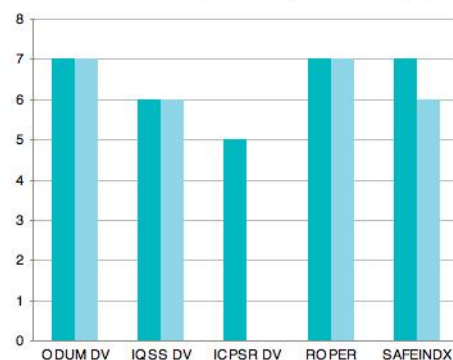
1 2

As of 2011-10-19T00-00-23

100% less than 100% Cancelled In Progress

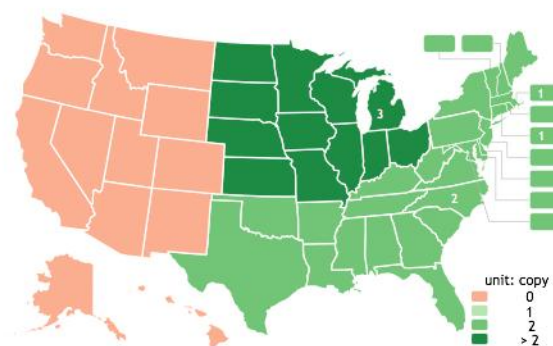
Experimental Showcase 1

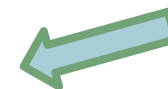
Archival Units: Available(dark blue) and Verified(light blue)



Experimental Showcase 2

Verified Odum-Dataverse Replicas by Region





SAFE PLN Statistics at a glance

| item | value |
|------------------------|----------|
| Group Name | safe |
| Institutions | 4 |
| Total Hosts | 7 |
| Total AUs | 5 |
| Total Repository Space | 21.56 TB |
| Used Space | 2.65 TB |
| Used Space | 12% |

As of 2011-10-18T00-00-33

SAFE PLN: Poll Agreement Results by AU

| AU | Latest-3 | Latest-2 | Latest-1 | Latest Poll |
|--------------------------|----------|----------|----------|-------------|
| ICPSR Dataverse | | | | |
| IQSS Dataverse | | | | |
| Index of /lockss/plugins | | | | |
| Odum Dataverse | | | | |

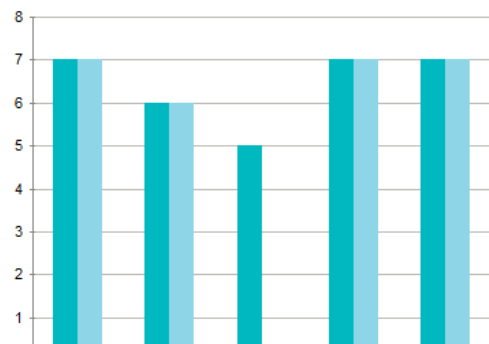
1
2

As of 2011-10-18T00-00-33

100%
 less than 100%
 Cancelled
 In Progress

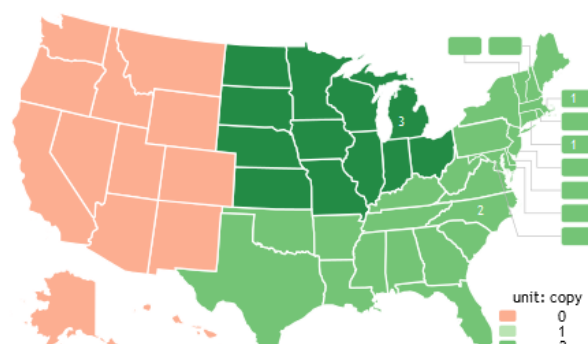
Experimental Showcase 1

Archival Units: Available(dark blue) and Verified(light blue)



Experimental Showcase 2

Verified Odum-Dataverse Replicas by Region



Detailed Network Status Data

Audit Report

Generate TRAC Audit Report

SAFE PLN

| item | value |
|------------------------|----------|
| Group Name | safe |
| Institutions | 4 |
| Total Hosts | 7 |
| Total AUs | 5 |
| Total Repository Space | 21.56 TB |
| Used Space | 2.65 TB |
| Used Space | 12% |

As of 2011-10-18T00-00-33

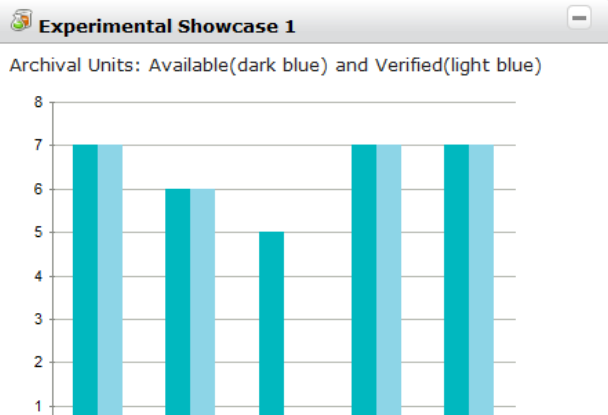
SAFE PLN: Poll Agreement Results by AU

| AU | Latest-3 | Latest-2 | Latest-1 | Latest Poll |
|--------------------------|----------|----------|----------|-------------|
| ICPSR Dataverse | | | | |
| IQSS Dataverse | | | | |
| Index of /lockss/plugins | | | | |
| Odum Dataverse | | | | |

