

The “M” Word: Exploring File Format Migration with Open Source Tools

LISA GREGORY

Digital Information Management Program, State Library of North Carolina

Abstract: As the official permanent depository for all North Carolina state publications, the State Library of North Carolina is concerned with preservation and access of these materials, regardless of file format. This paper describes that institution’s efforts at file format migration. With a limited budget and programming resources, migration file formats that match current and projected needs as well as free and open source tools that would normalize and document that migration were investigated. Although far from perfect, the tools described can effectively migrate a number of prevalent formats on a case-by-case basis. Work still needs to be accomplished to scale migration up to production level.

Introduction

“It can be argued that unless an object is accessible, it cannot be said to be preserved, as an inaccessible chunk of zeroes and ones is of no use whatsoever. Thus, any talk of preserving digital objects must include ways to access the objects” (Clausen 2004, 3).

The State Library of North Carolina (SLNC) is the official permanent depository for all North Carolina state publications. SLNC currently uses a number of third-party tools to manage short-term access to its digital collections and staff are working on refining a long-term access plan. The most recent installment in this plan is an investigation of file format options for migration as well as available tools that would match an achievable workflow and available resources. Similar to the reasons mentioned in Lawrence et al. (2000), it was decided to test migration rather than emulation (the other often-proposed long-term file access option) because of the environment and the programming expertise at staff disposal.

Before embarking on this process, the “m” word invoked a fair bit of trepidation – not quite an expletive, but one that caused a bit of a cringe. But between the library’s legal mandate and passion for maintaining access to government information, the process of digital preservation can’t be relegated to polite company forever. This paper recounts a first foray into migration testing, and it turned out a lot better than expected.

Problem Statement

The SLNC, and most government depositories in general, know that many of the files they are mandated to care for come in multitudes of formats. And yet funds for staff and technology are



© by Lisa Gregory.

Published by Society of American Archivists, April 2011.

rarely in abundance. With a very limited budget and a digital collection that could grow exponentially in the coming years, SLNC staff set out to determine whether or not currently available free and open source tools are a viable option for migrating data as part of a program for long-term access and, if so, which ones function best for the needs of libraries with similar resources.

Methodology

Investigating file formats and mapping transformations

At the SLNC, current digital collections are predominantly textual in nature with a mixture of born-digital and digitized content. Files are received from several locations: (1) digitized in house, (2) digitized at a local partner institution, (3) retrieved from state agency web sites, and (4) submitted by agencies through email or on disc. A significant amount of web content is also stored with the Internet Archive's Archive-It through an ongoing web harvesting program.¹ With digital publications increasing in state government, staff want to be prepared to accept and manage whatever range of file formats are submitted to the collection.

A selection of file formats currently in repository storage was chosen for testing, as well as a number of formats considered likely to be received in the future. After identification, file formats were loosely grouped by type. The bulk of current and anticipated formats fell into the "Images and Structured Graphics" and "Document-Like" categories. Also included in the test were a spreadsheet, audio/video files, and a web archive file. In addition to a variety of files, candidates that represented a range of file attributes were selected. Those attributes included some from varying dates, .pdfs with security settings and embedded content, a range of file sizes, multi-page .tifs, and converted files (such as .pdfs created from images or text files). The code was manipulated in two of the .pdf files as well, to determine if corrupted files could be migrated and if the corruption would be caught during any part of the process.

After file formats were chosen, a review of current literature as well as other institutions' recommendations for file format transformations was conducted to determine the recommended output format for each type. Regardless of the content of a file, formats are generally preferred for migration when they are uncompressed, encoded in an open standard, widely used, and interoperable (Brown 2003). The SLNC fully intends to keep the original file formats for all files preserved; migrated files will help keep content accessible. [Table 1](#) shows the desired transformations for the target files and the supporting references from which decisions were made, if any.

Out of all of the file formats chosen for investigation, no recommended migration format for Microsoft Publisher (.pub) and Adobe Photoshop (.psd) files could be found. Each could be migrated to a newer version using proprietary software, but no open format accommodates the content, look, and interactivity of these files. These were migrated to more static formats (.pdf and uncompressed .tif, respectively) using proprietary software as the only solution on hand which could be incorporated into the current workflow.

¹ See <http://webarchives.ncdcr.gov> for more information.

