Preserving the Records of the E-Legislature

Final Project Report
National Historical Publications and Records Commission Grant #5416-MN

2008

Robert Horton
Shawn Rounds
Elizabeth Lighthipe

State Archives
Collection Department
Minnesota Historical Society
345 Kellogg Boulevard West
Saint Paul, Minnesota 55102

www.mnhs.org/elegislature

Table of Contents

Progression of Work

Funding for the "Preserving the Records of the E-Legislature" project was awarded to the Minnesota Historical Society (MHS) in 2005.[1] Project activities fell into five categories:

- analysis and appraisal of systems and records
- research into discovery and access models
- technology testing for long-term preservation
- cost-benefit analysis
- collaboration and communication.

Work in these categories proceeded concurrently.

Analysis and Appraisal of Systems and Records

To start the collaboration process, as well as the research phase of the project, the MHS team began meeting with local project partners in July 2005, with the first order of business to understand the universe of Minnesota's legislative records in general and the records produced by the Minnesota Office of the Revisor of Statutes (ROS) in particular.[2] ROS staff[3] offered presentations, demonstrations, and discussions on the role of the office in the legislative process, as well as the supporting information systems which enable the work, most notably, the XTEND XML-based bill drafting system.[4] Detailed discussions focused on the ROS' bill drafting workflow to identify possible points of record capture.

During the workflow analysis, ROS staff described the various document types that are created and managed within the XTEND system. They identified three categories that they thought would be good candidates to use to test the technology of off-site storage, management, and retrieval through the San Diego Supercomputer Center's Storage Resource Broker (SRB) system[5]: the finalized session laws, statutes, and administrative rules, along with accompanying tables, indexes, and graphic figures. ROS staff explained that all drafts prior to the finalized documents are considered confidential, and that they did not feel comfortable storing non-public documents in a non-ROS system for the purposes of testing even though they were interested in the potential of the system for secure remote disaster recovery storage.[6] As part of the system

[1] A summary of the project proposal is presented in Appendix A of this report.
[2] For a comparative glossary of Minnesota and California legislative terms, see Appendix C of this report.
[3] Michele Timmons and Tim Orr participated in all ROS discussions. Kyle Johnson led the ROS' technology tests.
[4] The XTEND system was put into production in 2005. For more details on the system, see Appendix D of this report.
[5] For a complete explanation of the SDSC's SRB technology and related work, see www.sdsc.edu/srb/index.php/Main_Page.
[6] For a complete summary of the appraisal of the ROS document types, see Appendix E of this report.

and record analysis work, ROS staff completed an inventory of the XTEND system using the Trustworthy Information Systems methodology and worksheet.[7]

To understand the ROS' work in the larger context of the legislature as a whole, the MHS team met with three other groups – the House and Senate information technology (IT) departments and staff from the Joint Departments and Commissions – to discuss the records that each creates and manages, and to promote the project in general. The discussions revealed important issues.[8] First, only a handful of document types are created and used in XML; most are generated in software like Word and then converted to HTML and/or PDF for presentation online. A number of other applications and formats are also used, including PowerPoint, Excel, Access, PageMaker, SQL database, MP3, WMA, WMV, and JPEG. Document availability over time is not consistent – sometimes newer content replaces older, while in other cases all content is kept online. The preservation of digital audio and video over time is a particular concern.

The MHS team also met with staff from the Legislative Reference Library (LRL) and the MHS library to learn what types of legislative documents are collected and managed by each repository. The gathered information was compiled into a summary table that included date ranges, formats, locations, future plans. These records, along with the other documents created by legislative bodies, are of interest to a wide range of audiences and are generally available through various means. There is a partnership role for MHS, however, as coordinator between these resources, a knowledgeable intermediary to evaluate resources and recommend linkages.[9]

Technology Testing for Long-Term Preservation

The technology component of the project proceeded in four phases: (1) installation of hardware for local storage, (2) installation of the SRB software, (3) connection to the SDSC, and (4) file management testing.

The SRB is a data grid management system that can be thought of as a sort of middleware between the systems of creation and distributed storage resources, providing seamless file management and access across the systems. A strength of the SRB is that is handles files, which are grouped into collections and containers, through logical persistent identifiers rather than physical names and locations. This means that the storage is "infrastructure independent" – a file's physical location in the distributed system may change, but that is invisible to the user. This allows for changes in hardware, for instance, to occur without disruption to file management and access. [10]

[7] The Trustworthy Information Systems (TIS) methodology helps agencies develop systems for maintaining authentic and reliable records. The TIS Handbook is available online at www.mnhs.org/preserve/records/tis/tis.html. For more information about the XTEND system, see Appendix D of this report.
[8] For summaries of these Minnesota legislature document types, see Appendix F of this report.
[9] Exploration of this role in detail was considered beyond the scope of this project. However, the MHS' NDIIPP project will build on this research and evaluate the issues in providing access to contextual legislative information residing in different locations.
[10] For a complete explanation of the SDSC's SRB technology and related work, see www.sdsc.edu/srb/index.php/Main_Page.

File management and access with the SRB can be carried out in a number of ways, including inQ, a graphical interface for Windows with drag-and-drop functionality, mySRB for web-based access, and Scommands, which are utility routines that can be used in a Unix shell or in a Windows DOS-command shell. The Scommands are very powerful and offer the most functionality, but are difficult to use, requiring very strict syntax.

By September 2005, the MHS team and ROS staff had each successfully installed the local storage hardware within their IT environments. The hardware at each site consisted of a custom-built, two-terabyte, RAID 5 server called a "grid brick."[11] The next step in the process was to install an operating system and then the SRB software on each grid brick. The ROS installation went very smoothly since the SRB was loaded along with UNIX, the platform the SRB was originally designed for.

However, MHS, which runs a Windows-based network, experienced some significant technological difficulties with the SRB installation and use. At the time, the Windows version of the SRB was a new development, largely untested outside of the SDSC environment.[12] As a result, MHS staff found bugs in the Windows SRB software as well as the Windows version of the Scommand set, difficulties that were compounded by a paucity of detailed documentation. Additionally, there were problems getting MHS and SDSC systems to communicate through the MHS firewall. Addressing these issues was a very time consuming process of give and take between MHS and SDSC staff by telephone and e-mail, with some solutions requiring custom scripts and intensive, technically detailed work that was well beyond MHS skill levels. Even the simple task of registering files in the SRB was stymied by bugs in the Scommands, with SDSC staff finally having to do the process from their end. Over the course of the project, new versions of the SRB were released offering enhanced performance, but the upgrades also made it difficult to keep up with the technology and set a routine, a requirement for a stable, sustainable preservation system.[13]

The obstacle which finally proved to be the limiting factor on MHS' participation in transfer tests was bandwidth. The MHS has a centralized IT department and network system that serves the entire Society, from staff to on-site patrons to web site visitors. Network bandwidth is shared between all MHS functions and simply proved to be insufficient for the transfer of large files over the Internet, even when the tests were conducted at off-peak times.[14] Bandwidth also proved to be an issue for the ROS as well, with system timeouts often occurring during file transfers.[15]

---

[11] RAID technology features redundant drives to protect from data loss. In March 2006, one of the RAID drives in the MHS grid brick failed during replication tests and had to be replaced with hardware from the manufacturer. While testing had to halt for a time, no files were corrupted or lost.

[12] The Windows version of the SRB also cannot be configured with a metadata catalog or MCAT. www.sdsc.edu/srb/index.php/FAQ

[13] Since 2003, there have been 17 SRB versions released. MHS began by installing SRB version 3.3.1 (released April 2005). Upgrades were issued in October 2005 (version 3.4), April 2006 (version 3.4.1), June 2006 (version 3.4.2), and December 2007 (version 3.5, current). For more information, see www.sdsc.edu/srb/index.php/Release_Notes.

[14] The MHS network runs at approximately 2-3 Mbits/second. SDSC staff have suggested work-arounds for bandwidth limitations, including bulk loads of files and using multiple transfer streams (breaking things up into small packets and sending all together at the same time), which have not been tested.

[15] For details of the ROS' transfer tests, see Appendix H of this report.

Given MHS' bandwidth issues, the project team developed a workflow that put the burden of file transfers on the ROS, but which accommodated both disaster recovery as well as archives requirements and functionality. The process begins with ROS staff transferring selected files through the SRB directly to the SDSC into a disaster recovery data grid (storage area) once they had been declared final within the XTEND system. Upon notification of the transfer, MHS staff performs necessary archival management activities remotely in SDSC's network environment, including replication of files into an archival data grid.[16] These copies are then considered the "master files," with only MHS staff having access to them.[17]

Dr. Reagan Moore, SDSC, notes, "Each preservation environment is unique: local network configuration (firewalls, network infrastructure, storage systems, operating systems), local management policies (arrangement, description, preservation, access), preferred access mechanisms (web browser, portal, DSpace, digital library, Windows browser). [The] challenge is to provide unifying infrastructure that enables a preservation environment to span multiple sites."[18] He has also identified that a crucial further step is the automation of preservation policies/processes. Certainly the MHS-ROS workflow offers several points where automation would create more efficient processes.

This project demonstrated that the ability to adopt, customize, and sustain technologically advanced preservation systems such as the SRB was beyond the reach of the MHS and possibly most archival repositories save those within larger institutions with extensive technological capacity. A number of factors contribute to this conclusion, including the high level of customization required to fit such a system into existing infrastructure and workflow, the high level of technical skill needed to manage the system, and cost in terms of time and money for responding to new releases and technology changes. As well, such advanced systems usually require advanced infrastructure, lack of which may degrade or completely prohibit implementation. Partnerships between archives and/or records creators is, of course, one possible way to address these issues in a cost-effective manner.[19]

Researchers pursuing further work with SDSC technologies, however, will be faced with a difficult choice: as of September 2007, the SDSC has frozen the SRB at version 3.5 and no further upgrades will be issued. At some point, presumably, the system will no longer be supported except, perhaps, by user communities. To take its place, the SDSC has developed the iRODS system, which features a rule-oriented architecture that supports the automation of data management rules/policies.[20] While the enhanced functionality offered by iRODS is appealing for disaster recovery and preservation activities, the issues outlined above still apply with the

[16] For a complete discussion of the workflow, see Appendix G of this report.
[17] The ROS provides public access to these records through its web site (https://www.revisor.leg.state.mn.us/pubs/).
[18] Reagan Moore, et al., "Partnership for Preserving the Records of the E-Legislature," presentation given at NAGARA conference, July 2007. Available online at www.mnhs.org/elegislature.
[19] These conclusions are supported by other work, including the Persistent Archives Testbed (PAT) project (NHRPC grant no. 2004-008) as described by Richard Marciano, SDSC, in the project's final report (May 2007, www.sdsc.edu/PAT). See also Martha Anderson, et. al., "Data Center for Library of Congress Digital Holdings: A Pilot Project," final report (2008, http://chronopolis.sdsc.edu/assets/docs/SDSC_LC_data-storage_report_2.pdf).
[20] For more information about iRODS, see https://www.irods.org.

added cautionary note that, as a relatively new system, it is not yet in stable production and will be going through a series of version releases as it is refined and enhanced.

Research into Discovery and Access Models

As outlined in the project proposal, it was the MHS team's original intention to pilot a web-enabled architecture for public access to legislative records, beginning with those from the ROS' XTEND system, which have the advantage of being in XML instead of a proprietary format. However, as research into legislative records and systems in Minnesota and elsewhere commenced, it quickly became apparent that any such application would have limited portability because there is no consistent or standardized architecture within Minnesota government or even within the legislature, let alone across states. XML, which offers a great deal of flexibility in terms of file management, presentation, and access, is only in limited use, with many record creators continuing to opt for formats such as Word and PDF offered through HTML-based web sites enhanced by home-grown search engines. While some commercially available software has the ability to bring together varied records from disparate sources, such applications inevitably require customization which often cannot be leveraged when migrating to new versions or to different software.