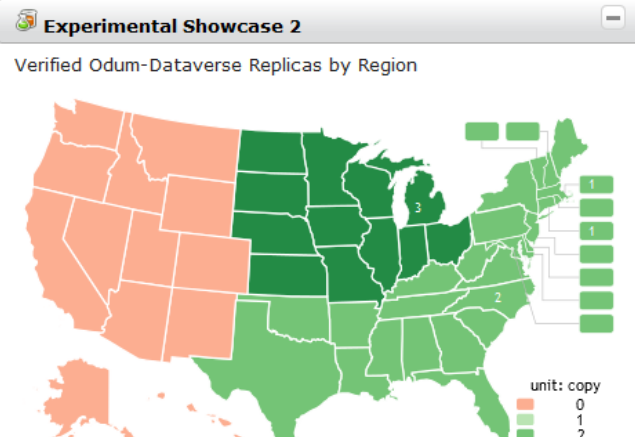
1 2

As of 2011-10-18T00-00-33

100% less than 100% Cancelled In Progress



<http://dev.irss.unc.edu:8080/safearchiveauditsystem-web/faces/index.xhtml#>





Network Data Tables

Network Data Tables

| Table Name | |
|----------------------------|----------------------------|
| lockss_box_table | Show Table |
| archival_unit_status_table | Show Table |
| repository_space_table | Show Table |
| au_overview_table | Show Table |
| crawl_status_table | Show Table |
| polls_table | Show Table |
| successful_polls_table | Show Table |

Data collection as of : 2011-10-18T00-00-33

Refresh Network-Status Data Tables

To browse the latest network-status data

- 1 Click the refresh button below.
- 2 Wait for the refresh-completion message.
- 3 **Log out and log in the SAAS again.**

[Refresh Tables](#)

Operation Reports

By Host

[View report](#)

By Archival-unit (AU)

[View report](#)



Archival-Unit Overview: List

| | PlnAuId | AuName | AuSizeMax | AuNReplica\$ | AuNVerifiedReplica\$ | LastSPollEnd | LastSCrawlEnd | CrawlDuration |
|---|---------|---------------------------------|-----------|--------------|----------------------|----------------------------|----------------------------|---------------|
| 1 | 2342 | Odum Dataverse | 62,747.0 | 7 | 6 | 2011-10-19 02:50:57 EDT | 2011-10-14 10:28:34 EDT | 0019:19:47 |
| 2 | 2341 | IQSS Dataverse | 95,744.0 | 6 | 5 | 2011-10-19 08:25:12 EDT | 2011-10-17 21:29:38 EDT | 0019:00:19 |
| 3 | 2339 | ICPSR Dataverse | 299,432.0 | 5 | 0 | | 2011-10-18 13:47:49 EDT | 0019:00:31 |
| 4 | 2343 | Roper's Data-PASS subset | 1,725.0 | 7 | 6 | 2011-10-18 20:02:46 EDT | 2011-10-14 15:04:18 EDT | 0019:00:04 |
| 5 | 2340 | Index of /lockss /plugins | 0.0 | 7 | 5 | 2011-10-18 20:51:16 EDT | 2011-10-18 23:55:41 EDT | 0019:00:00 |

< << 1 >> >

[Back to the network data viewer top page](#)

The SAFE Archive Audit System (SAAS) is a JEE6-based web application powered by



PRIMEFACES



Archival-Unit Overview: List

| | PlnAuId | AuName | AuSizeMax | AuNReplicas | AuNVerifiedReplicas | LastSPollEnd | LastSCrawlEnd | CrawlDuration |
|--|---------|---------------------------|-----------|-------------|---------------------|-------------------------|-------------------------|---------------|
| 1 | 2342 | Odum Dataverse | 62,747.0 | 7 | 6 | 2011-10-19 02:50:57 EDT | 2011-10-14 10:28:34 EDT | 0019:19:47 |
| <div> <div>PlnAuId:</div> <div>2342</div> </div> <div> <div>AuId:</div> <div>edu harvard iq dvn lockss plugin DVNOAIPlugin&base_url~http%3A%2F%2Farc%2Eirss%2Eunc%2Eedu%2Fbase_url2~http%3A%2F%2Farc%2Eirss%2Eunc%2Eedu%2Fblankmanifest%2Ehtml%3F%3A%2F%2Farc%2Eirss%2Eunc%2Eedu%2Fdvn%2FOAIHandler&oai_spec~Odum</div> </div> <div> <div>AuName:</div> <div>Odum Dataverse</div> </div> <div> <div>AuSizeMax(MB):</div> <div>65795159691</div> </div> <div> <div>DiskUsageMax(MB):</div> <div>6.72198033408E10</div> </div> <div> <div>AuNReplicas:</div> <div>7</div> </div> <div> <div>AuNVerifiedReplicas:</div> <div>6</div> </div> <div> <div>LastSPollEnd:</div> <div>2011-10-19 02:50:57 EDT</div> </div> <div> <div>LastSCrawlEnd:</div> <div>2011-10-14 10:28:34 EDT</div> </div> <div> <div>CrawlDuration:</div> <div>0019:19:47</div> </div> <div> <div>PollId:</div> <div>o10ZL2w2iUzelRgnBBYRPug6Mig=</div> </div> | | | | | | | | |
| 2 | 2341 | IQSS Dataverse | 95,744.0 | 6 | 5 | 2011-10-19 08:25:12 EDT | 2011-10-17 21:29:38 EDT | 0019:00:19 |
| 3 | 2339 | ICPSR Dataverse | 299,432.0 | 5 | 0 | | 2011-10-18 13:47:49 EDT | 0019:00:31 |
| 4 | 2343 | Roper's Data-PASS subset | 1,725.0 | 7 | 6 | 2011-10-18 20:02:46 EDT | 2011-10-14 15:04:18 EDT | 0019:00:04 |
| 5 | 2340 | Index of /lockss /plugins | 0.0 | 7 | 5 | 2011-10-18 20:51:16 EDT | 2011-10-18 23:55:41 EDT | 0019:00:00 |

