April 2009

People Profile: Michael Altman

Editor

these hooks are designed to interface with the Dataverse libraries in order to navigate the click-through style authentication used in the DVN.

• The plugin can process metadata schemas other than OAI-DC, when an appropriate library is supplied. We have supplied a library for DDI version 2.0 metadata. This preserves the metadata, and harvests any URI's listed as resources in it. We have also included performance enhancements for handling portions of large metadata objects.

Discussion

The ability to express replication requirements and inter-archival replication commitments using a formal schema, and to automatically audit a **LOCKSS** network for dynamic consistency with these requirements is a significant advance. This provides an organization with automatic, continuous and compelling evidence that accurate, timely, and complete replicas are being maintained. Moreover this approach does not require a central administrator or homogenous configuration of the **LOCKSS** network, or create a single point of failure, either in terms of individual machines or entire institutions.

This work is in a prototype stage, and two questions remain before it could be used in production: First, the **LOCKSS** cache manager, which plays a much more prominent role in a **PLN** than it does it in the public **LOCKSS** network, is still in a "beta" stage, and in our experience, must be manually triggered regularly in order to update its state — this can trigger "false alarms" when the state of the cache manager database becomes stale, and does not reflect the actual network state. It is unclear when the cache manager will be robust enough to support automated auditing. Second, we are investigating the extent to which the **PLN** architecture can support reconfiguration of host nodes by a source that is not completely trusted.

In future work, we plan to investigate how the network might adapt automatically to changes to commitments through harvesting requests to participating hosts to perform additional harvesting. We also intend to identify ways to make the network self-repairing, so that deviations from policy commitments are repaired using the same request mechanism. That said, having the ability to audit the network against a formal policy is a useful innovation on its own. Our prototype serves as a proof-of-concept of the ability of **LOCKSS** to accommodate the institutional needs of archives as well as libraries. For more information on **Data-PASS's** approach to archival replication, including our policies and practices, software,[13] and schemas see the **Data-PASS** Website: *http://data-pass.org*. 🌱

Micah Altman, Bryan Beecher & Jonathan Crabtree

Micah Altman, Senior Research Scientist
Institute for Quantitative Social Science, Harvard University
<Micah_Altman@harvard.edu> • *http://maltman.hmdc.harvard.edu*

Bryan Beecher, Director of Computing & Network Services
Inter-university Consortium for Political and Social Research
University of Michigan

Jonathan Crabtree, Assistant Director for Archives and Information
Technology, Odum Institute for Research in Social Science
University of North Carolina, Chapel Hill

MICAH ALTMAN (Ph.D. California Institute of Technology) is Senior Research Scientist in the **Institute for Quantitative Social Science** in the Faculty of Arts and Sciences at **Harvard University**, Associate Director of the **Harvard-MIT Data Center**, and Archival Director of the **Henry A. Murray Research Archive**. **Dr. Altman** conducts research in social science informatics, social science research methodology, and American politics, focusing on the intersection of information, technology, and politics; and on the dissemination, preservation, and reliability of scientific knowledge. His current research interests include survey quality; computationally reliable and efficient statistical methods; the collection, sharing, citation and preservation of research data; the creation and analysis of networks of scientific knowledge; and computational methods of redistricting. His over thirty-five publications and five open-source software packages span informatics, statistics, computer science, political science, and other social-science disciplines.

WHERE DOES MICAH SEE THE INDUSTRY IN FIVE YEARS: The next five years will see a further application of **Melvin Kranzberg's** second law of technology: "Invention is the mother of necessity." We are now seeing a flood of new inventions for the individual and social creation, recording, and dissemination of knowledge. Information is becoming ambient — continually created, and recorded. Libraries, by necessity, will need to take a larger role in integrating with the workflows around this knowledge, and especially around tagging, social annotation, and other forms of social creation of secondary data.

BRYAN BEECHER is Director of Computing & Network Services in **ICPSR**. **Mr. Beecher** is responsible for the delivery of information technology systems to support the mission of **ICPSR**, and for the ongoing support of those systems. He supervises staff responsible for the day-to-day support of **Windows**, **Macintosh**, and **UNIX** computing systems, common business applications, and specialized data analysis applications, as well as staff responsible for designing and developing customized internal data systems, and Web-based software systems used by the member community.

JONATHAN CRABTREE is Assistant Director for Archives and Information Technology at the **Odum Institute for Research in Social Science** at **University of North Carolina, Chapel Hill**. **Crabtree** joined the institute 15 years ago and is responsible for designing and maintaining the technology infrastructure that supports the institute's wide array of services. Before moving to the social science side of campus he was an information systems technologist for the **UNC School of Medicine**. In addition to his work at the institute he is working part time on an advance degree in the **School of Information & Library Science** at **UNC**. 🌱