Even as more state legislatures move toward XML-based bill drafting systems, there is no nationally recognized standard XML schema.[21] Most state legislatures are using customized systems put together by vendors. While these systems generally offer the advantage of records in a non-proprietary format, information between systems cannot be as easily shared as would be the case if everyone was using a basic set of common elements.[22]

Given that large sets of legislative records are available online in various forms, the project team tested the Archive-It web harvesting service to see if the system could be used as an access option that would offer information within its original online context.[23] However, MHS tests targeting several Minnesota state government pages highlighted a lack of functionality of the service for the team's purposes, namely that it did not pick up dynamically generated web content.[24] This was a crucial shortcoming in the case of Minnesota's legislative content in that the ROS served up statutes and session laws only as .asp and .php pages.[25]

---

[21] The National Association of Legislative Technology (NALIT) newsletter (www.ncsl.org/programs/lis/nalit/nalitnws.htm) occasionally features articles about XML system development within state legislatures. While groups like NALIT and the National Conference of State Legislatures (NCSL) have sponsored presentations on XML, they have not been able to move forward successful cross-state XML work for legislative content.

[22] The Library of Congress NDIIPP-funded project, "A Model Technological and Social Architecture for the Preservation of State Governement Digital Information, led by MHS, is developing an legislative information interchange schema. For more information about the project, which got underway in January 2008, visit the project web page at www.mnhs.org/legistlativerecords.

[23] Access to Archive-It's collections is available through the service's home page at www.archive-it.org.

[24] The Archive-It service continues to be refined and expanded with new version releases. For up-to-date, general information on the service, see www.archive-it.org/public/faq.

[25] The ROS has changed its web site and now offers online access to statutes, laws, and administrative rules through a secure interface at https://www.revisor.leg.state.mn.us.

As access technologies continue to advance, content owners must periodically analyze a number of factors to be sure that they are offering their information in the most appropriate, cost-effective manner. Analysis should include identifying and prioritizing: (1) user groups, (2) what user groups want in terms of search, content, and presentation/delivery, (3) what content is available, (4) partnership opportunities, (5) funding sources and/or revenue streams, (6) infrastructure gaps in terms of standards, policies, etc., (7) infrastructure technology gaps, (8) points of possible enhancement, value-added opportunities.

Collaboration and Communication

At the beginning of the project, a web site was established to serve as a communication tool between MHS and project partners and the public, providing access to resources such as reports, meeting summaries, and presentation materials as they were created.[26]

The MHS team met several times with ROS and LRL members, as well as other knowledgeable legislative staff, over the course of the project to discuss all aspects of the work. This level of close communication resulted in strong archives-government partnerships that have continued into new efforts such as the Library of Congress NDIIPP-funded project "A Model Technological and Social Architecture for the Preservation of State Government Digital Information."[27]

In order to engage all project partners, a two-day meeting was held in Saint Paul in September 2005. Representatives from every partner organization attended, as well as a staff member from the Minnesota Senate Information Systems office. Over the course of the meeting, participants discussed the goals and structure of the project, learned about the key technologies involved (e.g., the XTEND system and the SDSC's SRB software), and shared information with each other about current and planned legislative records initiatives.[28] Several issues were raised for consideration during the project, including expanding captured formats beyond XML, audience/use studies/appraisal to target popular and critical record sets to enhance and contextualize those in the bill drafting system, the challenges of cultural and institutional changes, metadata standards for management and access, researching bill drafting systems and initiatives from other states, and sustainability. Also of interest was testing the proposed technology solution offered by the SDSC.

The MHS team got a chance to share information about the SDSC's SRB system with other users when they were invited to participate in a meeting discussing the "Persistent Archives Testbed (PAT) project, an NHPRC-funded effort.[29] The meeting, which was hosted by the SDSC and held in San Diego on February 9-10, 2006, allowed the MHS team to present its preliminary

[26] The project web site is available at www.mnhs.org/elegislature.
[27] For more information, visit the project web site at www.mnhs.org/legislativerecords.
[28] A summary of the meeting is available on the project web site at www.mnhs.org/elegislature.
[29] Information about the PAT project is available online at www.sdsc.edu/PAT

findings regarding the technology testing to SDSC staff and to hear about the experiences of others working in similar environments with government records.[30]

The project partners met as a group twice more. In June 2006, they gathered in Sacramento for a two-day session organized by the staff at the California Legislative Counsel. Presentations, demonstrations, tours, and discussion focused on California legislative systems and records as a point of comparison to Minnesota.[31] The last partners' meeting was held in Saint Paul in September 2007 as a project wrap-up session, reviewing the work completed, conclusions, and next steps.[32]

To promote the project to a wider audience and gain input from legislative system experts, Robert Horton, the project director, and Robbie LeFleur, the director of the LRL, attended the 2005 Professional Development Seminar hosted by the National Association of Legislative Information Technology (NALIT) in Rapid City, South Dakota.[33] They spoke on the general topic of "Managing the Creation, Usage and Storage of Data," describing the project work to the assembled group of legislative managers and technical specialists.

Michele Timmons, the Minnesota Revisor of Statutes, traveled to Chicago in 2007 to attend a conference sponsored by the American Association of Law Libraries, "Authentic Legal Information in the Digital Age: A National Summit." Timmons discussed the project in her presentation which focused on her experiences authenticating legal information on the Web.

The project team also reached out to other states. Horton traveled to Vermont in November 2005 to meet with a work group of legislators and legislative and state archives staff interested in the work as they moved ahead towards their own electronic legislature initiative. Shawn Rounds and Elizabeth Lighthipe attended two national archives and records conferences in 2007. At the "Best Practices Exchange 2007: Libraries and Archives in the Digital Era" conference held in Arizona, Rounds and Lighthipe gave separate presentations on the technical aspects of the project as well as the record appraisal and selection process.[34] Rounds and Lighthipe teamed up with Reagan Moore, Director for Data and Knowledge Systems at the SDSC, to talk about all aspects of the project at the annual conference of the National Association of Government Archives and Records Administrators (NAGARA)[35]

Conclusion: Cost-Benefit Analysis

There are several aspects to consider in the cost-benefit analysis. As noted, each institution and state might well come to a very different conclusion, working with the same facts, as each has an

[30] Other participants in the meeting included staff from the Kentucky Department of Library and Archives, the Ohio Historical Society, and the Michigan Department of History, Arts and Libraries.
[31] A summary of the meeting is available on the project web site at www.mnhs.org/elegislature.
[32] A summary of the meeting is available on the project web site at www.mnhs.org/elegislature.
[33] Information about NALIT is available online at www.ncsl.org/nalit
[34] Conference information is available online at www.bpexchange.org/2007
[35] The Council of State Archivists (COSA) held its annual meeting in conjunction with NAGARA's conference. For more information about NAGARA, visit www.nagara.org. Information about COSA is available at www.statearchivists.org.

individual and unique context in which it works. But, every entity has to work with those facts, in its specific context, to determine a compelling business case or rationale that makes sense to it.

Out of pocket costs did not apply here, as the grant generously provided funding for storage costs at the SDSC, a grid brick server for the ROS, and expert consulting from the SDSC, ROS, and MHS as each contributed knowledge to the effort. Without that support, though, the initial and ongoing investments would have been considerable; indeed, they would represent, more or less, nearly the total costs of the project.

It is quite clear that the Minnesota legislature would not invest this sum or anything like it of its own volition. Simply, the idea of long-term preservation does not have sufficient appeal to warrant institutional support at this level, especially at a point when the legislature was investing funds in technologies that enhanced its basic business routines and provided additional access to legislative content to its constituencies. As we moved forward with the project, it became increasingly evident that preservation was, at best, a concept that had appeal primarily when paired with disaster recovery.

This is partially a question of semantics, partially of priorities. The latter recognizes the increasing restrictions on state budgets, leading to the understanding that all funding requests struggle for attention in a highly competitive arena. What we could conclude from this project is that the most attractive requests delivered a tangible outcome and supported current needs; we could go further and add that, ideally, they also received the support of an important audience and constituency.

What we could see in this project is that the tangible outcome was not evident. The technology worked, but we could not envision its integration into the normal business routine of the ROS and the MHS. The difficulties we faced in its implementation made its viability as a disaster recovery solution uncertain, especially as there were simpler options at hand. The first was simply an elaboration of current practices: off-site tape storage, with the site removed from the conceivable disaster area. While not ideal from the archival perspective, this is affordable, manageable, part of the current work routine, and within the realm of best practices for IT staff.

The second option is offered by the state's Office of Enterprise Technology (OET). It provides a centralized storage system and is developing a proposal for an enterprise-wide document management system to complement it. Costs are relatively low: between $3.00 and $4.50 per gigabyte per month contingent on decisions about storage media and retrieval time. As well, OET offers a "state of the art" disaster recovery solution: a fully replicated site with the estimate for the hardware and infrastructure at about $225,000.

Again, these are not archival, in the sense that an archival custodian is involved and archival care is applied. But, they meet the needs of the legislature which, of course, would be meeting the costs. Adding the State Archives to the equation would require it to provide an investment. We explored those options as well, looking at in-house storage and management, which currently relies on storage area network (SAN) solutions. The first SAN solution, accommodates online, immediate retrieval at a cost of about $55,000, which provides approximately 8 terabytes of

usable storage space. The second SAN solution, which is near-line with a slower response time, costs about $16,000 for similar storage capacity.

While storage costs at the SDSC are much less at roughly $400 per terabyte annually, the difference lies in the relative ease of implementation. All the local options are essentially manageable as part of the everyday routines of the staff and institutions involved. They do not represent another technology to learn, another partner to accommodate, another installation to manage, etc.

As the appraisal of legislative digital content indicates that the essential records to acquire and preserve are sessions laws, statutes, and administrative rules, all of which can be relatively easily extracted from an XML-based bill drafting system at the close of a session, copied to a portable hard drive and transported to the State Archives for ingestion, then the difficulties involved in working with the current manifestation of grid technology become ever more apparent. At the moment, looking at the entire spectrum of necessary infrastructure, skills capabilities, and other requirements, it seems clear that the total costs of ownership outweigh the benefits that records creators appreciate and the costs that the records custodians are able to bear.

That is not to say that a grid technology solution is inappropriate, simply that it demands more capacity than we can muster at this point. Newer applications like the SRB and iRODS are difficult to put into routine production, requiring complex IT skills and nimble IT infrastructure to adapt to changing requirements as new version are released and older versions abandoned. The level of service and assistance from the SDSC was extraordinary for this project, but we never reached the point where we could translate those into ordinary sustainable efforts.

Appendix A:
Summary of Project Proposal

1. Purpose and Goals of the Project

The purpose of the project is to preserve the records of the electronic legislature. As the legislative branches of government introduce new technology to their routine functions, they will create records in digital formats that have a tremendous potential. In 2005, the Minnesota Office of the Revisor of Statutes will move to an XML-based bill drafting application. This transition is an opportunity for a state archives to work with the right partners to preserve electronic records of undoubted legal and historical value. These are key records to the legislative process and, through that, to government and the history of government in Minnesota.

The goals of the project, working through a broad based partnership of Minnesota agencies and entities, are to develop and implement a plan to preserve these and other digital legislative records in a web enabled architecture that greatly improves their value. Many state legislatures are likely to move in the same directions as Minnesota's, as the National Conference of State Legislatures (NCSL) has recommended XML as the best technology for developing new bill-drafting systems and has drafted recommended standards for these systems. As a result, Minnesota can develop a sustainable model that can inform and support similar efforts across the country. To establish the value to other states, staff from corresponding California agencies will participate in the project to observe and evaluate Minnesota's implementation, to determine its pertinence to a different environment.

In the course of this grant, the State Archives Department of the Minnesota Historical Society, working with the Minnesota Revisor's Office and the Legislative Reference Library in Minnesota, along with the San Diego Supercomputer Center, will:

- analyze the legislative systems and appraise the resources they create and manage;
- plan for their long term preservation in a digital format, using grid technology;
- identify and evaluate other resources which could enhance the value of these records;
- test the information and technical architecture using grid technology, standard metadata and XML;
- pilot a web enabled architecture for public access;
- develop for the state legislature cost and benefit analyses for enhancing and supporting the system; and
- write and disseminate reports on the progress and products of the project.

To evaluate the model Minnesota creates, staff from the California State Archives, State Library and Legislative Counsel will work with their Minnesota counterparts over the course of the project. The California team will:

- serve on the project advisory board;
- analyze and appraise comparable California legislative resources;

- review the project work products from Minnesota; and
  evaluate their applicability and pertinence to California's systems and organizations.