Defining Policy

- ◉ Storage commitment
- ◉ Space allocation requirement
- ◉ Geographic replication
- ◉ Replication
- ◉ Verification frequency
- ◉ Data update frequency
- ◉ TRAC audit policies

List Audit Schema Instances

Current Audit Schema Instance

| Schema Instance Name | Last Modified Date | |
|----------------------------------|-------------------------|--|
| lockss_audit_schema_instance.xml | 2011-10-17 10:08:27 EDT | view Edit Download |

[Create a New Audit Schema Instance](#)

[Back to the Home Page](#)

Audit Schema Instances under Version Control

| Revision Number | Committed Date | Saved By | Log Message | |
|-----------------|-------------------------|----------|---|--------------------------|
| 146 | 2011-10-17 10:08:27 EDT | curator | Audit Schema Instance:Id=1.0.0; Version=asiId_89xL; date=2011-10-17T10-08-27; saved by curator; | Download |
| 145 | 2011-10-17 10:01:25 EDT | curator | Audit Schema Instance:Id=1.0.0; Version=asiId_f4YU; date=2011-10-17T10-01-24; saved by curator; | Download |
| 144 | 2011-10-17 09:58:58 EDT | curator | Audit Schema Instance:Id=1.0.0; Version=asiId_gX8l; date=2011-10-17T09-58-58; saved by curator; | Download |
| 143 | 2011-10-17 09:53:48 EDT | curator | Audit Schema Instance:Id=1.0.0; Version=asiId_3Fy5; date=2011-10-17T09-53-48; saved by curator; | Download |
| 142 | 2011-10-17 09:52:42 EDT | curator | Audit Schema Instance:Id=1.0.0; Version=asiId_jn7r; date=2011-10-17T09-52-42; saved by curator; | Download |

(1 of 9)

[1](#)
[2](#)
[3](#)
[4](#)
[5](#)
[6](#)
[7](#)
[8](#)
[9](#)
[5](#)

Edit Audit Schema Instance

About the PLN

Question 1: What is the name of the private LOCKSS network (PLN) that you would like the SAFE system to audit?

Question 2: What is the email address of the PLN administrator?

About PLN member Hosts

Question 3: How much storage disk space do you wish to commit on each server for the SAFE system to use (in gigabytes)?

| Discovered Hosts | | | | |
|--|--------------------------------|-----------------|-----------------------------------|---------------------|
| <input type="checkbox"/> | host Name | host Ip Address | Space Committed (Gb) | Geographic Location |
| <input checked="" type="checkbox"/> | lockss-3.ropercenter.uconn.edu | 137.99.36.160 | <input type="text" value="1740"/> | Connecticut |
| <input checked="" type="checkbox"/> | lockss.hmdc.harvard.edu | 140.247.115.220 | <input type="text" value="39"/> | Massachusetts |
| <input checked="" type="checkbox"/> | lockss-0.icpsr.umich.edu | 141.211.146.29 | <input type="text" value="120"/> | Michigan |
| <input checked="" type="checkbox"/> | lockss-1.icpsr.umich.edu | 141.211.146.52 | <input type="text" value="1843"/> | Michigan |
| <input checked="" type="checkbox"/> | lockss-2.icpsr.umich.edu | 141.211.146.53 | <input type="text" value="1843"/> | Michigan |
| <div>1 2 > <</div> | | | | |

The above table shows discovered servers only. To add a missing server or update a host's attributes, click the following link:

[Edit the Host List](#)

About Archival Units

Note: Before preservation content can be audited, it must be placed on a Host as an Archival unit or AU. If your content is not listed below in an AU, click the following link to add it to a LOCKSS host:

[Edit the Host List](#)

Question 4: The following AUs have been found on the specified hosts. Please check the archival units you would like to audit and answer the following questions in respective columns of the table below:

The above table shows discovered servers only. To add a missing server or update a host's attributes, click the following link:

[Edit the Host List](#)

About Archival Units

Note: Before preservation content can be audited, it must be placed on a Host as an Archival unit or AU. If your content is not listed below in an AU, click the following link to add it to a LOCKSS host:

[Edit the Host List](#)

Question 4: The following AUs have been found on the specified hosts. Please check the archival units you would like to audit and answer the following questions in respective columns of the table below:

4.1 How frequently should this archival unit AU be re-crawled (in days)? Re-crawling adds new content to the AUs as changes occur keeping your preservation copies current this is the Update Frequency.

4.2 What is the maximum size that you expect this archival unit AU to be (in gigabytes)? The default value is the current size. Please add your growth expectations. This is your storage required.

Question 5: How many preservation copies of each collection do you want to maintain? This will be the Number of Replicates.

Question 5a: How many regions do you want to maintain a replica of a collection? This will be the value of Geographic Redundancy.

Question 6: Archival Units AUs will be tested on a regular basis for problems to ensure your content is not corrupt. What is the maximum time you would allow between these verifications, in days? This is your Verification Frequency.

