

# Preserving State Government Digital Information Illinois Partners Meeting



## Minnesota Historical Society

Tuesday, February 5, 2008  
Room 309, Gwendolyn Brooks Building, Illinois State Library  
Springfield, Illinois

*Due to inclement weather, representatives from the Minnesota Historical Society (MHS) were unable to meet in person their Illinois partners of the NDIIPP-sponsored project, Preserving State Government Digital Information. However, project director Bob Horton gave a Powerpoint presentation via speakerphone, and the Illinois participants discussed matters amongst themselves afterward.*

*Illinois participants included Tim Rice and Mark Wenda from the Legislative Information System (LIS); Dave Joens, director of the Illinois State Archives (ISA); Suzanne Schriar, Kathy Bloomberg, Anne Craig, Andrew Bullen, Connie Frankenfeld, and Cherryl Walker from the Illinois State Library (ISL); and Larry Jackson, a researcher at the Graduate School of Library and Information Science (GSLIS) at the University of Illinois Urbana/Champaign (UIUC).*

*The following account is an amalgamation of original minutes, questions subsequently proffered by MHS staff, and Illinois personnel's responses to those questions.*

### ISL CURRENT STATUS ON PRESERVING LEGISLATIVE MATERIALS

Tim affirmed that LIS is using XML base bill drafting base at this time.

Andrew stated that ISL has a system that is working. However, ISL staff felt that we could use help making a business case to our own funders (e.g., Secretary of State's office).

Law includes electronic publication in the depository program.

Tim explained that LIS is saving material since the 93<sup>rd</sup> General Assembly (i.e., 2003-2004 legislative session).

LIS retains all the electronic content from its systems—text of bills and amendments, including versions from introduced through enrolled; text of public acts; statutes; history of action on each bill; journals; transcripts of floor action; chamber roll calls/votes; member information, including bio, bills sponsored, and committee membership; committee lists with members; final versions of reports such as the status of all bills, bills per sponsor, and public acts by effective date. The material is retained so as to be available on the General Assembly Web site and is backed up per standard LIS procedures for all data. However, this cannot be considered archiving per se.

LIS does not store audio or video for either chamber. Minutes are not retained electronically.

Use example: Tim shared that people (reporters) who wanted information about Barack Obama's tenure in the state legislature had to go through paper volumes because that information was not saved electronically.

Tim continues to explain what legislative information is available tangibly and electronically.

1. Any recording of committee activity is done internally by the House and Senate without the involvement of LIS. The results of committee action are entered into the LIS bill tracking system, and some items are entered into the journals, but LIS has no access to any minutes that might be taken in committees.
2. LIS streams audio and video of floor proceedings to the Internet in real time only. LIS does not capture or archive these streams in any way per current policy. The House does record the stream and makes it available on DVD. Whether and how they retain that is unknown at this time.
3. LIS currently streams one committee room for the House in real time without recording it. More will likely be added in the future. The Senate as yet has expressed no interest in this capability.
4. Both the House and the Senate use third-party solutions from FTR [For the Record, a digital asset management systems company<sup>1</sup>] to record audio of floor proceedings and produce the floor transcripts via LIS, which are made available electronically on the General Assembly Web site maintained by LIS.
5. Floor transcription began in 1971 (77<sup>th</sup> General Assembly). Paper copies of transcripts from that point through the 91<sup>st</sup> General Assembly have been scanned and made available on the Web site. This would include the material sought by reporters concerning Senator Obama's time in the Illinois State Senate.
6. The House and Senate journals are available on the Web site from the 91<sup>st</sup> General Assembly forward.

As to an evaluation of the cost/benefit for including electronic versions of this material, it may be helpful, although decisions on what is to be captured, retained, and made available, how that is to be done, and in what forms, is as much or more a policy decision than a financial one.

#### AUDIENCES

ISL and ISA agreed that their primary audience is State Government (and LIS, the General Assembly), followed by libraries (public, academic, special, law, and school libraries), historians, and the general public.