2. Significance and Relationship to NHPRC Goals and Objectives

The NHPRC has long supported work on electronic records. Many of those efforts have addressed individual aspects of the challenges technology presents to archivists. As the Commission notes in its suggestions for electronic records proposals:

> *Research Issues in Electronic Records* defines four categories of activities—research, basic program development, analysis, and advocacy—all of which are equally competitive for Commission funding. The distinction the report and these suggestions make between the four areas is, to a certain degree, artificial. There is a relationship among research, analysis, implementation, and advocacy.

With this grant, the partners propose to cross these permeable lines among the four categories in a project that addresses and integrates their concerns. The research and analysis will focus on the legislative systems that are producing electronic records; the implementation will comprise the appraisal of the records and their preservation and description within the technological infrastructure currently maintained by the partners; and, last, advocacy will follow from the model provided for other states to study.

3. Plan of Work for Grant Period

If the grant is awarded, the project will begin in April 2005, with preliminary analysis and preparation. The full array of partners in Minnesota will be available to participate after the completion of the state's legislative session, in June 2005. In all, the plan includes five phases, with tasks related to education and reporting running continuously through all of them. The project will end in September 2007.

The five phases are: 1) preparation (April-June 2005); 2) research and analysis (June 2005-June 2006); 3) testing and implementation (June 2006-January 2007); 4) evaluation and advocacy (January 2007-June 2007); and 5) completion of final reports and products (June 2007-September 2007).

4. Products and Publications to be Completed During Grant Period

The project will:

- appraise and preserve Minnesota's legislative electronic records from the Revisor's bill drafting system;
- maintain a project web page, with all written materials (e.g., project reports) available online;
- complete reports on the model developed in Minnesota, with evaluations from the California perspective;

- provide reports on best practices, to help other states position themselves to learn from this model, covering systems, records, legislative and archival business functions, data formats, metadata, technological and organizational infrastructure;
- present reports on the project at professional meetings, as, for example, the 2006 SAA and NAGARA joint meeting;
- evaluate and appraise additional and ancillary legislative records as the context to bills and laws; and
- develop a cost/benefit analysis and recommendations for enhancement to the legislative and archival systems.

Appendix B:
Project Participants

The following people contributed in a significant way to the project and/or attended a project meeting as a representative of their organization.

Minnesota Historical Society
- Robert Horton, Project Director
- Elizabeth Lighthipe, Project Archivist
- Shawn Rounds
- Dave Sagstetter

Minnesota Office of the Revisor of Statutes
- Michelle Timmons
- Tim Orr
- Kyle Johnson
- Cory Oman
- Robyn Ender

Minnesota Legislative Reference Library
- Robbie LaFleur
- Elizabeth Lincoln
- Leif Eischen
- Mike Schatz
- Paul Van Cura

Minnesota Senate Information Systems
- Jim Greenwalt

San Diego Supercomputer Center
- Richard Marciano
- Reagan Moore
- Sheau-Yen Chen
- Bing Zhu
- Chien-Yi Hou

California Legislative Counsel
- Diane Boyer-Vine
- Bill Behnk
- Linda Heatherly
- Annie Anderson
- Mendora Servin
- David Duarte
- Maureen McRandle
- Cheri Meadows

California State Archives
- Nancy Zimmelman Lenoil
- Lucy Barber (moved to NHPRC in 2006)
- Chris Garmire
- Melodi Andersen

California State Library
- Janet Coles
- Kristine Ogilvie

Appendix C:
Glossary of Legislative Terms

Administrative Rules:

Minnesota[36]

An administrative rule is a general statement adopted by an agency to make the law it enforces or administers more specific or to govern the agency's organization or procedure. An agency may adopt a rule only after the legislature has enacted a law granting this authority to the agency. An agency rule that is adopted under the rulemaking provisions of Minnesota Statutes, Chapter 14, had the force and effect of law.

California[37]

Similar term, see California Code of Regulations.

Act:

Minnesota

Official name for a bill that has been signed into law. Each act is assigned a chapter number and every chapter for a particular year is published in the bound volume called the *Session Laws of the State of Minnesota*.

California

A bill passed by the Legislature and enacted into law.

Assembly:

Minnesota

Similar term, see House of Representatives

California

The house of the California Legislature consisting of 80 members, elected from districts determined on the basis of population. Two Assembly districts are situated within each Senate district.

Biennium:

Minnesota

The two-year legislative term, which begins in January of an odd-numbered year and ends in December of an even-numbered year.

California

Similar term, see Session.

[36] Minnesota definitions taken from "Frequently Asked Questions About the Minnesota Legislature" at www.house.leg.state.mn.us/leg/faqtoc.asp?subject=7 and "Glossary of Terms" at www.house.leg.state.mn.us/leg/faqtoc.asp?subject=18, accessed on October 19, 2007.

[37] California definitions taken from "Glossary" at www.legislativecounsel.ca.gov/Legislative+Counsel/Glossary/_Glossary.htm, accessed on October 19, 2007.

Bill:

Minnesota

A proposal calling for a new law, a change in current law, the repeal of current law, or a constitutional amendment. It consists of a title, enacting clause, and body (text), which is examined and approved by the Revisor of Statutes.

California

A proposed law, introduced during a session for consideration by the Legislature, and identified numerically in order of presentation; also, a reference that may include joint and concurrent resolutions and constitutional amendments.

Bill Analysis:

Minnesota

Similar term, see Bill Summary.

California

A document prepared by committee and/or floor analysis staff prior to hearing the bill in that committee or on the floor of the Assembly or Senate. It explains how a bill would change current law and sometimes identifies major interest groups in support or opposition.

Bill Cover:

Minnesota

A two-page, paper document created by the Revisor's Office prior to a bill's introduction. The bill cover contains all pertinent information describing the bill and its authors. House and Senate staff use the bill cover to record all official actions to the bill. Bill covers are green for the House and yellow for the Senate.

California

Legislative Counsel's system of tracking the bill. Contains title, author, bill numbers, date of introductions and first reading. Bill covers are white for the Assembly and goldenrod for the Senate.

Bill Summary:

Minnesota

A document prepared by either House Research or Senate Counsel and Research (both non-partisan offices) to summarize significant or lengthy bills.

California

Similar term, see Digest.

Calendar

Minnesota

The House and Senate each create their own daily calendars. The daily calendar contains all bills, resolutions, and other matters to be considered for final passage that day.

California
Similar term, see Daily File.

Chapter:

Minnesota
In preparing an enrollment of a bill passed at a legislative session, the Revisor's Office assigns the bill a chapter number. As far as practical, the numbers are assigned in the order of the date of the legislature's last vote of the bills before presentment to the governor. As soon possible after a session of the legislature has adjourned each year, the Revisor's Office publishes the laws of the session in a publication called *Laws of Minnesota.* Each bill is shown with the chapter number assigned to its enrollment.

California
When a bill has been passed by the Legislature and enacted into law, the Secretary of State assigns the bill a "chapter number" such as "Chapter 123, Statutes of 1992," which is subsequently used to refer to the measure in place of the bill number.

Codes:

Minnesota
Similar term, see Statutes.

California
Bound volumes of law organized by subject matter. Code sections to be modified by a bill are identified in the title of the bill.

Committee Report:

Minnesota
After a bill has its first reading it goes to a legislative committee. The committee reviews the bill and returns it to the House or Senate with a recommendation for action in the form of a report.

California
No similar term.

Conference Committee:

Minnesota
A group of legislators from the House and Senate appointed to reach a compromise between the House and Senate versions of a bill. Once finished, a conference committee report (suggests changes and compromises to the bill) is issued to both houses.

California
Usually composed of three legislators (two voting in the majority on the disputed issue, one voting in the minority) from each house, a conference committee meets in public session to forge one version of a bill when the house of origin has refused to concur in amendments to the bill adopted by the other house. For the bill to pass, the conference committee version must be approved by both Assembly and Senate. Assembly conferees

are chosen by the Speaker; Senate conferees are chosen by the Senate Committee on Rules.

Consent Calendar:

Minnesota
Bills on the Consent Calendar are placed there by recommendation of a standing committee which considers the bills to be non-controversial.

California
File containing bills that received no dissenting votes in committee.

Daily File:

Minnesota
Similar term, see Calendar of the Day.

California
Publication produced by each house for each day the house is in session. The publication provides information about bills to be considered at upcoming committee hearings and bills that are eligible for consideration during the next scheduled floor session. . . . The Daily File also contains useful information about committee assignments and the legislative calendar.

Daily History:

Minnesota
Similar term, see Journal.

California
Produced by the Assembly and Senate respectively the day after each house has met. The History lists specific actions taken on legislation. Each measure acted upon in that house the previous day is listed in numerical order.

Daily Journal:

Minnesota
Refers to either the Journal of the Senate or the Journal of the House, which are the officials records of the respective bodies. Each body's journal contains the proceedings for every day of the session.

California
A publication that is produced for each legislative day that contains the minutes of the session, vote information, motions, parliamentary inquiries and letters of legislative intent.

Digest:

Minnesota
Similar term, see Bill Summary.

California
The digest is a brief summary of changes the proposed bill would make to current law. The digest is found on the front of each printed bill.

Engrossment:

Minnesota
The current text of a bill or resolution which includes or incorporates all adopted amendments.

California
When a bill is amended, the printed form of the bill is proofread by staff to assure that the amendments are inserted properly. After being proofread, the bill is "correctly engrossed" and is thereupon deemed to be in proper form.

Enrollment:

Minnesota
A bill that has been passed by both houses and has been put in final form to be presented to the governor for his signature. The Revisor's Office creates the enrollment and presents the enrollment to the governor.

California
Whenever a bill passes both houses of the Legislature, it is ordered enrolled. In enrollment, the bill is again proofread for accuracy and then delivered to the Governor. The "enrolled bill" contains the complete text of the bill with the dates of passage certified by the Secretary of the Senate and the Chief Clerk of the Assembly. A resolution, when enrolled, is filed directly with the Secretary of State.

File Notice:

Minnesota
No similar term.

California
Bills that are scheduled for a committee hearing must be listed in the Daily File for not less than four days prior to the hearing.

File Number:

Minnesota
Similar terms, see House File and Senate File.

California
The number assigned to a measure in the Assembly or Senate Daily File. The file number changes each day as bills move on or off the Daily File. File numbers are assigned to measures on second and third reading; in conference; unfinished business (a bill amended in the other house and awaiting concurrence in amended form); and Governor's appointments. Legislation is taken up on the Assembly or Senate floor in

chronological order according to file number. Items considered on the floor are ordinarily referred to by file number.

Final History:

Minnesota
No similar term.

California
The publication printed at the end of every session showing the final disposition of all measures.

General Orders:

Minnesota
In the Senate, General Orders is a listing of bills that have had two readings, have been acted upon by one or more standing committees, and have been printed. If approved by a vote of the Senate, the bill is placed on the Senate Calendar to be considered for final passage.

California
No similar term.

General Register:

Minnesota
In the House, the General Register is a list of bills that have had their second reading and await action by the full House.

California
No similar term.

Handbook:

Minnesota
Similar term, see Red Book.

California
The 3" x 5-3/4" hardbound edition of "California Legislature" published for each two-year legislative session. Contains indexed versions of the Assembly, Senate, and Joint Rules; biographies of members; and other useful information. Published by the Chief Clerk of the Assembly and Secretary of the Senate.

House File:

Minnesota
A number assigned to a bill by the House before it is introduced. It is listed at the top of the bill.

California
Similar term, see File Number.

House of Representatives:

Minnesota

The Minnesota House of Representatives consists of 134 members elected from districts determined on the basis of population. Each Senate district is divided into two sections, made up of about 36,713 people each and identified with an "A" or a "B." Voters elect one House member, or representative, from each section.

California

Similar term, see Assembly.

Inquiry System:

Minnesota

No similar term.

California

A database containing bill analyses, bill status, bill text, votes, and other useful information for bill tracking and research by legislative employees in Sacramento and district offices. The system is maintained by the Legislative Data Center, which is a part of the Office of Legislative Counsel. To be replaced by the Legislative Information System, an XML bill-drafting system.

Law:

Minnesota

A law is an idea, placed in bill form, that has passed both the House of Representatives and the Senate and has been signed by the Governor.

California

[Laws are] the rules adopted by formal governmental action that govern our lives in various respects.