Question 7: What is the maximum amount of time that the preservation system should take to ingest/crawl your content (in days)? LOCKSS uses crawling to add your content to an AU. The maximum number of days this should take is your Update Duration.

| Discovered Archival-Units | | | | | | | | | |
|---|---------------------------|--------------------------------|----------------------------------|--------------------------------|--------------------------------|---------------------------------|---------------------------------|-----------------|--------------------|
| <input type="checkbox"/> | Archival Unit Name | Update Frequency | Storage Required (Gb) | Number of Replicates | Geographic Redundancy | Verification Frequency | Update Duration | Subject | Owner Institution |
| <input checked="" type="checkbox"/> | ICPSR Dataverse | <input type="text" value="7"/> | <input type="text" value="500"/> | <input type="text" value="3"/> | <input type="text" value="2"/> | <input type="text" value="21"/> | <input type="text" value="21"/> | Social Sciences | UMich-ICPSR |
| <input checked="" type="checkbox"/> | Index of /lockss /plugins | <input type="text" value="7"/> | <input type="text" value="1"/> | <input type="text" value="3"/> | <input type="text" value="2"/> | <input type="text" value="21"/> | <input type="text" value="21"/> | Social Sciences | HU-IQSS |
| <input checked="" type="checkbox"/> | IQSS Dataverse | <input type="text" value="7"/> | <input type="text" value="500"/> | <input type="text" value="3"/> | <input type="text" value="2"/> | <input type="text" value="21"/> | <input type="text" value="21"/> | Social Sciences | HU-IQSS |
| <input checked="" type="checkbox"/> | Odum Dataverse | <input type="text" value="7"/> | <input type="text" value="300"/> | <input type="text" value="3"/> | <input type="text" value="2"/> | <input type="text" value="21"/> | <input type="text" value="21"/> | Social Sciences | UNC-Odum Institute |
| <input checked="" type="checkbox"/> | Roper's Data-PASS subset | <input type="text" value="7"/> | <input type="text" value="10"/> | <input type="text" value="3"/> | <input type="text" value="2"/> | <input type="text" value="21"/> | <input type="text" value="21"/> | Social Sciences | UConn-Roper |
| | | Q4.1 | Q4.2 | Q5 | Q5a | Q6 | Q7 | | |
| <div><=< < 1 > >=></div> | | | | | | | | | |

To update or correct an AU's subject or ownership information, click the following link:

[Edit the Archival-Unit List](#)

[Audit Schema Instance List](#)

[Edit TRAC Audit Checklist](#)

at a glance

| item | value |
|---------------------------|----------|
| Group Name | safe |
| Institutions | 4 |
| Total Hosts | 7 |
| Total AUs | 5 |
| Total Repository Space | 23.26 TB |
| Used Space | 2.83 TB |
| Used Space | 12% |
| As of 2011-10-19T00-00-23 | |

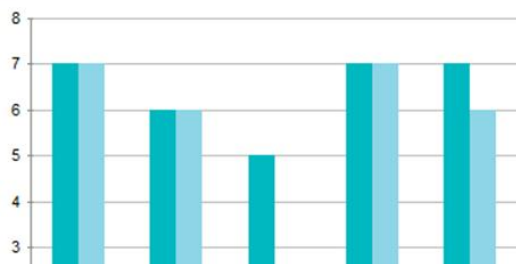
SAFE PLN: Poll Agreement Results by AU

| AU | Latest-3 | Latest-2 | Latest-1 | Latest Poll |
|--|----------|----------|----------|-------------|
| ICPSR Dataverse | | | | |
| IQSS Dataverse | | | | |
| Index of /lockss/plugins | | | | |
| Odum Dataverse | | | | |
| <div> << < 1 2 > >> </div> | | | | |
| As of 2011-10-19T00-00-23 | | | | |

100%
 less than 100%
 Cancelled
 In Progress

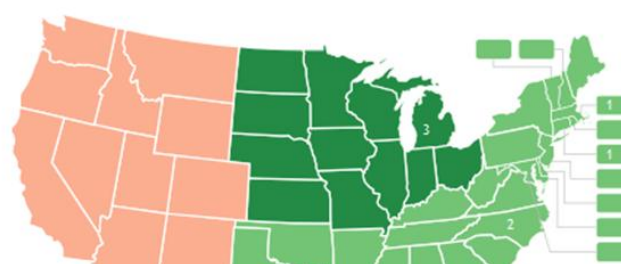
Experimental Showcase 1

Archival Units: Available(dark blue) and Verified(light blue)



Experimental Showcase 2

Verified Odum-Dataverse Replicas by Region



Edit TRAC Audit Checklist

Section A

Section B

Section C

Confirmation and Save

Organizational Infrastructure

Section A

| Id | Criterion | Evidence (Documents) Examined |
|------|---|--|
| A1.1 | Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information. | http://www.icpsr.umich.edu/icpsrweb/DATAPASS/about.jsp#overview , Data-PASS Articles of Collaboration: http://www.icpsr.umich.edu/files/DATAPASS/pdf/collaboration.pdf |
| A1.2 | Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope. | Data-PASS Articles of Collaboration |
| A2.1 | Repository has identified and established the duties that it needs to perform and has appointed staff with adequate skills and experience to fulfill these duties. | Data-PASS Articles of Collaboration, <u>Odum</u> Staff Page and Bios |
| A2.2 | Repository has the appropriate number of staff to support all functions and services. | Data-PASS Publications and Presentations: http://www.icpsr.umich.edu/icpsrweb/DATAPASS/presentations.jsp , partner's Staff Pages |
| | Repository has an active professional | Data-PASS Publications and Presentations, <u>Odum</u> Internal Courses |