After deciding on each file’s migration equivalent, characteristics of the original file format and significant properties hoped to be retained after migration were identified. These mapped closely to those described in Clausen (2004), which are listed below in parentheses. For each file format, the following were desired

1. No visual loss of content (readability);
2. No loss of metadata;
3. Minimal degradation in quality (appearance, “look & feel”);
4. Minimal degradation in structure (comprehensibility);
5. Minimal degradation in interactivity (functionality).

Partly because expectations varied depending on file type and partly because the reasonable expectation of loss during this process was unknown, exact characteristics for each format were not quantified.

Table 1. Migration Formats and Associated Tools

Original Format	Desired Migration Result	Selected Tools	Sources (see References)
Document-Like			
.css	.txt	XENA, PLANETS	
.doc (all versions)	.odt	XENA, PLANETS	2 , 11
.docx	.odt	XENA, PLANETS	2 , 11
.html	.txt	XENA, PLANETS	
.pdf	.pdf/a	XENA, PLANETS	2 , 3 , 15
.ppt	.odp	XENA, PLANETS	2 , 11
.pub	.pdf*	<i>Acrobat</i>	
.rtf	.pdf/a	PLANETS, <i>Convert Doc</i>	8
.txt	.txt	XENA, PLANETS	2 , 11 , 15
Images and Structured Graphics			
.ai	.svg	Inkscape	2 , 11
.psd	.tif (uncompressed)*	XENA, PLANETS	
.gif	.tif (uncompressed)	XENA, PLANETS	2 , 3 , 8 , 11
.jpg	.tif (uncompressed)	XENA, PLANETS	2 , 3 , 8 , 11 , 15
.tif (compressed)	.tif (uncompressed)	XENA, PLANETS, <i>AVS Image Converter</i>	2 , 3 , 8 , 11 , 15
Audio/Video			
.mov	.mpeg-2 + mxf wrapper	FFmpeg	2 , 3 , 8 , 11 , 15 , 17

.mp3	.wav file + bwf header	FFmpeg	1 , 2 , 3 , 5 , 8
Spreadsheets			
.xls	.odf	XENA, PLANETS	2 , 11
Web Archives			
.arc	.warc	*	3 , 18

* No recommended format or tool found.

NOTE: Tools in italics were used, but do not fit into the requirements for the final workflow (see the “Tools” section below).

Choosing Tools

Tools were chosen if it was felt that they could be used in practice within the institution and be incorporated fairly quickly into the current workflow with a minimal level of resources, expertise, and technology support. With this in mind, the following requirements were established. The tools would need to:

1. Be free,
2. Be open source,
3. Be relatively well documented,
4. Be currently maintained,
5. Provide an audit trail and, of course,
6. Perform the required transformation(s) successfully.

In addition to the requirements above, it was hoped that tools could be located that could also:

1. Be used easily (preferably with a GUI) and
2. Transform multiple formats.

After surveying a range of options, FFmpeg, Inkscape, the PLANETS Testbed, and XENA appeared to be the best options meeting the list of requirements above for the range of file formats chosen for testing.² Although the PLANETS Testbed itself is not open source, it takes advantage of a number of open source tools to complete its experiments. A brief description of the most relevant characteristics of these tools for this project (in addition to the criteria mentioned above) can be found in the [Appendix](#).

A number of other tools were also reviewed, some of which are rolled into the tools used (like ImageMagick, JHOVE, DROID) but the tools chosen for testing could accommodate the largest range of file formats, with a few exceptions to fill in the gaps. There were two additional tools originally in the testing plan. The first was PLATO³ (Planets Preservation Planning Tool), described on the project website as “a decision support tool that implements a solid preservation planning process and integrates services for content characterisation, preservation action and automatic object comparison in a service-oriented architecture to provide maximum support for

² For several files, FITS (File Information Tool Set) was also used. Created by the Harvard University Library, FITS “identifies, validates, and extracts technical metadata for various file formats” and then outputs the results in an XML file. FITS does not migrate files, but is useful for verifying metadata or identifying the type of file you have on hand. For more information, see <http://code.google.com/p/fits/>.

³ <http://www.ifs.tuwien.ac.at/dp/plato/intro.html>

preservation planning endeavours.” Staff were unsuccessful in connecting to PLATO, and decided not to pursue access for this project. The second was warc-tools,⁴ which was the only option that fit the above criteria and could convert .arc files to the .warc format. At the time of this project, warc-tools was still being developed and the code had not been released for public use.