Legislative Counsel:

Minnesota

Similar term, see Revisor of Statutes

California

The Legislative Counsel (who is appointed jointly by both houses) and his or her legal staff are responsible for, among other things, drafting all bills and amendments, preparing the Digest for each bill, providing legal opinions, and representing the Legislature in legal proceedings.

Red Book:

Minnesota

The official directory of the Minnesota legislature. It contains House and Senate member information, permanent rules of the House and Senate, committee information, and other staffs contact information.

California
Similar term, see Handbook.

Revisor of Statutes:

Minnesota
The Office of the Revisor of Statutes is a legislative office that provides services to members of both houses of the legislature as well as all constitutional offices, and all state agencies and departments. The services of the office are non-partisan and confidential. The office assists members and legislative staff at most stages of the legislative process, from the preparation of a draft of a bill through its presentation to the governor. The office also creates and maintains computer systems used by the legislature and the public. These systems are used to create, edit, publish, and track the status of legislative documents.

California
Similar term, see Legislative Counsel.

Rules:

Minnesota
The regulating principles and procedures of the Legislature. Three sets of rules exist: Rules of the House, Rules of the Senate, and Joint Rules.[38]

California
Principles formally adopted to govern the operation of either or both houses. These include Standing Rules of the Assembly, Standing Rules of the Senate, and Joint Rules of the Senate and Assembly.

Senate:

Minnesota
The Minnesota Senate consists of 67 members elected from districts determined on the basis of population. There are about 73,425 people in each district. Voters elect one senator from each of those districts.

California
The house of the California Legislature consisting of 40 members elected from districts apportioned on the basis of population, one-half of whom are elected or re-elected every two years for four-year terms. Also known as the "Upper House."

Senate File:

Minnesota
A number assigned to a bill by the Senate before it is introduced. It is listed at the top of the bill.

[38] Permanent Rules of the House are online at www.house.leg.state.mn.us/cco/rules/permrule/permrule.htm. Permanent Rules of the Senate are online at www.senate.leg.state.mn.us/rules/index.php?ls=85#rulesofthesenate. Joint Rules are online at www.house.leg.state.mn.us/cco/rules/jtrule/jtrule.htm.

California
Similar term, see File Number.

Session:

Minnesota
1) The biennial period during which the Legislature meets; 2) regular session, the annual meeting of the Legislature between the first Tuesday after the first Monday in January and the first Monday after the third Saturday in May; 3) special or extra session, a meeting of the Legislature after the end of 120 legislative days in the biennium or after the date set by law for adjournment; 4) daily session, a meeting of the House or Senate in its chamber.

California
The period during which the Legislature meets. The Legislature may meet in either regular or special (extraordinary) session. Each regular session lasts two years.

Session Laws:

Minnesota
Session laws are the laws enacted at each session of a legislature. They are arranged and published by the Revisor's Office in the order of chapter number.

California
Similar term, see Codes.

Side by Side:

Minnesota
A comparison of selected portions from two bills. It is published as a PDF document with the corresponding portions from the two bills positioned side by side on the page.

California
No similar term.

Statutes:

Minnesota
A compilation of the general and permanent laws of the state, incorporating all new laws, amendments and repeals of old law. They do not include budgetary laws or local counties, cities, or temporary laws. *Minnesota Statutes* are published in their entirety every two years by the Revisor's Office.

California
Enacted bills, which are chaptered by the Secretary of State in the order in which they become law.

Summary Digests:

Minnesota

Similar term, see Session Laws.

California
Digests of each bill enacted in a two-year session, as prepared and compiled by the Legislative Counsel. The measures are listed by chapter number, reflecting the order in which they were signed into law.

XTEND:

Minnesota
A software system built and maintained by the Revisor's Office. The system is used to create many types of XML-formatted documents for the legislature, as well as Administrative Rules for the executive branch and Court Rules for the judicial branch.

California
Similar term, see Inquiry System.

Appendix D:
The XTEND Bill Drafting System

Summary of Functions

As described online, the "Office of the Revisor of Statutes is a legislative office that provides service to members of both houses of the [Minnesota] legislature as well as all constitutional offices and all state agencies and departments. The services of the office are nonpartisan and confidential…."[39] The duties of the ROS falls into a number of areas as follows:[40]

Legislative Duties

- Bill Drafting for Agencies
- Amendment Drafting
- Resolution Drafting
- Revisor's Bills
- Form Approvals of Bills
- House Committee Reports
- Conference Committee Reports
- Comparison Reports
- Conference Committee Comparison Reports
- Desk Comparison Reports
- Engrossing and Enrolling

Administrative Rule-Related Duties

- Rule Drafting
- Form Approvals of Rules
- Proposed Rules
- Modifications
- Notices of Adoption
- Adopted Rules

Publications

- Laws of Minnesota
- Minnesota Statutes and Supplement
- Indexes
- Minnesota Rules and Supplement
- Court Rules
- Local Laws Tables
- Bill Drafting Manual
- Rule Drafting Manual
- Rulemaking Guide

[39] www.revisor.leg.state.mn.us/
[40] Quoted directly from www.revisor.leg.state.mn.us/office/duties/

Legal Assistance and Liaison

- Advice Concerning the Effect of Bills
- Counsel to Subcommittee on Claims
- Court Opinions Report
- Uniform Laws Conference
- Compiling Data on Operation and Effect of Laws
- Counsel and Information for Other State Office and the General Public

Computer Services and Access to Data

- Maintaining a Computer System
- Computer Searches
- Copies of, and Access to, Public Data

With respect to the last category, the work of the ROS is varied and highly technical: "…maintain hardware and develop software to support the work of the legislature….provide support to users through training and through our Help Desk. Systems supported in House and Senate journals, calendars and agendas, House and Senate index, and bill drafting….also provide network and server support for the legislature's Internet services and provide Internet information on bills, rules, and laws."[41] Two ROS systems play center-stage roles in the legislative process: the XTEND bill drafting system and the bill status system, which is a web-accessible system for identifying and tracking current legislation.

XTEND is an XML-based legislative document processing system built upon commercial products such as Arbortext and Oracle, and customized for Minnesota's legislative environment. The system was put into production in 2005 and now offers service to over one hundred users who rely on it for creating and retrieving over two dozen document types. As described by ROS staff, "The system provides all document lifecycle operation. The ability to perform these functions is essential for the operation of the legislature. The operations are: Security, Document Editing, Document Management, Content Searching, Document Operations, Composition, and Publishing."[42]

System Analysis

The Trustworthy Information Systems (TIS) Handbook was developed by the Minnesota State Archives under NHPRC Grant No. 98-001. The goal of the system analysis methodology outlined in the handbook is to guide agencies through developing, implementing and maintaining information systems to meet records management and access requirements over time.

In March 2006, ROS staff completed a TIS analysis of the XTEND system, which walked them through examination of their system documentation and metadata, security and audit trail

[41] www.revisor.leg.state.mn.us/
[42] Timothy Orr, Office of the Revisor of Statutes, TIS analysis summary, 24 March 2006.

capabilities, and disaster recovery plans. This last item was of particular note as the ROS had interest exploring how the SDSC might serve as an off-site disaster recovery data storage center. At the time, the ROS did not have a full disaster recovery/business continuity plan, although by 2007 ROS staff were at work on a number of related initiatives, including securing backup generator capacity and stockpiling retired servers and network equipment for use in emergencies. As well, the ROS secured off-site storage at two locations for system backups on tape. [43]

[43] Timothy Orr, Office of the Revisor of Statutes, e-mail correspondence, 13 June 2007.

Appendix E:
Appraisal of Minnesota Revisor's Office Document Types

Summary

Shawn Rounds and Elizabeth Lighthipe, Minnesota Historical Society, met with Michele Timmons and Tim Orr, Office of the Revisor of Statutes (ROS), on September 27, 2005. Discussion focused on the ROS' paper and electronic workflow and associated document types.

In summary, Timmons and Orr identified a number of document types that they would like to have archived at the San Diego Supercomputer Center (SDSC) through the Storage Resource Broker (SRB) system:

- Final edited session laws
  - text
  - tables
  - index
- Final edited statutes
  - text
  - tables
  - index
- Administrative rules
  - text
  - tables
  - index
  - graphic figures

The legislature's web site offers session laws back through 1994[44], as well as the most current version of the statutes and administrative rules. These are considered unofficial versions; the official copies are the bound, printed volumes published by the ROS. These copies parallel the electronic versions and are printed after the legislative session and after an editing and proofing process.

Timmons and Orr expressed concern about the security of the SDSC system, especially in light of the fact that many of the working documents within XTEND are confidential. For the purposes of this project, the ROS did not store any confidential documents at the SDSC. While this resulted in only a partial backup at SDSC for disaster recovery purposes, all files were captured in locally stored daily system backups.

Workflow Notes

The primary role of the ROS is summarized as follows:

[44] Accurate as of the time of the interview. In 2007, the ROS scanned session law books dating back to 1849 and created PDF files of the images. These PDFs are available online at https://webrh12.revisor.leg.state.mn.us/laws.

> The Office of the Revisor of Statutes and staff from other legislative offices work with legislators in putting the idea for a new law into proper legal form. The Revisor's Office is responsible for assuring that the proposal's form complies with the rules of both bodies before the bill can be introduced into the Minnesota House of Representatives and the Minnesota Senate.[45]

The office also compiles, edits, and prints Laws of Minnesota, Minnesota Statutes, and Minnesota Rules. The documents are published in bound paper volumes as well as in digital form on the ROS web site. Each publication contains user aids and indexes, and are the official text of the law.

The ROS maintains a parallel system of paper and electronic files for bills moving through the legislative process. Quality control tasks are recorded on the paper folders, which are color-coded by document type.

Bill Drafts

Bill drafts are created in the XTEND system and have been in XML format since the 2006 session. Bill drafts are confidential. The ROS stores them in their own database. 2005 bill drafts are being converted from a proprietary file format into XML. The ROS keeps bills that die or fail in the system so that they can pull them up and reuse them if needed. The ROS has a records retention schedule for paper bill files which stipulates destruction after two biennium.

The ROS assigns an ID number for each draft in chronological order. This ID number is not the same as the House or Senate file numbers, which are assigned when the bill is introduced. The ROS ID number consists of a two-digit year code plus a four-digit number that is consecutively assigned (e.g., 05-0001, 05-0002). House Research and Senate Counsel also drafts bills and use the XTEND system. The bills they draft go to the ROS, where each is assigned an ROS identification number.

Each bill draft goes through multiple checks. Drafts first go to a drafting attorney, move to a non-attorney legal editor, and then go to a drafting and editing assistant for data entry and proofing. After these steps, drafts go back to a legal editor for quality control checks, then to the drafting attorney, and finally to another attorney for peer review. Quality control tasks include checking for repealed language, spelling, correct citations and cross-references, correct title, consistent internal references, and correct statutes.

Each bill draft is considered confidential and is recorded in the “Bill Track” database. This is an Oracle database used exclusively by ROS staff. Each bill introduced in the house or senate is also entered in the “Bill Status” database, an Oracle database that records significant legislative action and is available to the public through the legislature’s web site. The Bill Status database contains information from the 1995 legislative session to the present.

Once the final bill draft is done, the bill is “jacketed.” Both houses receive identical copies with the same ROS identification number in color-coded jackets (yellow for Senate, green for House).

[45] Minnesota House of Representatives Public Information Office, “How a Bill Becomes a Law in Minnesota,” 29 March 2000 (www.leg.state.mn.us/leg/howbill.asp).

The ROS delivers the bill to the legislators who are sponsoring it ("authors"), and the authors introduce the bill in their house. Both the House and Senate assign their own ID numbers to the bill, using different schemes than the ROS. Once the bill is introduced, it is considered a public record.

Committee Reports

After the bill is introduced and has received its first reading, it goes to a legislative committee. The committee reviews the bill and returns it to the house with a recommendation for action in the form of a committee report. A committee report recommends whether or not the bill should pass, and if so, gives suggestions on how to amend the bill if necessary. The ROS completes the official committee report out of the XTEND system. Committee report numbers are based on the house or senate file ones.