Generate Audit Reports

- PLN Replication status report
- TRAC audit information
- Individual institutional reports
- Manual & automatic reports
- Audit reports stored and versioned on regular basis

[Detailed Network Status Data](#)

[Audit Report](#)

[Generate TRAC Audit Report](#)

SAFE P

| item | value |
|---------------------------|----------|
| Group Name | safe |
| Institutions | 4 |
| Total Hosts | 7 |
| Total AUs | 5 |
| Total Repository Space | 21.56 TB |
| Used Space | 2.65 TB |
| Used Space | 12% |
| As of 2011-10-18T00-00-33 | |

SAFE PLN

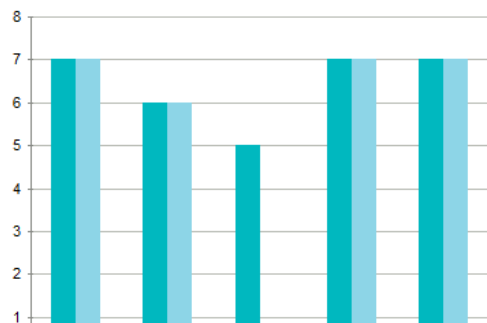
Poll Agreement Results by AU

| AU | Latest-3 | Latest-2 | Latest-1 | Latest Poll |
|--|----------|----------|----------|-------------|
| ICPSR Dataverse | | | | |
| IQSS Dataverse | | | | |
| Index of /lockss/plugins | | | | |
| Odum Dataverse | | | | |
| <div> 1 2 </div> | | | | |
| As of 2011-10-18T00-00-33 | | | | |

100%
 less than 100%
 Cancelled
 In Progress

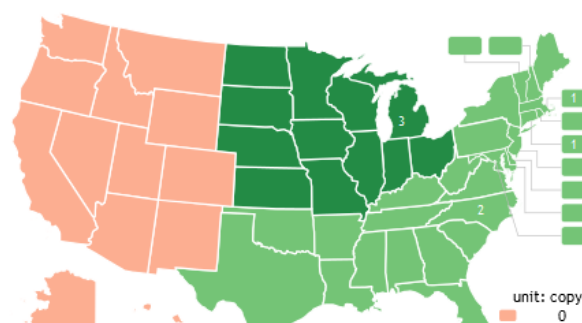
Experimental Showcase 1

Archival Units: Available(dark blue) and Verified(light blue)



Experimental Showcase 2

Verified Odum-Dataverse Replicas by Region





Safe Archive Audit System (SAAS)

[Audit Schema
Manager](#)

[Audit Reporting](#)

[Management
Tools](#)

[Configuration
Console](#)

[Preservation
Enforcer](#)

[Support](#)

curator - [Logout](#)

TRAC Audit Report

Generate TRAC Audit Report

To generate a new TRAC audit-report, click the following button

[Generate a new TRAC Audit Report](#)

Current TRAC Audit Report: Summary and Full Report by BIRT Viewer

To view the latest TRAC audit report, click one of the following links:

[TRAC Audit Report: Summary](#)

[Full TRAC audit report by BIRT viewer](#)

The SAFE Archive Audit System (SAAS) is a JEE6-based web application powered by



PRIMEFACES



Trusted Repositories Audit & Certification (TRAC) Report

Oct 19, 2011 4:20 PM

TRAC Audit Summary

PLN Group Name: safe
TRAC version: 1.0
TRAC Report data generation date: Tue Oct 18 13:26:18 EDT 2011

Complete

[Section A: 24 out of 24 entered](#)

[Section B: 44 out of 44 entered](#)

[Section C: 16 out of 16 entered](#)

TRAC Audit Checklist

Section A

Incomplete

- A1.1 Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information.
Data-PASS About Page: <http://www.icpsr.umich.edu/icpsrweb/DATAPASS/about.jsp#overview>,
Data-PASS Articles of Collaboration: <http://www.icpsr.umich.edu/files/DATAPASS/pdf/collaboration.pdf>
- A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.
Data-PASS Articles of Collaboration
- A2.1 Repository has identified and established the duties that it needs to perform and has appointed staff with adequate skills and experience to fulfill these duties.
Data-PASS Articles of Collaboration, Odum Staff Page and Bios
- A2.2 Repository has the appropriate number of staff to support all functions and services.
Data-PASS Publications and Presentations: <http://www.icpsr.umich.edu/icpsrweb/DATAPASS/presentations.jsp>, **partner's Staff Pages**
- A2.3 Repository has an active professional development program in place that provides staff with skills and expertise development opportunities.
Data-PASS Publications and Presentations, Odum Internal Courses
- A3.1 Repository has defined its designated community(ies) and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation service requirements will be met.
Data-PASS Articles of Collaboration, Data-PASS Confidentiality (pg. 6-7): <http://www.icpsr.umich.edu/files/DATAPASS/pdf/confidentiality.pdf>, **Odum Mission Statement**

[Detailed Network Status Data](#)
[Audit Report](#)
[Generate TRAC Audit Report](#)

SAFE PLN

| item | value |
|----------------------------------|----------|
| Group Name | safe |
| Institutions | 4 |
| Total Hosts | 7 |
| Total AUs | 5 |
| Total Repository Space | 23.26 TB |
| Used Space | 2.83 TB |
| Used Space | 12% |
| As of 2011-10-19T00-00-23 | |