Testing Procedure

Before file migration began, files were checksummed, copied to a single folder on a local server, and then re-checksummed to verify the files were intact. All of the tools mentioned above (with the exception of the web-based PLANETS Testbed) were installed. (It should be noted that XENA also requires installation of OpenOffice.⁵) During migration testing, each file was run through its associated tool(s). Any particular difficulty using the tools was noted. The viewers in the associated tools as well as the output from FITS were used to compare the retention of significant properties and retention of metadata as discussed above.

Results

Following are the results of the migration tests. All files with mapped migration transformations were tested at least once through at least one of the tools. The tests were done over the course of two days. See [Table 2](#) for a brief overview of the test results and conclusions.

Document-Like Files and Spreadsheets

Almost all of the document-like files were migrated using both XENA and the PLANETS Testbed, with the exception of Microsoft Publisher files. As might be expected due to their lack of interactivity and relational data, .txt and similar files (.css and .html) rendered best during this process. Both content and structure remained intact, and XENA packaged them nicely with a metadata wrapper.

XENA, although successful at migration, caused some consternation with its viewer. For both .doc and .docx files, details like tables, tabs and bullets did not render exactly in the viewer, but were fine when the migrated file was opened directly in OpenOffice. The .ppt file also migrated successfully in XENA,

⁴ <http://code.google.com/p/warc-tools/>

⁵ <http://www.openoffice.org>

Table 2. Migration Results



© by Lisa Gregory.

Published by Society of American Archivists, April 2011.

Original Format	Desired Migration Result	Tools Attempted	Converted Successfully?	Rendered Well?	Metadata Retained?	Results Considered Acceptable?	Notes
Document-Like							
.css	.txt	XENA, PLANETS	Y	Y	Y	Y	
.doc (all versions)	.odt	XENA, PLANETS	Y	*	Y	Y	Tables & tabs did not render exactly in XENA; fine in OpenOffice.
.docx	.odt	XENA, PLANETS	Y	*	Y	Y	Bullets did not render exactly in XENA; fine in OpenOffice.
.html	.txt	XENA, PLANETS	Y	Y	Y	Y	
.pdf	.pdf/a	XENA, PLANETS	N	n/a	n/a	N	Neither tool could accommodate migrating .pdf to .pdf/a.
.ppt	.odp	XENA	Y	*	Y	Y	Page numbers missing, some font and footers did not render exactly.
.pub	.pdf ?	**	N	n/a	n/a	N	No tool found.
.rtf	.pdf/a	PLANETS, <i>Convert Doc</i>	Y	Y	N	N	Errors when migrated file was checked for pdf/a-1a and pdf/a-1b compliance.
.txt	.txt	XENA, PLANETS	Y	Y	Y	Y	
Images and Structured Graphics							
.ai	.svg	Inkscape	Y	*	Y	Y	Font formatting and color subtly different.
.psd	.tif (uncompressed)	XENA, PLANETS	Y	N	Y	N	Rendered, but no .psd functionality (layers, etc.) retained.
.gif	.tif (uncompressed)	XENA, PLANETS	Y	Y	Y	Y	
.jpg	.tif (uncompressed)	XENA, PLANETS	Y	Y	N	N	File header's "modified date" was changed to experiment date.
.tif (compressed)	.tif (uncompressed)	XENA, PLANETS, <i>AVS Image Converter</i>	Y	Y	Y	Y	
Audio/Video							
.mov	.mpeg-2 + mxf wrapper	FFmpeg	Y	N	Y	N	Considerable degradation in video and audio quality.
.mp3	.wav file + bwf header	FFmpeg	Y	Y	Y	Y	
Spreadsheets							
.xls	.odf	XENA	Y	Y	N	N	Author, manager, company metadata lost.
Web Archives							
.arc	.warc	**	N	n/a	n/a	N	No tool found.

* Yes, but with some loss.

** No tools found within project criteria.

although page numbers were missing and some fonts did not render exactly. The Microsoft Excel (.xls) file came through surprisingly well as an .odf file. Content, tabs, formulas and a text box appeared to be the same as the original. The only drawback was the loss of properties (author, manager, company, etc.) metadata, which were absent in the .odf file.

The .pdf/a file type presented more challenges than expected. XENA does not normalize .pdf to .pdf/a, but simply wraps the .pdf file in XML. Of the two manually corrupted files, neither rendered in the XENA viewer and only one could be identified as corrupted from the XENA report. While .pdf/a is an output option in the PLANETS Testbed, .pdf is not one of its corresponding input options. It appears that the Testbed does not allow users to input a file as undefined or unidentified. In the end, .pdfs were not successfully converted to .pdf/a using any of the open source tools in this experiment.