As stated in the *How a Bill Becomes a Law in Minnesota*:

> After the full House or Senate accepts the committee report, the bill has its second reading and is placed on the House agenda called the General Register or the Senate agenda called General Orders. (A committee can recommend that non-controversial bills bypass the General Register or General Orders and go onto the Consent Calendar, where bills usually pass without debate.) After this point, House and Senate procedures differ slightly.
>
> In the House, the General Register serves as a parking lot where bills await action by the full body. Bills chosen to appear on the Calendar for the Day or the Fiscal Calendar are drawn from the General Register.
>
> In the Senate, a different procedure is used. Bills are listed on the General Orders agenda. Senate members, acting as the "committee of the whole," have a chance to debate the issue and offer amendments on the bill. Afterwards, they vote to recommend: passage of the bill, progress (delay action), or further committee action. And sometimes they recommend that a bill not pass. From here, the bill is placed on the Calendar.[46]

Amendments

The second reading of the bill occurs after the body adopts the committee report. The House then discusses the bill and votes on any amendments coming from the report or from the floor. If there are amendments to the bill, the bill returns to the ROS. Amendments are created by the ROS in the XTEND system. From this point, the ROS will refer to the bill by its Senate or House file number, although it can still be pulled up in XTEND by its original ROS-assigned ID. Bill and amendment drafts are saved (even if the bill dies in the committee or house) and are confidential.

Engrossed Bills

The ROS will engross the bill (add amendments) and return it to the House. It is considered an engrossed bill until it is passed in both houses. XTEND has an auto-engrossment feature.

[46] Ibid.

Conference Committee Reports
The third reading of the bill occurs after the ROS adds any amendments to the bill and the House has finished debating. If the bill passes, it goes to the second house where it must be passed in identical form. If both houses do not agree on the content, a conference committee may be appointed to resolve differences. Members of both houses discuss and suggest amendments to the bill. A conference committee report is returned to both houses, and the bill is voted on again. The ROS is responsible for creating the committee report, which is generated from the XTEND system.

Bill Side-by-Sides and Comparison Reports
The ROS also produces side-by-sides, or comparison reports, if either house requests them so that both houses can compare bills, thus facilitating conference committee work. Comparison reports are also requested by the House or Senate desks to compare House and Senate companion bills. These documents are created in the XTEND system as working documents with no life beyond their session.

Enrollments
Once a bill passes in both houses, it is assigned a chapter number in addition to its House or Senate file number. The ROS is responsible for the final version of the bill. Once it passes, it becomes a session law unless the governor vetoes it as a whole or does a line-item veto (in which case the ROS edits the bill).

Session Laws
The final edited session laws (not the final enrolled laws) are the top priority for the ROS in this project. The ROS has session laws for each year since 1994 in electronic format (e.g., HTML and XML). Sessions laws are printed as the *Laws of Minnesota* by the ROS at the end of each legislative session as the official version.

Minnesota Statutes
Statutes are a compilation of permanent and general state laws. The official versions of *Minnesota Statutes* are published in printed form by the ROS every even-numbered year. Electronic copies are available in XHTML, which are updated annually.

Indices and Tables
The software application Cindex, from Indexing Research, is used to create the statute and session law indexes in ASCII, which is converted to HTML. Various statute, session law, and administrative rule tables are created in XTEND.

Administrative Rules
The legislature grants permission to state agencies to form laws that govern the agency's organization and procedures. These are not actually passed by the legislature, but have the same authority as a law. The official version, *Minnesota Rules* is published in printed form by the ROS every odd-numbered year. The chapters are organized alphabetically by agency or department name. Administrative rules reside electronically in the XTEND system, and are put online weekly with frequent updates.

House and Senate Journals

Journals are the official, legal record of the history of the bill. Created by House and Senate desk staff, they record details of each day of the session, including data pulled out of the ROS's system. The House produces its journals in Microsoft Word, while the Senate creates its journals in XTEND. The ROS does not have administrative control of the journals.

Appendix F:
Other Minnesota Legislative Documents[47]

1. Minnesota Senate Documents

Summary

Bob Horton, Shawn Rounds, and Elizabeth Lighthipe, Minnesota Historical Society, met with Robbie LaFleur, Legislative Reference Library; Peter Wattson, Senate Counsel, Research and Fiscal Analysis; Steve Senyk, Senate Media Services; and Jim Greenwalt, Linda Schmitt, and Charlie Fastner from Senate Information Systems, on 13 October 2005. Bob Horton began the meeting by introducing Senate staff to the project and answering any questions. Discussion about legislative documents that the Senate staff creates and/or manages followed.

Document Types

Senate Journals

The Journal of the Senate is a record of Senate floor proceedings. Daily session journals are available on the web following adjournment each session day, on the same day of session adjournment or the following day[48]. At the end of each calendar year, the contents of the daily session journals are compiled to create a published collection of books referred to as the permanent journal, the official, legal record of all floor action. Senate journals from 1996 to the present are available online in PDF (prior to 1996, the Senate web site was not available). During the final editing process following each session, corrections may be made to the daily journals. Beginning in 2005, the edited daily journals are placed online, and an index is attached for each day. From 1996-2004, the online daily journals may not completely match the final, official, edited permanent journal

Amendments

Amendments considered in committee are created in the Revisor of Statutes' system, but are not part of any public file. Some of the more important amendments are posted online in HTML or PDF. The list of amendments is available in a section of the Senate site labeled "Select Committee Amendments," arranged by committee, or on the individual committees' pages. However, this process is not uniform or comprehensive.

Floor amendments are also created by the Revisor of Statutes in XML and are converted by Senate staff into HTML and posted on the Web. These are not archived because they become part of the published journal.

[47] All notes regarding Senate, House, and Joint Commissions documents are current only as of 2005 when the interviews took place.
[48] The Minnesota Senate is online at www.senate.leg.state.mn.us.

Bill Summaries

Senate Counsel and Research staff members create bill summaries when requested, usually for bills that are being heard in committee. When the Senate receives a new version of the summary, the updated version replaces the old one on the Web. Bill summaries dating from 1997 are online in HTML.

Committee Minutes/Books

Paper copies of the committee minutes/books are transferred to the Legislative Reference Library. The Senate committee minutes/books are not published electronically. The Senate would eventually like to create electronic versions and set up some kind of system to archive these. They would archive the final version since committee books are often edited (e.g., agenda is usually rewritten).

Fiscal Notes, Fiscal Tracking Documents, Spreadsheets

Fiscal notes are created and maintained by the Department of Finance. They identify possible monetary impacts of a proposed law. Senate fiscal analysts create fiscal tracking documents in Excel for omnibus appropriations bills. These are kept in spreadsheets on the Senate Counsel, Research, and Fiscal Analysis web page and also published in PDF.

Senate Media Coverage

The Senate records committee hearings, floor sessions, and other meetings held in Senate hearing rooms as audio MP3 files. Selected coverage in video is available as RealVideo files. Available online are select committees, floor sessions, press conferences, special events, Capitol Reports, and civic education beginning in 2001. A few additional civic education videos produced in the 1990s are also included. The digital files of committee hearings and floor sessions are kept one calendar year on servers at the Department of Administration's Intertech group and are then transferred to the Legislative Reference Library for long-term retention and access. Previously, when committee hearings and floor sessions were only available in audiotape, the Legislative Reference Library kept the files for eight years and then sent them to the Minnesota Historical Society. The Minnesota Historical Society destroyed the tapes after an additional eight years, per approved records retention schedule.

Senate Schedules

Weekly schedules since 2000 are available online in HTML.

Photographs

The Senate takes candid photos and legislators portraits. These digital images are in JPEG format and date from 2001.

Senate Briefly

Senate Briefly is a weekly newsletter that summarizes Senate committee and floor proceedings. Paper copies are stored at the Legislative Reference Library. The 1995-1996 issues of *Briefly* are posted online in text file format. The 1997-2002 issues are posted online as PDF files. The 2003 issues are available online as PDF and HTML. Issues of *Briefly* from 2004-2005 are available online as PDF and HTML and include JPEG images.

Rules
The 1994-2005 Permanent Rules of the Senate and the Joint Rules from 2005 are available online in HTML. Current rules are always available online.

Senate Executive Appointments
Senate confirmations of executive appointments dating from 1999 are available online in HTML.

Biographies of Senators and Staff Lists
Biographies of current senators are maintained on the Senate web site in HTML. Current staff contact lists are maintained, with entries created or deleted as staff changes.

Official Directory (Red Book)
The House Public Information Services Office produces a print legislative directory. Electronic versions are not available. The Senate IT group does not want to issue the Red Book in PDF because of format issues.

2. Minnesota House of Representatives Documents

Summary

Bob Horton, Shawn Rounds, and Elizabeth Lighthipe, Minnesota Historical Society, met with Robbie LaFleur, Legislative Reference Library, Mike Speiker, Brenda Van Dyck, Pat Murphy, Don Crosby, Al Mathiowetz, Barry LaGrave, House of Representatives, and Greg Hubinger, Legislative Coordinating Commission, on 3 November 2005. The meeting began by introducing House staff to the project and answering any questions. Discussion about legislative documents that the house staff creates and/or manages followed.

Document Types

House Journal
The House Journal is an important publication that is created daily. It records the actions of the Legislature in session for that particular day and is considered, in its final printed and bound form, the official, legal record of all floor action. Before the final, published journals are created, the Chief Clerk must approve any changes made to the journal. Working versions of the journal are posted online at the end of each day.[49] Once the final version is approved, it replaces all other versions.

The House Journals are available online beginning with the 1994 Legislative session. Those from 1994-2005 were created using Word and then converted to HTML. From 2005 to the present, the journals are available in HTML and PDF. Journals prior to1994 were created using WordPerfect. It is important to note that the House Journal is tied to the bill status system.

[49] The Minnesota House of Representatives is online at www.house.leg.state.mn.us.

Bill Summaries
Bill summaries are created by the House Research Department. Summaries are generally prepared for bills heard in public hearings. The bill summaries from 1998 are in PDF format. Summaries created since 1999 are posted online once they are converted from Word to HTML. The House maintains each version of House bill summaries and offers all of them online, at this point indefinitely. Bill summaries that become laws are never taken off the Web.

Bill Status System
The Bill Status system is an Oracle database that goes back to 1993-1994 as part of the Revisor of Statutes' system. It is also used by the Senate.

Act Summaries
Act summaries are created by the House Research Department to give the public and legislators a basic overview of bills that became laws (chapters). Act summaries since the 1999 legislative session are available online in HTML.

Committee Minutes/Books
A secretary submits committee minutes two days after the meeting. These are in a SQL database in text format without agendas or attachments. They are dynamically converted to HTML and posted online. Much of the material created in committee is working documents and not something the House wants to archive. Hardcopy committee notebooks with attachments are also created.

Fiscal notes, Fiscal Tracking Documents, Spreadsheets, Publications
Fiscal notes, fiscal tracking documents, and spreadsheets are done by the House's Fiscal Analysis Department and are posted online as PDFs on the House website. The Fiscal Analysis Department also offers the publications entitled, *Money Matters* and *Issues Briefs*, which summarize fiscal issues and topics related to the Legislature.

House Media Coverage
Selected committee meetings and floor sessions are recorded and offered online to the public. Digital video content is offered as Windows Media Player files (56 kilobytes per second quality) back to 1999. Metadata is added to the audio/video files that are converted to digital format. The House also started producing MP3 audio files in 2005.

New Laws
At the end of every session the House publishes *New Laws*, which summarizes legislative activity including new laws, vetoed bills, and other items of interest that session. The House Public Information Services Department posts the issues online as database-generated PDF files. Issues dating from 1993 are available.

Session Weekly
Session Weekly is a newsletter that summarizes House committee and floor action for each week during session. Issues since 1995 are offered online as PDF files, with 1994 issues of *Session Weekly* available as text files.

Session Daily
Session Daily issues are posted online as HTML files, which offer dynamic links to committee information, bills, video and audio files, legislators, and other specific related sites.

RSS Feeds
The house offers RSS feeds for the *Session Daily*, House standing committee audio, and bills.

3. Joint Departments and Commissions Documents

Summary

Shawn Rounds and Elizabeth Lighthipe, Minnesota Historical Society, met with Robbie LaFleur, Legislative Reference Library (LRL), Michele Timmons, Office of the Revisor of Statutes, Greg Hubinger, Legislative Coordinating Commission, Lee Meilleur, Geographic Information Services, Michelle Pryce, Office on the Economic Status of Women, Denice Malone, Office of the Legislative Auditor, Lisa Diesslin, Legislative Commission on Pensions and Retirement, Susan Thornton and John Velin, Legislative Commission on Minnesota Resources.