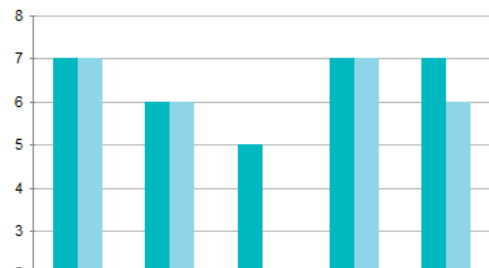
SAFE PLN: Poll Agreement Results by AU

| AU | Latest-3 | Latest-2 | Latest-1 | Latest Poll |
|----------------------------------|----------|----------|----------|-------------|
| ICPSR Dataverse | | | | |
| IQSS Dataverse | | | | |
| Index of /lockss/plugins | | | | |
| Odum Dataverse | | | | |
| 1 2 | | | | |
| As of 2011-10-19T00-00-23 | | | | |

100% less than 100% Cancelled In Progress

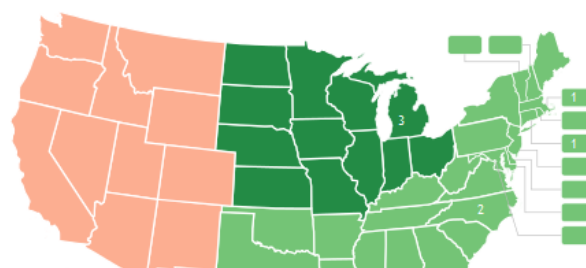
Experimental Showcase 1

Archival Units: Available(dark blue) and Verified(light blue)



Experimental Showcase 2

Verified Odum-Dataverse Replicas by Region





Audit Report Listing

Update Audit Report Data

Network-daemon-status data were last updated at:

2011-10-19T00-00-23

Audit-report data were last updated at:

2011-10-18T20-09-52

The state of required data tables:

Network data are ready for audit-reporting

To generate a new audit-report, click the following button to update audit-report data

[Update Audit Report Data](#)

Current Audit Report

To view the latest audit report, click one of the following links:

[view the current report without filters](#)

[view the current report with filters](#)



Archived Audit Reports

| Archived Audit Reports | | | |
|------------------------|-------------------------|---|--------------------------|
| Revision Number | Committed Date | Log Message | |
| 137 | 2011-10-16 01:00:02 EDT | audit report xml: generation date=2011-10-16T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |
| 132 | 2011-10-09 01:00:06 EDT | audit report xml: generation date=2011-10-09T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |
| 129 | 2011-10-02 01:00:03 EDT | audit report xml: generation date=2011-10-02T01-00-00; audit schema instance: version=; saved by | download |

BIRT Report Viewer



Showing page 1 of

Go to page:

Parameter

Parameters marked with * are required.

Select owner institution(s) for au: *

| |
|--------------------|
| All |
| HU-IQSS |
| UConn-Roper |
| UMich-ICPSR |
| UNC-Odum Institute |

OK Cancel

**Audit Summary Report**

Report Generation Date: Oct 19, 2011 10:32 AM

Network ID: safe

Reporting Host: dev.irss.unc.edu

Audit Data Date: Tue Oct 18 2011 20:09:52 GMT-0400 (EDT)

Preservation Network Summary

| | |
|---|---------------|
| Mean Up Time for LOCKSS Hosts | 11d:2h:26m:1s |
| Number of Hosts in the PLN | 7 |
| Number of Hosts NOT Reporting | 0 |
| Number of Unique AUs in the PLN | 5 |
| Total Number of AU Replicates | 32 |
| Number of AUs NOT Policy Compliant | 2 |
| Total Disk Space in the PLN | 21.555 T |
| Total Disk Space in Use | 2.651 T |
| Total Disk Space Free | 18.904 T |
| Number of Hosts NOT Meeting Storage Commitments | 1 |

Archival Unit (AU) Summary

[HU-IQSS]

[Click here](#) to view Archival Unit Detail Report.**AU Disk Size (estimated)²**

| AU Name | Policy | Actual | Audit Result | Notes |
|--------------------------|-----------|----------|--------------|-------|
| IQSS Dataverse | 500.000 G | 93.500 G | Pass | - |
| Index of /lockss/plugins | 1.000 G | 38.055 K | Pass | - |

Number of Verified AU Replicates²

| AU Name | Policy | Actual | Audit Result | Notes |
|--------------------------|--------|--------|--------------|-------|
| IQSS Dataverse | 3 | 5 | Pass | - |
| Index of /lockss/plugins | 3 | 6 | Pass | - |

BIRT Report Viewer



Showing page 1 of

Go to page:

Parameter

Parameters marked with * are required.