#### MISSION

All agreed to the value of the material and that ISL and ISA have a mandate to preserve it. Currently, LIS displays General Assemblies back just to the 93<sup>rd</sup> General Assembly. They store backup tapes in a rented space in the ISA building. Some of the data is not published, and LIS is not allowed to disseminate it.

---

<sup>1</sup> <http://www.fortherecord.com/page.asp?PageID=11>

Tim elaborated that, as noted above, all electronic data is currently backed up. Some data from previous systems (i.e., mainframe-based) is on tapes stored in that location, but may no longer be accessible due to the obsolescence of the creating system. In that case, hard copy is stored as well, but it's unclear whether it is all-inclusive.

As for non-dissemination, certain information on LIS systems is restricted by policy in means of access and format. There are also some confidentiality requirements, but these are not extensive quantities of data.

The published data is printed and sent to ISA for archiving after each biennial legislative session. LIS intends to maintain online availability of all pertinent data from the 93<sup>rd</sup> General Assembly forward.

ISL is trying to spider the LIS Web site for the ISL search engine and dark archive. (The dark archive consists of files harvested from all the State of Illinois Web sites, but it's not publicly available because we haven't found a user-friendly way to present the multi-terabyte material.) The spider gets into loops and doesn't seem to copy everything. In order to facilitate archiving, Larry has experimented with picking up material from the LIS FTP site and automating the metadata, but that has also been problematic.

What data do we need and how frequently do we need it?

Dave Joens will share a list of legislative materials that ISA is required to archive. ISL will compare that list to the items available from the LIS FTP site. Then, ISL can define exactly what should be harvested from the LIS Web site.

#### ILLINOIS PRIORITIES FROM THE GRANT PROJECT

- A persuasive business model/case to present to funders.  
Primary funders for ISA and ISL are the budget officials for the IL Secretary of State because they make requests to the General Assembly. The leadership of the House and Senate would control funding to LIS.
- Staff expertise level and number employed in successful programs.
- Rosetta stone for transforming complex data sets to a collectable format.  
An XML schema may help, but it involves coding custom scripts.
- How can you authenticate electronic documents?
- How are scalability issues handled?
- Is there a way to avoid human intervention on every document?
- Things we can share (i.e., growth).

#### INITIATIVES

LIS is not planning any major system changes. There may be some updates or minor enhancements. Currently, LIS provides streaming audio and video, but is not allowed to capture audio or video. If the legislature decides to capture those formats and offer recorded

files from the LIS Web site, it would be a tremendous change in bytes to transfer, store and serve.

As far as the assembled participants know, the House does not retain the audio/video file for long after it has been transferred to DVD (and the Senate does not record anything). Also, no one present knows the retention period for the audio files used by the FTR solution to create the floor proceedings transcripts.

### CAPACITY

ISL has concerns about

- Scalability
- Staff
- Storage space

How do we judge how much we'll need?

### RESOURCES

Again, how much is enough?

- Human Resources
  - The cost?
  - How many FTE?
- Bandwidth
  - The cost?
  - How much is enough if Audio and Video files start coming to the depository?

### ISA LIST (PER DAVE JOENS)

The statute for what records the general assembly is to give to the Archives is pretty vague, stating:

The Secretary of the Senate and Clerk of the House of Representatives, at the close of each session of the General Assembly, shall deliver to the Secretary of State all books, bills, documents and papers in the possession of either branch of the General Assembly, correctly labeled, folded and classified, according to the subject matter of such documents respectively; and the Secretary of State shall preserve the same in his office.

That said, here is what we have that I would recommend for an electronic archives (Tim noted that the paper copies of bills, amendments, resolutions, votes, transcripts, and journals all come from electronic sources):

- House and Senate Bills, Resolutions and Amendments arranged numerically by bill or resolution number within each biennium

- Beginning around 1929, we have committee reports, recommendations, and roll call votes.
- We have transcripts of debates
- House and Senate Journals

In miscellaneous, we have:

- Original minutes rules adopted by the House and Senate
- Communications between the two chambers
- Reports and communications from Constitutional officers, departments, boards, commissions, and committees, as well as circuit court judges and Court of claims and incoming correspondence to each house and the entire General Assembly