Michele Timmons, the Revisor of Statutes, is mainly interested in using the SDSC's SRB system for disaster recovery purposes. Michele would also like to identify data in print that could be digitized for archiving and access (e.g., scanning the session law books since many of them are deteriorating and will not last forever).

One concern brought up numerous times during the meeting was the differences in perspective on what electronic and paper records to archive. The State Archives is interested in preserving documents with historical value, while some agencies would rather save everything and connect it all. Another issue brought up was the longevity and trustworthiness of electronic records.

Document Types

Legislative Commissions on Pensions and Retirement
The Legislative Commissions on Pensions and Retirement create a background document on each bill in Word. Depending on the document, these are posted online in either PDF or HTML format. Beginning in 2005, committee hearings are recorded in MP3 format and are also available online.[50]

There is no records retention schedule. Electronic documents from the early to mid 1980s are irretrievable because of access issues with format. Staff also believes there are files on floppy discs from the 1980s. The LRL has reports mandated by the Pensions Commission; most are in Word, with a few are PowerPoint and Excel.

[50] Joint Departments and Commissions web sites are accessible through www.commissions.leg.state.mn.us/depts.htm

Office of the Legislative Auditor
Most of the office's electronic records are in HTML, PDF, or Word. They are moving towards Linux, but are having issues with Access databases.

Geographic Information Services
Staff consider map and redistricting data the most important, including ESRI and caliber files. The geospatial community is considering using GML, which is a subset of XML, as format for ongoing use (converting to it). Maps are created in PDF for printing. The department also uses Word and HTML files and different databases (dbf, access, SQL server). They do not use digital audio files.

Legislative Commission on Minnesota Resources
Minutes of meetings and supplementary materials are online in HTML and PDF. They try to post everything, including abstracts and schedules. Summaries of past projects are kept online, along with files for the current two-year cycle. The LCMR also creates hard copy final reports, which include electronic files like maps. These are transferred to the LRL in a 2-year cycle, with some also going to the State Archives. Meetings are recorded in digital audio as MP3 files. The LCMR has an Access database that contains internal reports, which may be of historical value.

Legislative Coordinating Commission
Staff stated that there are not many records created that they consider unique from any other documents already mentioned, beyond documents from support task forces, which include reports and minutes of these groups. Most documents are created in Word and posted online in PDF. Digital audio is in MP3 format at present, with prior audio records on tape.

Office on the Economic Status of Women
This is no longer a commission. They create documents in Word, PageMaker, PDF, and HTML. They post only current information, such as fact sheets, and session summaries online.

Appendix G:
MHS-ROS Preservation Workflow

Figure 1 shows the basic workflow between the Office of the Revisor of Statutes (ROS), the San Diego Supercomputer Center (SDSC), and the Minnesota Historical Society (MHS). It is a generic workflow for acquiring and accessioning records that can be used with any file/storage management system. It is designed to satisfy both disaster recovery and archival requirements. Figure 1 represents a first iteration of the workflow, which can be much refined and revised through hands-on tests over time.

Briefly, ROS files are moved from the XTEND repository to SDSC using the SRB software. ROS staff handle the transfer since this portion is completely a disaster recovery operation, done to their specifications for integrity checks, and using their directory and naming conventions. Once the files are at SDSC, MHS staff appraise the records and replicate those chosen into a deep-storage archival holding area, where MHS staff can add additional metadata, enhance descriptive directory and file names, if necessary, and accession records into a collection. [51]

This process has two advantages. The first is project-specific, in that it overcomes MHS' bandwidth issue whereby the technical infrastructure of the Society is not sufficient to handle the requirements of large online data transfers.[52] The ROS has a much more robust network infrastructure, so it makes sense for staff there to handle the file transfer, taking full advantage of the SRB's capabilities. Second, and most important, MHS staff think it desirable to handle ROS digital files as little as possible once they leave the XTEND system. This is good practice from both archival and disaster recovery perspectives as the more times a digital file is transferred, copied and/or manipulated in some way, the more chances for errors to be introduced.

The workflow also parallels the MHS acquisition process for other paper and electronic government records. When MHS staff work with agencies to collect their records, they typically do not go into agency systems and offices to pull out the records they want. Instead, the records are transferred to MHS by the agency at the point when the agency deems the records to be final. The workflow makes further sense from the ROS' point of view. For this project, the ROS wanted to test this system for disaster recovery use. The ROS wanted internal staff handling the transfer operations from security and quality assurance.

The cooperative aspect of this workflow is important to note. It provides benefits to both the ROS and MHS with minimal extra effort on the part of ROS staff. Storage for disaster recovery and for archival preservation have aspects in common which make such a partnership appealing.

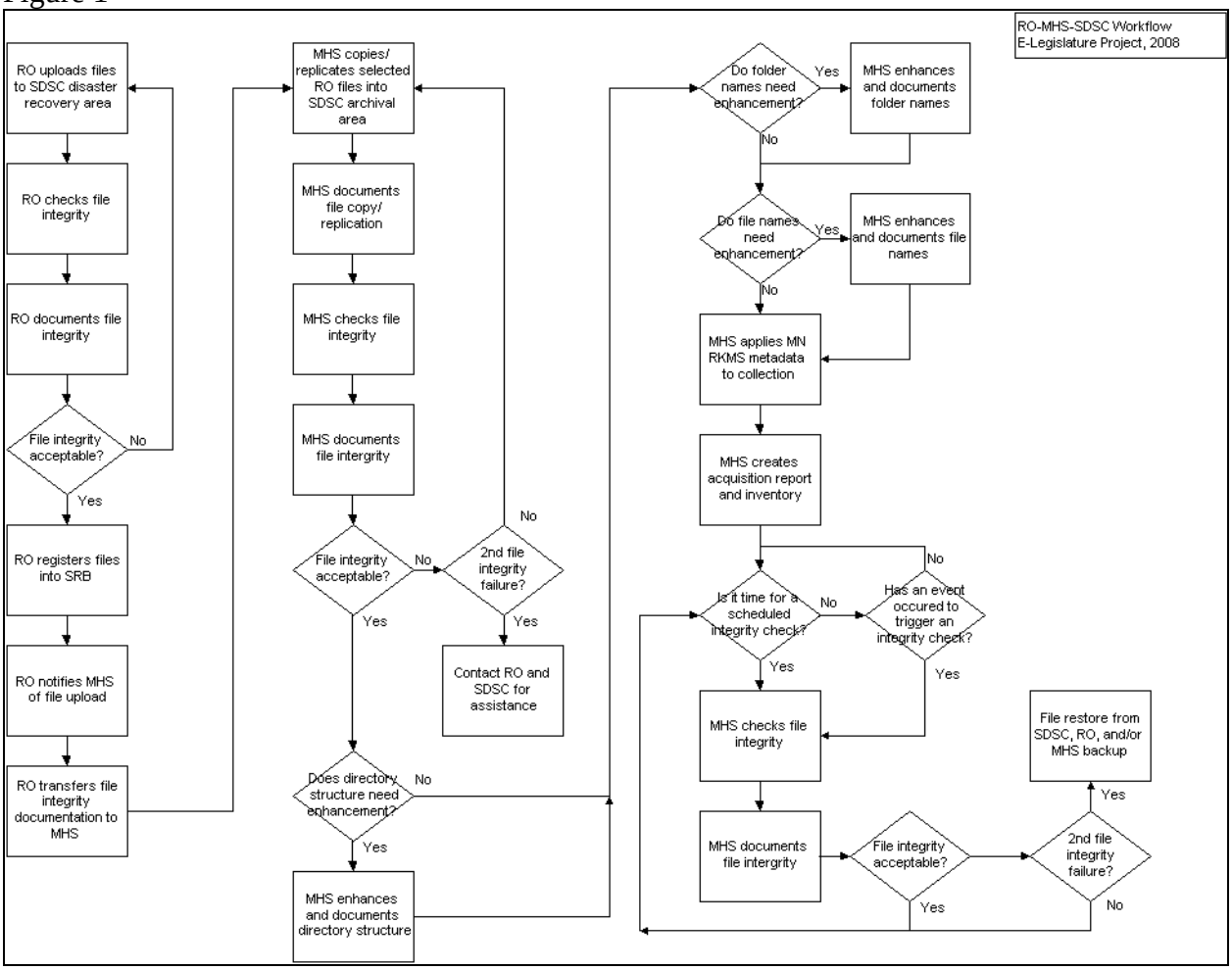
---

[51] The internal ROS repository uses a directory structure and naming scheme that follows the structures of particular document types. For example, the first section of Chapter 1 of the 2007 statutes would be represented by "Stat/2007/1/1.xml". The first sub-part of the first section would be "Stat/2007/1/1A/1.xml". The result is that there are several files within the repository with the same file name, distinguished from one another only by there directory location. Along with documenting the original files names and directory structure, MHS staff would create descriptive, standalone file names according to a standardized naming convention.

[52] As discussed in "Progession of Work," Section 2 of this report.

Most notably, both require high-security, limited-access environments and both hinge on establishing and maintaining file integrity over time.

Figure 1



Open Issues to Finalized in Production:

With respect to the ROS' disaster recovery process:

- Establish routine schedule for transfers
- Is the transfer process a complete over-write each time? If not, how many backup iterations are maintained?
- Establish routine file/transfer integrity check process. How will integrity checks be documented? Where is the documentation stored?
- Establish routine process for notifying MHS when archival files are transferred. What standard information is shared about the files and the transfer?

With respect to the MHS' archival storage process:

- How is the replication process documented? Where is the documentation stored?

- Establish documented naming conventions for enhanced directory/folder/file names. How are enhancements documented? Where is documentation stored?
- Which elements of the Minnesota Recordkeeping Metadata Standard are applied to collections? What is the format used for the metadata? Where is the metadata stored?
- Is a copy of the acquisition report and file inventory stored with the collection? What is the format of reports and inventories?
- Establish schedule for integrity checks of collections. Establish routine integrity check process. Is the integrity check process manual or automatic? How are integrity checks documented? Where is documentation stored?

General issues:

- What are the roles/functions within the processes? What are the allowable actions for each role, including editing/deletion rights to documentation? What are the corresponding security levels? What is the process for changing and adding roles/functions?
- Are there other trigger points or events for integrity checks? What is the process?
- How are access files created from master files? What is the documentation process? Where are access copies stored? What triggers creation of new access copies?
- What is the policy/procedure for deaccessioning and deletion? What is the documentation process?
- What is the policy/procedure for migration and conversion? What is the documentation process? What is the trigger?
- What about audit trails for actions/access to files (to be documented within metadata set)?
- What documentation is shared with ROS and when?
- What is SDSC's system maintenance and availability responsibility (to be documented in a service agreement)? What happens if there is a problem on the SDSC side?
- What processes can be automated?