Select AU(s) to view detail report: *

- All
- ICPSR Dataverse
- IQSS Dataverse
- Index of /lockss/plugins
- Odum Dataverse
- Roper's Data-PASS subset

OK Cancel



Archival Unit Detail Report

Report Generation Date: Oct 19, 2011 10:33 AM

Network ID: safe

Reporting Host: dev.irss.unc.edu

Report Data Date: Tue Oct 18 23:56:14 EDT 2011

Archival Unit (AU) Overview

[Click here](#) to view Audit Summary Report.

| AU Name | Max AU Replicate Disk Size | Max Allowable Disk Usage | Number of Replicates | Number of Verified Replicates | Last Successful Poll End Date | Last Successful Crawl End Date |
|--------------------------|----------------------------|--------------------------|----------------------|-------------------------------|-------------------------------|--------------------------------|
| Odum Dataverse | 61.277 G | 62.603 G | 7 | 6 | Oct 19, 2011 2:50 AM | Oct 14, 2011 10:28 AM |
| IQSS Dataverse | 93.500 G | 112.586 G | 6 | 5 | Oct 19, 2011 8:25 AM | Oct 17, 2011 9:29 PM |
| ICPSR Dataverse | 292.414 G | 323.422 G | 5 | 0 | - | Oct 18, 2011 1:47 PM |
| Roper's Data-PASS subset | 1.685 G | 1.802 G | 7 | 6 | Oct 18, 2011 8:02 PM | Oct 14, 2011 3:04 PM |
| Index of /lockss/plugins | 38.055 K | 204.800 K | 7 | 5 | Oct 18, 2011 8:51 PM | Oct 18, 2011 11:55 PM |

AU Replicate Locations and Details

[Click here](#) to view Host Detail Report.

| Host Name | AU Replicate Disk Size | Allowable Disk Usage | Number of Successful Polls | Status |
|---------------------------------|------------------------|----------------------|----------------------------|-------------------|
| ICPSR Dataverse | | | | |
| lockss-3.ropcenter.uconn.edu | - | - | 2 | Crawling |
| lockss-1.icpsr.umich.edu | 291.332 G | 303.990 G | 1 | 99.92% Agreement |
| lockss-2.icpsr.umich.edu | 292.299 G | 304.196 G | 1 | 99.92% Agreement |
| fong.irss.unc.edu | 292.414 G | 323.422 G | 1 | 98.37% Agreement |
| haar.irss.unc.edu | 292.222 G | 304.281 G | 0 | 99.91% Agreement |
| Index of /lockss/plugins | | | | |
| lockss-3.ropcenter.uconn.edu | 38.055 K | 102.400 K | 13 | 100.00% Agreement |
| lockss.hmdc.harvard.edu | 38.055 K | 204.800 K | 8 | 100.00% Agreement |
| lockss-0.icpsr.umich.edu | 38.055 K | 102.400 K | 13 | 100.00% Agreement |
| lockss-1.icpsr.umich.edu | 38.055 K | 102.400 K | 13 | 100.00% Agreement |
| lockss-2.icpsr.umich.edu | 38.055 K | 102.400 K | 13 | 100.00% Agreement |
| fong.irss.unc.edu | 38.055 K | 204.800 K | 10 | 100.00% Agreement |
| haar.irss.unc.edu | 38.055 K | 102.400 K | 13 | 100.00% Agreement |
| IQSS Dataverse | | | | |



Audit Report Listing

Update Audit Report Data

Network-daemon-status data were last updated at:

2011-10-19T00-00-23

Audit-report data were last updated at:

2011-10-18T20-09-52

The state of required data tables:

Network data are ready for audit-reporting

To generate a new audit-report, click the following button to update audit-report data

[Update Audit Report Data](#)

Current Audit Report

To view the latest audit report, click one of the following links:

[view the current report without filters](#)

[view the current report with filters](#)

Archived Audit Reports

| Archived Audit Reports | | | |
|------------------------|-------------------------|---|--------------------------|
| Revision Number | Committed Date | Log Message | |
| 137 | 2011-10-16 01:00:02 EDT | audit report xml: generation date=2011-10-16T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |
| 132 | 2011-10-09 01:00:06 EDT | audit report xml: generation date=2011-10-09T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |
| 129 | 2011-10-02 01:00:03 EDT | audit report xml: generation date=2011-10-02T01-00-00; audit schema instance: version=; saved by | download |

Audit Report Listing**Update Audit Report Data**

Network-daemon-status data were last updated at:

2011-10-19T00-00-23

Audit-report data were last updated at:

2011-10-18T20-09-52

The state of required data tables:

Network data are ready for audit-reporting

To generate a new audit-report, click the following button to update audit-reporting

[Update Audit Report Data](#)**Current Audit Report**


To view the latest audit report, click one of the following links:

[view the current report without filters](#)[view the current report with filters](#)**Archived Audit Reports**

| Archived Audit Reports | | | |
|------------------------|-------------------------|---|--------------------------|
| Revision Number | Committed Date | | |
| 137 | 2011-10-16 01:00:02 EDT | | |
| 132 | 2011-10-09 01:00:06 EDT | audit report xml: generation date=2011-10-09T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |
| 129 | 2011-10-02 01:00:03 EDT | audit report xml: generation date=2011-10-02T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |
| 126 | 2011-09-25 01:00:02 EDT | audit report xml: generation date=2011-09-25T01-00-00; audit schema instance: version=; saved by auditreportCJ; | download |

Opening lockss_audit_report_revision_137.xml

You have chosen to open

 lockss_audit_report_revision_137.xml

which is a: XML Document

from: http://dev.irss.unc.edu:8080

What should Firefox do with this file?

☒ Open with

XML Editor (default)

☐ Save File☐ Do this automatically for files like this from now on.

