

## Policy-based Preservation Environments

REAGAN MOORE, ARCOT RAJASEKAR, MIKE CONWAY, ANTOINE DE TORCY, JEWEL WARD, AND HAO XU

**Abstract:** The management of preservation policies for large collections can be an onerous task. Typical policies may control ingestion of records, extraction of required provenance information, replication of records, retention and disposition, and validation of assessment criteria. Within the iRODS integrated Rule-Oriented Data System, policies are turned into computer-actionable rules that control the execution of procedures. The procedures are composed from functional units called micro-services, and are executed at each storage location linked by the data grid. The policies are applied on each manipulation of a record. The policies can also be executed periodically, making it possible to automate policy enforcement. An archivist can implement a policy/procedure set that enforces policies that are unique to the desired preservation environment. We will present a demonstration of the NARA Transcontinental Persistent Archive Prototype (TPAP) preservation environment, including the composition of preservation policies and automated enforcement of policies. A user interface framework will be demonstrated that manages ingestion of records into a staging area, automates application of preservation policies, and tracks the status of the processing steps.

### About the author:

*Reagan Moore* is the Director of the Data Intensive Cyber Environments Center at the University of North Carolina at Chapel Hill, professor in the School of Information and Library Science, and Chief Scientist at the Renaissance Computing Institute. Moore coordinates research efforts in development of policy-based data management systems that are used to support data grids, digital libraries, processing pipelines and persistent archives. Moore is the principal investigator for the development of the integrated Rule Oriented Data System (iRODS). The iRODS technology automates the application of management policies, automates validation of assessment criteria, and minimizes the labor required to manage massive distributed data collections. Moore has a B.S. in physics from the California Institute of Technology (1967), and a Ph.D. in plasma physics from the University of California, San Diego (1978).