Appendix H:
ROS File Transfer Tests[53]

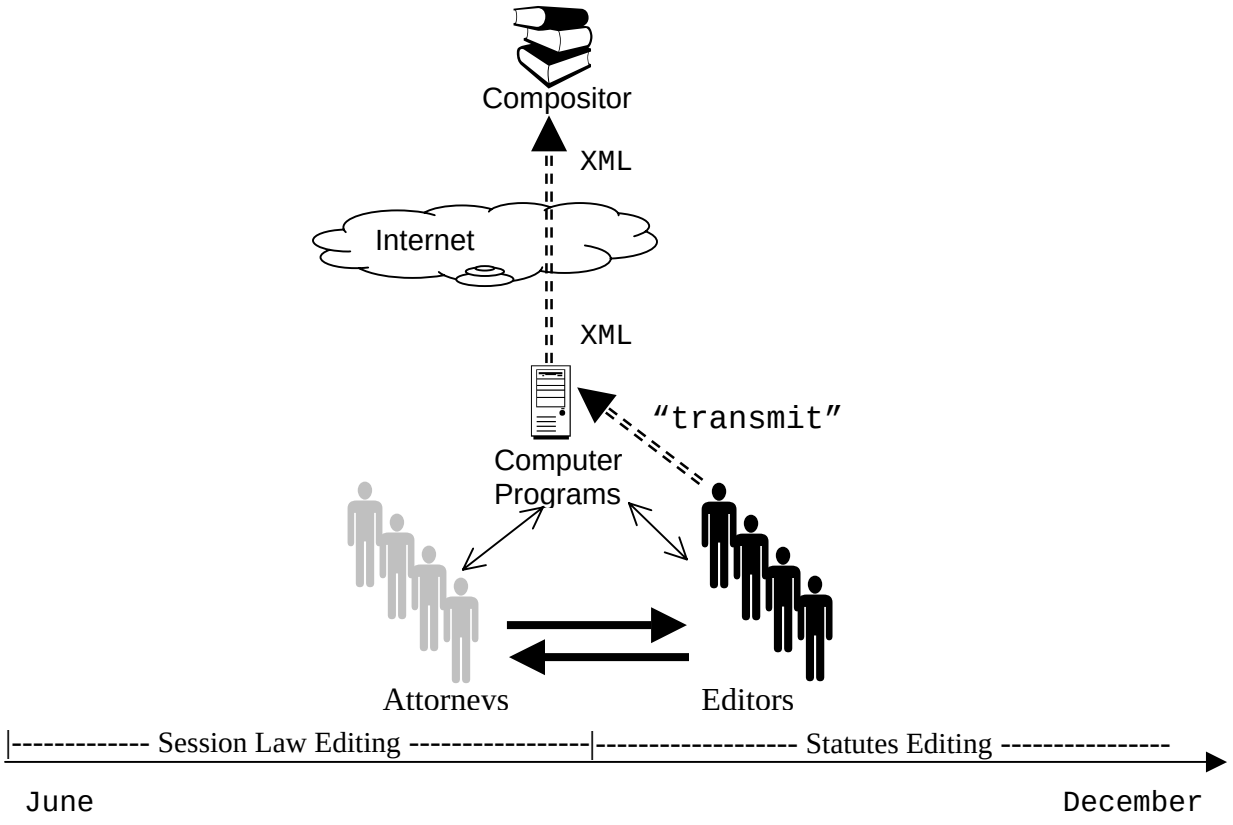
Publication Workflow

The Minnesota Office of the Revisor of Statutes (ROS) produces 24 document types. Multiple, interdependent workflows are used to create, edit, and publish these documents. A description of each workflow is beyond the scope of this paper. However, the workflow for Minnesota Session Laws and Statutes is relevant because: a) these documents were used by this project for data transfer testing; b) these are two of the primary candidates for data transfer in an operational archival system.

Current Workflow

The Session Law and Statutes workflows are sequential. Figure 1 shows the workflows. After each legislative session, staff attorneys and editors review the text of the enrolled acts implementing the Revisor's statutory authority to correct clerical errors. Computer programs are used to facilitate the checking of the integrity of the statutory language in Laws. As the processing and quality control inspections for a chapter are complete, the chapter is transmitted via the Internet to the compositor (composition contractor). This process is known as session law editing and takes from one to two months depending on the number and complexity of the session law chapters.

Figure 1. Session Law and Statutes Workflows

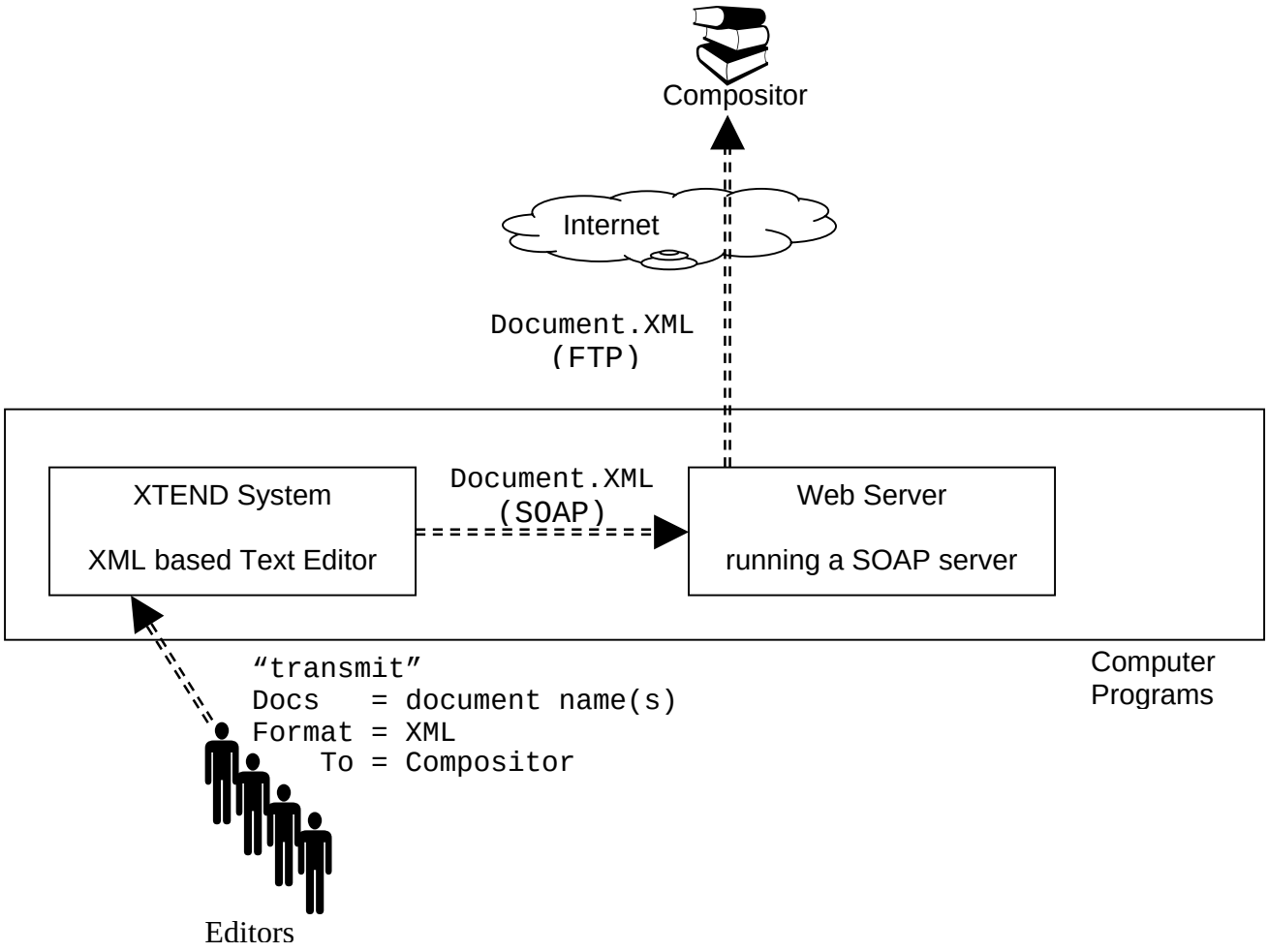


[53] This summary was created by ROS staff; statistics have not been verified.

When session law editing is complete, statutes editing begins. After each legislative session in even-numbered years, the office compiles and edits the complete, hardbound edition of Minnesota Statutes. After each legislative session in odd-numbered years, the office compiles and edits a supplement to the statutes. General and permanent laws added or amended during the session are extracted from the session laws and edited for inclusion in a new edition of the statutes. The repeal of statutory text is reported, multiple amendments are integrated, conflicting amendments are reported, and changes to the text directed by Revisor's instructions in bills are implemented. The publication includes section histories, tables, an index, and other editorial aids. As the processing and quality control inspections for a statutes chapter are complete, all of the sections in the chapter are transmitted via the Internet to the compositor. Statutes editing takes from two to three months depending on the number and complexity of the statutory changes.

Figure 2 shows the software process flow for transmitting documents to the compositor. When attorneys and editors have finished modifying documents, a Legal Editor supervisor sends the "transmit XML to compositor" command to the XTEND bill drafting system. XTEND retrieves the specified document(s) for the XML document repository then sends the document to a web service running on the Revisor's web server. The web service uses FTP to put the document on the compositor's FTP server.

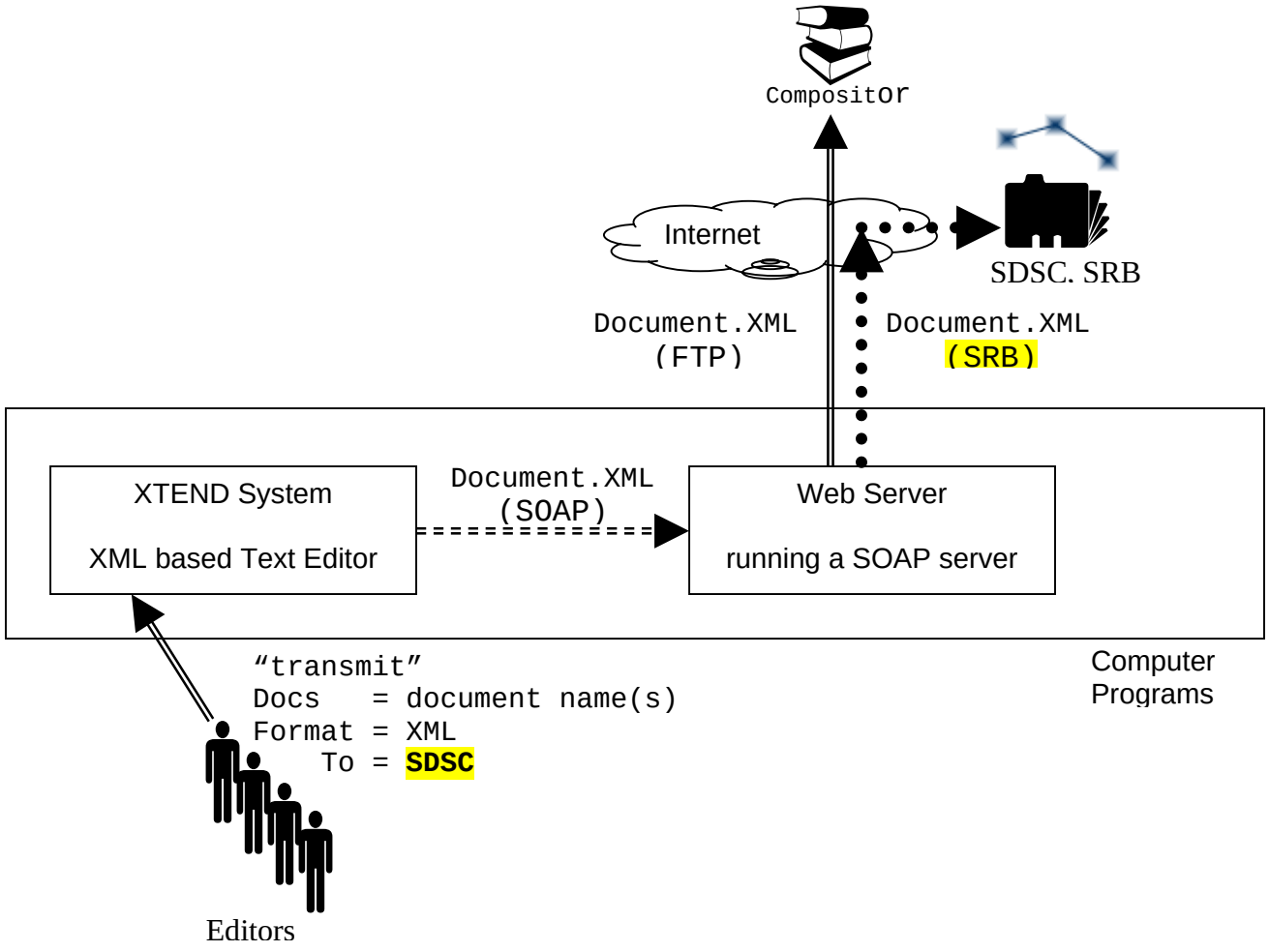
Figure 2. Document transmission to compositor



Future Workflow with Backup/Archive to SDSC

The existing workflow can be easily augmented so that documents can be transmitted to other destinations. Figure 3 shows a potential workflow that could be used to send documents to the SDSC's Storage Resource Broker (SRB). Two simple changes would be needed to the current workflow. First, a new user option would be added to XTEND. The new option would allow users to select the SDSC as the destination. Second, the web service running on the web server could be modified to execute SRB Scommands. These commands would use the SRB protocol to transmit documents to persistent archives in the SRB system.

