

A Model Technological and Social Architecture for the Preservation of State Government Digital Information

EXTENSION REQUEST, April 2009

Project manager Minnesota Historical Society (MHS)

Contact Robert Horton
State Archivist, Director of Library, Publications and Collections
345 Kellogg Blvd W
St. Paul, MN 55102
Voice: 651 259 3240, Fax: 651 296 9961
robert.horton@mnhs.org

Lead Partners California Digital Library (CDL), Patricia Cruse
National Conference of State Legislatures (NCSL), Jo Anne Bourquard
California Legislative Counsel, State Library and State Archives
Minnesota Revisor's Office, Legislative Reference Library and State Archives

State Partners Current: Illinois (Connie Frankenfeld), Kansas (Matt Veatch), Mississippi (David Pilcher), Tennessee, (Wayne Moore), Vermont (Gregory Sanford)

Proposed: Arkansas (Mary Brewer), Nebraska (Andrea Faling), North Dakota (Ann Jenks)

Project website www.mnhs.org/ndiipp.htm

Request

Since 2008, the Minnesota Historical Society has collaborated with its partners and the Library of Congress to: 1) create a trustworthy technological and social architecture for capturing, managing and providing access to legislative content from California and Minnesota; 2) test the capacity of other states to adopt that model; 3) analyze the results with the legislative and government records communities; 4) define the thresholds for participation, to identify how different states could position themselves to implement similar programs; 5) use education and outreach to promote further collaboration and sustainability; and 6) develop compelling business cases for further investment.

To build on the success of this project so far, the Society requests additional funding to extend its efforts to include:

- pilot implementations in two states, California and Kansas
- inclusion of those states in the collaboration with the California Digital Library, to test a preservation repository and web harvesting

- inclusion of those states in the evaluation and benchmarking process directed by Prof. Cal Lee of the University of North Carolina
- tests of the audio file repository developed by the Washington State Archives
- the addition of four more states to the partners' group (ND, AR and NE confirmed, one more to be identified)
- annual site visits to all state partners, along with annual all-partners meetings
- continued collaboration with the NCSL, with the offer to host workshops at its meetings
- expanded education and outreach to the partner states, including technology workshops now being developed in Minnesota
- travel and presentations to support and promote the project, such as continued collaboration on the NCCUSL authentication of legislative documents study group

The extension will not fundamentally change the original plan's goals, but will expand its scope and extend its timetable. Including two more states in the tests and adding four more partners will require an additional 18 months, pushing the completion date to July 2011. This will allow for more comprehensive testing and analysis of the project in Minnesota, through 2009, followed by implementation and analysis in California and Kansas in 2010, then evaluation, outreach, reporting and educational efforts in 2011.

2008-2009 Project Implementation and Developments

The project began with an intensive analysis of the needs and capacities of the various stakeholders. This included an ongoing series of meetings in Minnesota, along with site visits to all partner states, presentations at NCSL forums, discussions with user groups and professional communities, such as staff at the Library of Congress, members of the NDIIPP partnerships and archival and library conferences.

These conversations informed the development of the legal, policy and technological framework for the project. Products, available on the project website, were developed in response to partner concerns. Projects to date include:

- Authentication Practices and Policies
- Digital Audio and Video Options
- Legislative History Resources
- Preservation Options for Digital Materials
- Survey of Partners' Legislative Records Online
- Web Content Accessibility
- XML Usage Survey
- Records Retention Policies for Selected Legislatures
- XML Schema for Legislative Content

The focus of current activity is on developing and implementing a technological model for providing access to and preserving at risk government digital content. Developers from Syntactica are working on the access portion of the project, and we will be using the California

Digital Library (CDL) digital archives for preservation. We are also in the process of testing Heritrix, a web harvesting application, and Bag-It, a tool that assists with authenticating data.

As the project moves further along, Prof. Cal Lee, of the University of North Carolina, has agreed to undertake the evaluation and analysis, benchmarking the model and its recommended practices against other preservation architectures, particularly those developed by other NDIIPP partners, such as the LOCKSS implementation sponsored in Arizona and the Washington State Digital Archives.

In the project so far, the partner states function primarily as an advisory board. They have had a great deal of influence in shaping the policy framework and products, as outlined above; these largely reflect the interests and questions the states have posed in a variety of meetings. As Minnesota develops its technological model, the states will help evaluate its capacity to work in different environments. The project assumes that “one size will not fit all,” so any method of capturing, preserving and providing access to digital legislative content will have to be practical and adaptable.

Extension

The extension request is for additional time and additional funding. With additional time, we would be able to further the work we can do with the model we are developing. With additional funding we can take the model developed in Minnesota and test it in two other states; these are tentatively identified as Kansas and California, both of which have agreed to undertake a more engaged partnership with this project, contingent on funding that would cover costs for staffing and equipment.

Based on the current work schedule, Minnesota will have a working model ready for evaluation discussion by the December all partners meeting. California and Kansas would then have the support for their implementations and tests through 2010. As noted, the model will undoubtedly need some modifications to fit the local, unique environments. This would affect the concomitant relationship with the CDL and be part of the evaluation by Cal Lee. The CDL and Prof. Lee have also agreed to participate in an extended collaboration.

The results of these efforts would demonstrate the potential of a federated partnership, with state entities managing the identification and acquisition of legislative content, within a common framework of standards and shared services that support cost-effective and efficient preservation and enhanced access. In particular, the partnership would address the needs expressed by various user groups, to support both search and discovery across state content and access to content in batch formats.

In addition, given the interest in the audio file preservation and search functions developed recently by the Washington State Digital Archives, we will test the transfer of legislative audio files to the NDIIPP sponsored project there. Virtually all the state legislatures in the partnership use the web to provide access to current deliberations; increasingly, these mean streaming audio and/or video from floor sessions and committee hearings. One state, Tennessee, has made a long

term commitment to capturing and preserving this content. Another, Kansas, is interested in expanding the capacity in order to promote accountability and transparency in government. These needs, together with the opportunity to work with another NDIIPP project and take advantage of the technology it has fostered, making this a compelling addition to the project.

We will also expand our educational and outreach efforts to the initial and the new partner states, to focus on facilitating dialogues with the now rather large array of partners. This will aim at initiating closer evaluation of the various policy products from the unique perspective of each state, as well as allowing us to plan and schedule more intensive educational offerings, as, for example, the three day technologies training course currently under development at the MHS. As we have found, education is a positive approach to encouraging collaboration locally, introducing staff from archives, libraries and legislatures to common concerns and solutions.

Deliverables

Along with the deliverables outlined in the initial proposal for this project, others would be added as part of the extension:

- Visit new partner states in 2009 to familiarize staff with the project and learn about their current environment and work, issues, and ideas regarding access to legislative content
- Implement access application in CA and KS
- Evaluate pilot implementations in MN, CA, and KS to provide benchmarking and cost-benefit analyses
- Develop technology training workshop/s for NCSL and other states
- Test CDL preservation repository option (dark archives) with selected MN, CA and KS legislative content
- Comparison and cost-benefit analysis of selected preservation options, including CDL dark archive, LOCKSS (PeDALs project), WA digital repository, and in-house storage (MHS)
- Test WA audio repository and search functions with selected MN and TN legislative files
- Site visits to all partner states in 2010
- All-partner meetings in 2010 and 2011
- Continued collaboration with and presentations to the NCSL and other professional associations