OK

Cancel

```

<?xml version="1.0" encoding="UTF-8" ?>
- <compare xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="lockss_audit_report.xsd">
- <whatsthis>
  <filename>/usr/safe/data/audit_output_files/lockss_audit_report.xml</filename>
  <date>2011-10-16 01:00:00</date>
  <audit_id>1.0.0</audit_id>
</whatsthis>
- <hosts>
- <host host="141.211.146.29" seq_num="1">
  - <test>
    <schema_xpath>/SSP/hosts/host/@host_id</schema_xpath>
    <schema_value>141.211.146.29</schema_value>
    <db_column>ip_address</db_column>
    <db_value>141.211.146.29</db_value>
    <match>true</match>
    <note />
  </test>
  - <test>
    <schema_xpath>/SSP/hosts/host/hostCapabilities/storageAvailable/@max_size</schema_xpath>
    <schema_value>128849018880</schema_value>
    <db_column>repo_size</db_column>
    <db_value>128849018880</db_value>
    <match>true</match>
    <note />
  </test>
  - <test>
    <schema_xpath>/SSP/hosts/host/hostIdentity/name</schema_xpath>
    <schema_value>lockss-0.icpsr.umich.edu</schema_value>
    <db_column>host_name</db_column>
    <db_value>lockss-0.icpsr.umich.edu</db_value>
    <match>true</match>
    <note />
  </test>
  - <test>
    <schema_xpath>/SSP/network/@network_id</schema_xpath>
    <schema_value>safe</schema_value>
    <db_column>group_name</db_column>
    <db_value>safe</db_value>
    <match>true</match>
    <note />
  </test>
</host>
- <host host="141.211.146.52" seq_num="2">
  - <test>
    <schema_xpath>/SSP/hosts/host/@host_id</schema_xpath>
    <schema_value>141.211.146.52</schema_value>
    <db_column>ip_address</db_column>
    <db_value>141.211.146.52</db_value>
    <match>true</match>
    <note />
  </test>
  - <test>
    <schema_xpath>/SSP/hosts/host/hostCapabilities/storageAvailable/@max_size</schema_xpath>
    <schema_value>128849018880</schema_value>
    <db_column>repo_size</db_column>
    <db_value>128849018880</db_value>
    <match>true</match>
    <note />
  </test>
  - <test>
    <schema_xpath>/SSP/hosts/host/hostIdentity/name</schema_xpath>
    <schema_value>lockss-0.icpsr.umich.edu</schema_value>
    <db_column>host_name</db_column>
    <db_value>lockss-0.icpsr.umich.edu</db_value>
    <match>true</match>
    <note />
  </test>
  - <test>
    <schema_xpath>/SSP/network/@network_id</schema_xpath>
    <schema_value>safe</schema_value>
    <db_column>group_name</db_column>
    <db_value>safe</db_value>
    <match>true</match>
    <note />
  </test>
</host>
</hosts>
</compare>

```

SafeArchive Adds Policy Management to Replicated Storage

- LOCKSS Lots of Copies Keep Stuff Safe

- Widely used in library community
- Self-contained OSS replication system, low maintenance, inexpensive
- Harvests resources via web-crawling (mostly)...
- Maintains copies through secure p2p protocol
- Zero trust & self repairing



- What does SafeArchive Add?

- Auditing –
easily monitor number of copies of content in network
- Collaboration –
coordinate across partners, monitor resource commitments
- Provisioning –
ensure sufficient copies and distribution
- Restoration guarantees
- Integrates with The Dataverse Network ®
(and other repositories through OAI-PMH + DDI)



Future Possibilities

- Support additional audit standards
 - Data Seal of Approval
 - ISO 16363
- Simplify SAFE audit standards interface
 - Single interview style questionnaire
 - Automatic crosswalk to all supported audit standards
- Enhance qualitative audit compliance
 - Interactive audit agency tools
 - Multiple levels of audit compliance
- Support additional replication networks
 - iRODS

Collaborators*

- Micah Altman, Leonid Andreev, Ed Bachman, Adam Buchbinder, Ken Bollen, Bryan Beecher, Steve Burling, Kevin Condon, Jonathan Crabtree, Merce Crosas, Gary King, Patrick King, Tom Lipkis, Freeman Lo, Jared Lyle, Marc Maynard, Nancy McGovern, Lois Timms-Ferrarra, Akio Sone, Bob Treacy



- Research Support

Thanks to the, IMLS (LG-05-09-0041-09), Library of Congress (PA#NDP03-1), the National Science Foundation (DMS-0835500, SES 0112072), the Harvard University Library, the Institute for Quantitative Social Science, the Harvard-MIT Data Center, and the Murray Research Archive.

* And co-conspirators



Related Work

- Reprints available from:
<http://futurelib.org>
- Altman, M., and J. Crabtree, 2011. "Using the SafeArchive System: TRAC-Based Auditing of LOCKSS", *Proceedings of Archiving 2011*.
- Thu-mai Christian, Jonathan Crabtree, Nancy McGovern et al., Overview of SafeArchive :An Open-Source System for Automatic Policy-Based Collaborative Archival Replication. *Proceedings of iPres 2011*. (Forthcoming)
- Altman, M., Beecher, B., and Crabtree, J.; with L. Andreev, E. Bachman, A. Buchbinder, S. Burling, P. King, M. Maynard. 2009. "A Prototype Platform for Policy-Based Archival Replication." *Against the Grain*. 21(2): 44-47.
- Altman, M., Adams, M., Crabtree, J., Donakowski, D., Maynard, M., Pienta, A., & Young, C. 2009. "Digital preservation through archival collaboration: The Data Preservation Alliance for the Social Sciences." *The American Archivist*. 72(1): 169-182

Questions

◉ Website

- www.safearchive.org

◉ Sourceforge

- <http://safearchive.sourceforge.net/>

◉ Contacts

- Michah.Altman@gmail.com
- [Jonathan Crabtree@unc.edu](mailto:Jonathan_Crabtree@unc.edu)