To explore the possibility of using .pdf/a as a fall-back migration format for textual documents, Rich Text Format (.rtf) was migrated to .pdf/a. Unfortunately, .rtf is not available as an input format for .pdf/a in the PLANETS Testbed. A free tool that would make the required conversion, Convert Doc by SoftInterface, Inc., was chosen.⁶ It successfully converted the file to .pdf/a. Although the content and formatting appeared to be correct, there were multiple errors when the output file was checked for compliance to pdf/a-1a and pdf/a-1b. There is no plan at this point to use .pdf/a as a migration format for .rtf, but this was a useful demonstration of some of the issues that would have to be resolved should this workflow ever be considered.

The last document-like file format chosen for conversion using an open source tool was a Microsoft Publisher (.pub) file. As mentioned above, no recommended open preservation format or tool that fit the tool criteria could be identified so this file was not converted successfully.

Images and Structured Graphics

As with the document-like file formats, this category was dominated by XENA and the PLANETS Testbed. While using the Testbed, the ImageMagick migration service was chosen.

Unlike XENA, which can convert .gif formats to .png, the PLANETS Testbed offered .tif transformation for .gif files. The Testbed also successfully migrated a .tif file with LZW compression to an uncompressed .tif. When it came to migrating a .jpg to .tif, the migrated content was correct, however the “modified date” metadata in the file header of the original object was changed to indicate the date of the experiment. This only happened with .jpg to .tif transformations.

The PLANETS Testbed accommodated transforming an Adobe Photoshop (.psd) file to a .tif file. The resulting file was excellent but any layers or objects in the .psd file were not retained, leaving a limited range of options for reuse or examination of the original structure of the file.

⁶ It should be noted that Convert Doc does *not* meet the criteria for tools described in the methodology. A free trial is available, which was used for this test.

Neither XENA nor the PLANETS Testbed could accommodate migrating .ai files, so Inkscape was used to perform this migration. The content of the .ai file remained intact; however the font formatting and coloring seemed mildly different in the resulting .svg.

Audio and Video Files

FFmpeg successfully converted QuickTime (.mov) and .mp3 files. Conversion of .mov to .mj2 was the preferred migration format, but after having trouble locating a tool that fit the project criteria and a further review of the literature, .mpeg-2 with an .mxf wrapper seemed to be an acceptable alternative. FFmpeg did successfully perform the conversion, however there was significant degradation in the quality of both the sound and video. FFmpeg also successfully converted the .mp3 file to .wav, and the quality in the resultant .wav file was much more comparable to the original than in the video file conversion.

Web Archives

Unfortunately, no tool could be found that met the criteria and that could transform .arc to .warc. While the warc-tools project looks promising, the product was not yet available for testing. The Internet Archive stewards the SLNC's .arc files for the time being, but hopefully a tool will become available at some point in the future to allow experimentation on those files.

Findings

In general, both XENA and the PLANETS Testbed performed well for document-like file types. In some cases, using the XENA viewer presented a different viewable result than exporting and viewing in OpenOffice. XENA inconsistently migrated the metadata (properties) from the original file format. Migrating .pdf to .pdf/a using XENA or the Testbed was not successful, and Microsoft Publisher files presented the most obstacles for migration.

XENA, the Testbed, and Inkscape all successfully converted the targeted image files. The primary issues were the incorrect metadata for the .jpg to .tif conversion and the difficulty in preserving structured data (layers, etc.) in the structured graphics formats. The latter was expected, but the former was not.

FFmpeg worked well for both .mov and .mp3 files, although .mpeg-2 does not seem to be an acceptable migration destination for .mov files due to the poor quality of the result. The different nature of time-based files (audio and video) highlighted SLNC staff's lack of expertise with terminology and technology in this area, making it more difficult to select the best migration alternatives. It was also concluded that the quality of audio and video files, because they include so much data layered in the format, should not be tested solely using human perception.

In addition to the success of each file format migration, the tools used left some general impressions. XENA, though limited in some of the migrations, worked well. Its GUI interface, which is not a given with open source tools, is more user friendly for those with less experience working from a command line. The primary difficulty with XENA was the XENA viewer, which did not always render things in the same way as OpenOffice.

FFmpeg can be intimidating for multiple reasons. Those not used to working at the command line will require time to get up to speed. In addition, between compressions, frame rates, and

codecs, repositories not predominantly dealing with audio visual material and that lack that expertise will find a definite learning curve. Lacinak (2010) is highly recommended as a get-started resource.

