Wisconsin Public Library Consortium
Digital Archives Backup Workgroup Notes
September 16, 2024, 10:00 am
via zoom*

ATTENDEES: Jean Anderson (SCLS), Keetra Baker (WLS), Alison Hoffman (MLS), Andrew Hoks (SCLS), Kristie Hauer (WVLS), Joshua Klingbeil (WVLS), Scott Prater (UW-Madison), Tamara Ramski (SCLS), Kristen Whitson (RW/WiLS)

ABSENT: Vicki Teal Lovely (SCLS)
Project Manager: Jennifer Chamberlain (WiLS), Melody Clark (WiLS)

The meeting started: 10:01 am

1. **Review Agenda – Changes or additions**
   There were no additions or changes to the agenda.

2. **Discussion Items**
   a. **Update: Budget and Dell Maintenance Extension**
      At the last meeting, the group decided to proceed with the Dell maintenance extension after discussing the Technology Steering Committee's feedback. Since the meeting, SCLS and LEAN WI staff have moved forward with the Dell Backup Maintenance Extensions.

      Total cost: $177,400.47
      IFLS will cover $133,475 from held LSTA funds: remaining balance $0
      SCLS will cover $43,925.47: remaining balance $66,419.53

      An updated version of the [Five-Year Plan](#) was shared and reviewed.

   b. **Update: Onboarding of Systems**
      [A document was shared](#) with each systems' onboarding update. It was noted that the estimated storage need is 26.25 TB. This may be needed in the next generation discussion. It was shared that K. Whitson did have to estimate for a few library systems; notes for how those estimates were made are in the document.

      J. Klingbeil shared that it is good to see the total storage needs so far, but it would also be good to know how much data is still out there in terms of items that haven't been digitized, etc. It was discussed that the option of preservation-level storage might inspire some systems or organizations to begin digitizing if they haven't already.

      There are a couple of systems, SWLS and Nicolet, that would love to participate, but they don't have digitization programs in place yet. Recollection Wisconsin is doing a digitization training program with SWLS this fall and has a digitization intern placed at Nicolet to design and document digitization workflows.

      In WVLS, there are several libraries that have content to be digitized still. There is yet to be a lot of content that still needs to be identified.

S. Prater noted that one thing that might be useful with libraries that want to digitize materials but do not yet have a workflow in place, is to communicate baseline standards like good formats, information rich documents, files organized for easy duplication into the storage system, etc. This will make it relatively easier to be able to predict the space needed in the future. Recollection Wisconsin is ahead of the game on standards and practices and so are many of the systems. Making sure that those standards and practices are made clear is important and, in terms of this project, there is an extra step or two to make sure the data is ready for storage.

It was noted that it is interesting that Winnefox needs 20TB, by far the lion's share. Winnefox has one library that has well over 100 years of newspapers and three others that have most of their newspapers digitized. In addition, Winnefox does have another library who has initiated the process of digitizing their newspapers with a vendor, so their needs could grow sooner rather than later.

The Historical and Local Digital Collections Committee historically has been a conduit for helping with newspaper digitization projects and served as a link between the Wisconsin Historical Society. The project has been on pause indefinitely. However, the group still exists to work on digitization efforts and that committee is available to help with the communication aspect of sharing standards and practices. This is turning into a community of practice now.

The next generation of hardware will need to be researched and purchased based on storage space needs. S. Prater noted the storage needs may need to include mirroring, so if the current need is 26.25, would the expectation be that double the amount would be the more accurate storage need? It depends on any compressed material, but duplication does need to be considered. Regarding file compression practices, S. Prater noted that compression is not recommended for preserved resources, as it's another technology layer/dependency that can affect recoverability of a resource in the future.

K. Whitson shared that her anecdotal impression is they won't grow this storage need very much in the future, maybe getting up to 50 TB before mirroring and duplication. It was also noted that during the onboarding, compression of data has not been discussed at all.

T. Ramski noted that after her initial upload of 2.3 TB, she has only added another ½ TB of data in a year and assumes that this will be a consistent rate of addition in the future.

It was shared that TB Scott Library in Merrill has done a yearbook archival project and the initial data range was 2-6 TB, and they ended up with the lower end of that from one yearbook project. This was just one project that took up 3 or 4 TB so the total need for all users could very well be much higher than the projected 26.25 TB. S. Prater noted that there is a natural bottleneck that will slow that growth, digitization, quality control, and writing metadata takes a lot of time and staff resources. It's unlikely that the total storage needs will quickly balloon although this is something to keep an eye on.

   c. **Discussion: Hardware Replacement Timeline**
The group discussed the timeline for hardware research and replacement. It was noted that a group consisting of the host sites and current users of the Technology Backup side of the project will be getting together in February to begin discussing hardware replacement. It was shared that the list of users provided by K. Whitson will help determine who needs to be involved and consulted on that process.

A. Hoks noted that most likely there will be two separate systems for the next-gen hardware as the integration that was thought would happen, has not happened.

It was asked that the scope of the needs of this group does need to be confirmed, is this simply backup storage for digitized material? K. Whitson noted her understanding is that it's just primary file storage and that there are no content management system needs for Recollection Wisconsin content partners. It was asked to clarify who Recollection Wisconsin content partners are, those with content on MPL CONTENTdm server? K. Whitson noted that it would include both those that have content on MPL's CONTENTdm system and partners that have access through Recollection Wisconsin but have their own storage areas. It was noted that Recollection Wisconsin would also be a catch-all for organizations not in a library system, including a public library in a system that is not participating.

3. **Next Meeting Dates**
The next meeting is December 9, 2024, at 1:00 pm.

The meeting ended at: 11:03 am