Figure 3. A candidate workflow to transmit documents to SDSC



Data Transfer Test Environment

A grid brick server was used to conduct document transmission testing. Figure 4 shows testing environment which consists of an Intel server running the Red Hat (Linux) operating system and using the SRB Scommand software.

Four automated tests were run. Addendum A contains listings for each script showing the commands executed and the results. Each automated test consisted of 3 parts:

① a cron file entry specifying the time(s) of day a test script should be run
② a test script containing the Scommands to execute and the data set to transmit
③ a specific data set consisting of:
- Minnesota Session Laws
- Minnesota Statutes
- a single, binary data file in .ISO format

Table 1 contains a summary of each script.

Figure 4. E-Legislature Project, Testing Environment

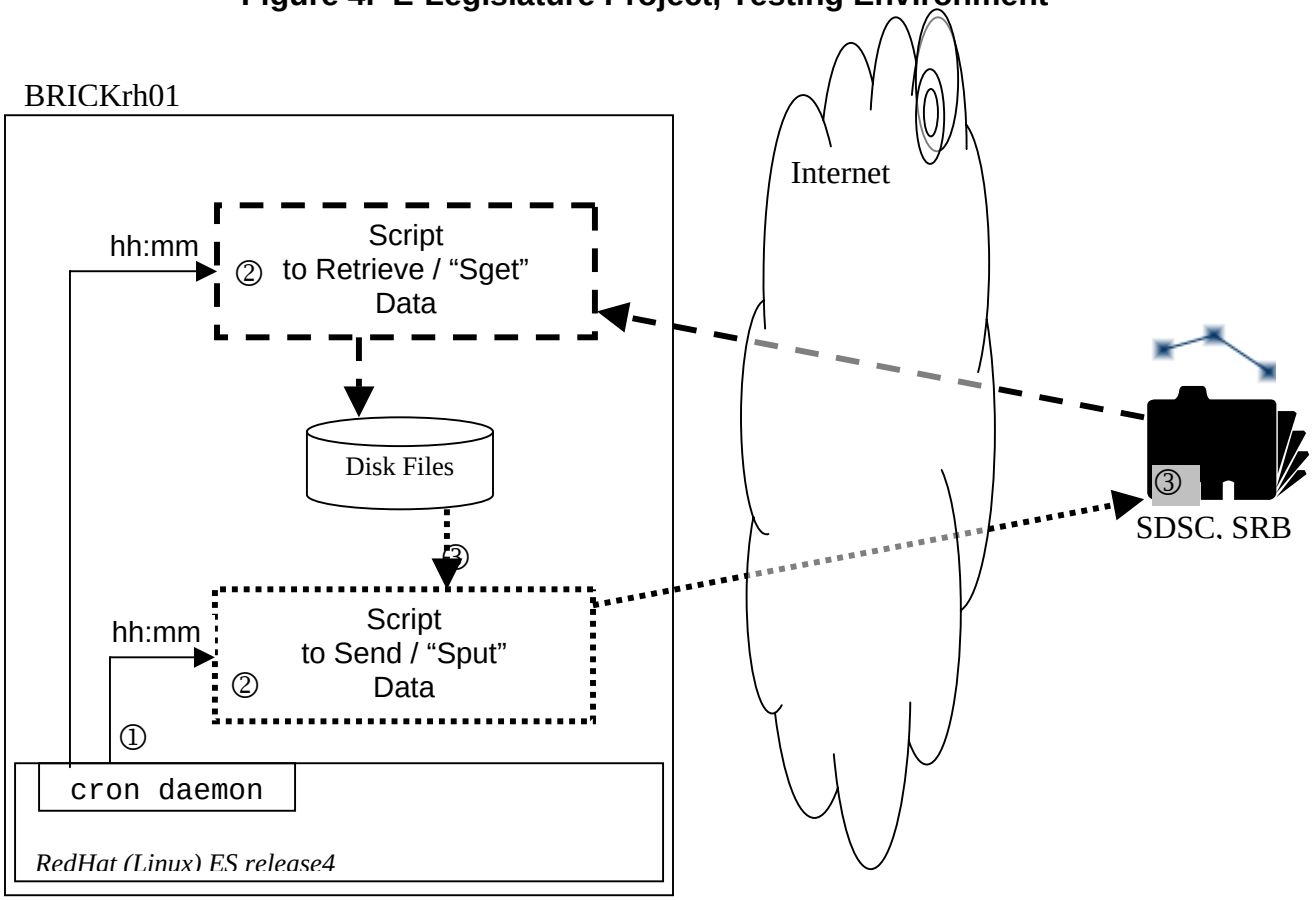


Table 1. Script Summaries.

Script Name	Execution Dates	Frequency	Scommand	Data Type	Transmitted
srbtest.sh	2006-05-09 to 2007-07-17	1 hour	Sput	executable, binary file	2.7 MB
srbtestbig.sh	2006-05-12 to 2007-07-17	1 hour	Sput	ISO 9600 CD binary file	687 MB
getsrb.sh	2007-03-06 to 2007-07-17	30 minutes	Sget	ISO 9600 CD binary file	687 MB
dataput.sh	2007-04-16 to 2007-07-17	3 times per week	Sput	2006 Session Laws, and Statutes 3 File Formats: .PDF .XHTML .XML	360 MB

Addendum A. Data Transfer Test Results

Script srbtest.sh

srbtest.sh 2006.05.09 - 2007.07.17 4924 observations

```
1  /home/srb/SRB3_4_1/utilities/bin/Sinit -v
2  /home/srb/SRB3_4_1/utilities/bin/Sinit -v
3  /home/srb/SRB3_4_1/utilities/bin/Srm /PatZone/home/srbAdmin.eleg/Sget
4  /home/srb/SRB3_4_1/utilities/bin/Sput -VKMS sfs-disk-pat
/home/srb/SRB3_3_1/utilities/bin/Sget >> /scripts/srbtest.txt
```

	MB	MB/s	s	Date	Time	chksum
Min	2.741	0.063	10.3			50284
Max	2.741	0.266	43.53			50284
Average	2.741	0.1645	26.915			
Median	2.741	0.1645	26.915			
Mode	2.741	#N/A	#N/A			
Std Dev.	0	0.143542677	23.49715834			

Script srbtestbig.sh

srbtestbig.sh 2006.05.12 to 2007.07.17 4768 observations

```
1  /home/srb/SRB3_4_1/utilities/bin/Sinit -v
2  /home/srb/SRB3_4_1/utilities/bin/Sinit -v
3  /home/srb/SRB3_4_1/utilities/bin/Srm /PatZone/home/srbAdmin.eleg/fc5.iso
4  /home/srb/SRB3_4_1/utilities/bin/Sput -VKMS sfs-disk-pat /data/fc5.iso >>
/scripts/srbtestbig.txt
```

	MB	MB/s	s	Date	Time	chksum
Min	687.235	0.191	93.26			8665
Max	687.235	7.369	3603.46			8665
Average	687.235	2.311052	808.282			
Median	687.235	2.632	261.075			
Mode	687.235	0.479	1408.98			
Std Dev.	0.000119	1.617901	896.5397			

Script getsrb.sh

getsrb.sh 2007.03.06 to 2007.07.17 1897 observations

```
1  rm -f /scripts/xxx.iso
2  /home/srb/SRB3_4_1/utilities/bin/Sinit -v
3  /home/srb/SRB3_4_1/utilities/bin/Sinit -v
4  /home/srb/SRB3_4_1/utilities/bin/Sget -VkM /PatZone/home/srbAdmin.eleg/fc5.iso
/scripts/xxx.iso >> /scripts/getsrb.txt
```

	MB	MB/s	s	Date	Time	chksum
Min	687.235	0.46	169.93			8665
Max	687.235	4.044	1493.46			8665
Average	687.235	3.087852	232.634			
Median	687.235	3.125	219.89			
Mode	687.235	3.048	193.59			
Std Dev.	0	0.488257	78.69618			

Script *dataput.sh*

```
dataput.sh      2007.04.16 to 2007.07.17   ?571,444 observations?

/home/srb/SRB3_4_1/utilities/bin/Sinit -v
/home/srb/SRB3_4_1/utilities/bin/Sinit -v
/home/srb/SRB3_4_1/utilities/bin/Srm -rf /PatZone/home/srbAdmin.eleg/data
/home/srb/SRB3_4_1/utilities/bin/Sput -rVKMS sfs-disk-pat /scripts/data >>
/scripts/dataput.txt
echo "@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@"  >> /scripts/dataput.txt



#/home/srb/SRB3_4_1/utilities/bin/Srm /PatZone/home/srbAdmin.eleg/Sget
#/home/srb/SRB3_4_1/utilities/bin/Sput -VKMS sfs-disk-pat
/home/srb/SRB3_3_1/utilities/bin/Sget >> /scripts/srbtest.txt
```

Script Execution Frequency	3 times per week
Data Analysis Dates	2007.04.16 - 2007.07.17
Data Directory Structure	2006_SessLaws/ PDF / 114 files XHTML / 114 files XML / 9 files -------------- 237 files 2006_Statutes/ XHTML/ 137 directories/ 46,349 files XML/ 0 directories 1,037 files --------------- - 47,386 files
Total Data Size (bytes) **du –bs / du –s --block-size=1**	237,992,313 / 377,417,728 bytes

Appendix I:
Analysis Recommendations

Successful, sustainable digital record preservation programs are the result of careful analysis and planning that is based on what is practical and acceptable rather than what is ideal but unattainable. Several factors must be considered in terms of content, infrastructure, and functionality.

Content considerations include:

- What record types are of interest?
- Who are the record creators and/or owners?
- What are the associated data formats?
- What are the retention requirements?
- What are the access/security/privacy requirements?
- What metadata exists for the records and/or record sets?
- Are the records indexed?
- Are the records compressed?
- Are the records encrypted?
- Are the records associated with digital signatures?
- What other information resources or record sets are related to the records?
- What new record types, formats, etc. may be added in the future?

Infrastructure considerations include:

- Where will the records be stored (e.g., off-line, near-line, on-line)?
- Who owns the storage system?
- Who maintains the storage system?
- What is the availability of the storage system for access?
- What is the back-up plan for the storage system?
- Is there a disaster recovery plan for the storage system?
- How will records physically be transferred into/from the storage system? Is there a potential bandwidth issue if the files are large or numerous?
- Projecting records quantities, when will the storage system reach capacity? What is the plan for acquiring more storage space (e.g., acquire another system, acquire additional storage space for existing system?)
- What is the staff time necessary to manage the system? What skills are needed?
- What is the staff time necessary to manage the records? What skills are needed?
- If necessary, how will record compression and/or encryption be handled over time?
- If necessary, how will digital signatures be handled over time?
- What is the plan for migrating the records to a new storage system and/or operating system?
- When data formats need to be converted to maintain accessibility over time, what will that process entail?

- What is the process for quality assurance tests carried out to ensure record integrity? What are the triggering events (e.g., record additions, system migration, record conversion, routine periodic tests)?
- Should there be a Service Level Agreement detailing roles and responsibilities with respect to the preservation infrastructure?
- What level of documentation should be maintained to describe the architecture of the storage system and management applications, as well as the supporting management processes?

Functionality considerations include:

- What levels of trustworthiness/authentication must be maintained for the records (for the archives, the record users, others with an interest)? At what levels (e.g., individual records, record set)?
- To support appropriate authentication level, how will access and security requirements be addressed?
- What level of detail is necessary for audit trails?
- What is the process for deleting records/record sets?
- Is indexing required to facilitate discovery and use?
- Does additional metadata need to be created to facilitate discovery/access, use, and/or management? At what level (e.g., individual record, record set)?
- How is metadata connected with the records/record sets (intrinsic, external/associated)?
- Where is documentation relating to the records and management actions stored?
- Is it necessary to indicate how the records/records sets relate to other information resources or records within and/or outside the system? If so, how should such linkages be described and maintained over time?

The point of such detailed analysis is to understand the current environment and describe the goals for preservation so that the gap between the two can be defined. Cost-benefit and risk analysis methods should be applied to prioritize goals and allocate resources.[54] At those points where the gap is insurmountable with available resources, partnerships/collaborations with other interested parties, such as records creators, should be considered as long as such an arrangement can be sustained over the long term.

[54] See additional discussion in Section 2 of this report, "Progression of Work: Cost Benefit Analysis."