As for the PLANETS Testbed, the extended documentation functions are detailed and flexible and offer the ability to comment and report often during the process. The user interface was straightforward, requiring users to go through each step of the “experiment” might be overwhelming to someone interested in a very specific function. The structure of the FTP area (used for processing batches of files) was somewhat confusing. During the experiment, it was unclear what a user would do if he or she had not validated the input file format or if the exact format version was unknown. It seems like it would be easy to incorporate FITS or one of its constituent tools as an extra step to suggest an input file format. Finally, the Testbed is precisely that – it does accommodate batch migration, although if it’s a large batch it may be scheduled to avoid overloading their resources. However all of the files in the batch must be of the same file type, which does not seem ideal for many real-life situations.

Limitations

The migration testing described here has a number of limitations. Some of these were self-imposed or imposed due to resource constraints. Proprietary and for-fee software would offer more and, possibly, more robust options, however purchasing such software for testing is not an option for the near future. Selected file formats were limited to objects in the SLNC repository or those created recently. None of the files were older than 2001, and yet it is likely that older legacy publications from a state agency will be received at some point in the future. By confining testing to real files at hand, “maxed-out” versions of each file format, such as a spreadsheet replete with macros and higher order functions or a presentation full of videos and audio clips, were not included. For an upcoming second stage of testing, additional file types from a partner agency will be used to expand the scope.

As mentioned above, staff also felt limited by a lack of knowledge regarding the general components and structure of video and audio formats. While education did happen on the fly, more information would be helpful to truly ensure the fewest significant properties of those formats were lost during migration.

Related to this idea, all of these file formats were only visually inspected to determine attribute retention. Metadata was examined, images were compared on the same monitor, and audio files were listened to as closely as possible. It is acknowledged that there are more exacting ways to determine the difference between two similar files. In most cases, however, a visual analysis was determined to be enough, or the difference between the original and migrated format was so marked that it was sufficient to make a determination about whether or not the migration could be considered successful.

Finally, as is apparent from the description above, this was a very manual process. In general, files were walked through each program one by one. Because these were open source tools, some automation could be achieved easily with limited scripting knowledge, simply by setting up recursive actions. In other cases, it would be helpful to have multiple programs work in concert, as FITS does for the purposes of validating and extracting metadata. For any repository

with a significant amount of content, and especially one that isn't planning to migrate on the fly, procedures and programs that not only ensure the integrity of the original file formats but also could be performed in batch would need to be developed.

Conclusions and Next Steps

This file format migration process was informative and (pun intended) refreshing. On the whole, both XENA and the PLANETS Testbed were easy to use, retained more file contents than anticipated, and supplied plenty of audit trail information. Institutions could definitely use the former in a production environment; the latter is best for getting used to the file migration process, solidifying local workflows, and refining which transformations result in the highest quality files. FFmpeg and Inkscape were also viable alternatives, albeit not as multi-purpose for SLNC collections as the other tools tested. On the ground, FFmpeg, Inkscape, XENA, and the PLANETS Testbed (or its constituent parts) could all be used to migrate some file formats quite successfully. In order to scale up to processing large batches of files, however, an institution would need to develop batching procedures to interact with these programs.

Going forward, these tools will be put through a second round of testing to verify results. The test batch of files will be broadened to include other formats and variations, perhaps including a sample of legacy items held at another state agency. In addition to keeping tabs on new tools, processes will need to be found or developed that can scale to production before incorporating them into the workflow on a regular basis. Perhaps the most valuable result of this testing was an increased comfort level with migration in general. Eventually, if everyone's saying it, it might not be such a bad word after all.

Appendix: Tools Used

FFmpeg

<http://ffmpeg.org/index.html>

FFmpeg can be used to transform, create, and stream audio and video files. Although FFmpeg does not have an institution behind its development, it has been widely used and incorporated into a variety of projects.

Creation Date: Unknown – at least since 2004

Documentation: Web page with intermediate and advanced materials (these do not introduce you to the concepts behind video and audio encoding, but do give you examples and specific commands)

Platform(s): Windows, UNIX, Mac

Interface: Command line

Inkscape

<http://inkscape.org/>

Inkscape was chosen to accommodate converting .ai files to .svg files, something that the other multi-purpose tools could not handle.

Creation Date: 2003

Documentation: Wiki, FAQ, tutorials blog, and robust user community

Platform(s): Windows, UNIX, Mac OS X

Interface: GUI

PLANETS Testbed

<http://testbed.planets-project.eu/testbed/>

According to the PLANETS Testbed website, the Testbed “provides a dedicated research environment where services and data can be experimented upon, results can be evaluated and outcomes shared with the wider community.” Released for broader use in 2010, the site provides robust features for documenting preservation experiments.

Creation Date: 2007 internally, 2010 for broad use

Documentation: User guide, some documentation on SourceForge

Platform(s): Independent

Interface: GUI

XENA (Xml Electronic Normalizing for Archives)

<http://xena.sourceforge.net/>

XENA is a tool specifically designed for digital preservation activities. It can detect file formats and then convert files into open formats for preservation. The National Archives of Australia has been developing this tool for quite awhile (when compared with other open source tools in this area). The last release was in December of 2009.

Creation Date: 2003

Documentation: Wiki with introductory and advanced materials

Platform(s): Windows, Linux, Mac OS X

Interface: GUI

References

1. Addis, Matthew, and Richard Wright. 2010. Audiovisual preservation strategies, data models and value-chains. PRESTOPrime project. https://prestoprime.ina.fr/public/deliverables/PP_WP2_D2.1.1_preservationstrategies_R0_v1.00.pdf (accessed August 19, 2010).
2. Archivemata. 2010. Media type preservation plans. http://www.archivemata.org/wiki/index.php?title=Media_type_preservation_plans (accessed August 19, 2010).
3. Arms, Caroline R. and Carl Fleischhauer. Format descriptions. Sustainability of Digital Formats Planning for Library of Congress Collections. <http://www.digitalpreservation.gov/formats/fdd/descriptions.shtml> (accessed August 19, 2010).
4. Brown, Adrian. 2003. Digital preservation guidance note: Selecting file formats for long term preservation. United Kingdom: The National Archives. http://www.nationalarchives.gov.uk/documents/selecting_file_formats.pdf (accessed August 19, 2010).
5. CARLI. 2009. Digitization best practices for audio. Consortium of Academic and Research Libraries in Illinois. http://www.carli.illinois.edu/mem-prod/contentdm/guidelines_for_audio.pdf (accessed August 19, 2010).
6. Clausen, Lars. 2004. Handling file formats. Århus, Denmark: State and University Library / Copenhagen, Denmark: Royal Library. <http://netarchive.dk/publikationer/FileFormats-2004.pdf> (accessed August 19, 2010).
7. Fleischhauer, Carl. Format considerations in audio-visual preservation reformatting: Snapshots from the Federal Agencies Digitization Guidelines Initiative. *Information Standards Quarterly* 22, no. 2 (2010): 34-40. http://www.digitizationguidelines.gov/audio-visual/documents/IP_Fleischhauer_AudioVisual_Reformatting_isqv22no2.pdf (accessed August 19, 2010).
8. Harold B. Lee Library. Digital formats. <http://www.lib.byu.edu/departs/dp/guidelines/formats.html>
9. Lacinak, Chris. 2010. Primer on codecs for moving image and sound archives. New York: AudioVisual Preservation Solutions. http://www.avpreserve.com/wp-content/uploads/2010/04/AVPS_Codec_Primer.pdf (accessed August 19, 2010).
10. Lawrence, Gregory W., William R. Kehoe, Oya Y. Rieger, William H. Walters, and Anne R. Kenney. 2000. Risk management of digital information: A file format investigation. Washington, D.C.: CLIR. <http://www.clir.org/pubs/reports/pub93/pub93.pdf>. (accessed August 19, 2010).
11. Prom, Chris. 6. Develop preservation/access plans. Practical E-Records. http://e-records.chrisprom.com/?page_id=581 (accessed August 19, 2010).
12. ---. May 7, 2010. Planets Testbed review. Practical E-Records. <http://e-records.chrisprom.com/?p=1183> (accessed September 21, 2010).
13. ---. April 25, 2010. PLATO (Digital Preservation Planning) software review. Practical E-Records. <http://e-records.chrisprom.com/?p=1082> (accessed September 21, 2010).
14. ---. April 22, 2010. Review of XENA normalization software. Practical E-Records. <http://e-records.chrisprom.com/?p=1081> (accessed September 21, 2010).

15. Rohde-Enslin, Stefan and Keith R. Allen. 2009. Nothing lasts forever. nestor / Germany: Institute for Museum Research. http://files.d-nb.de/nestor/ratgeber/ratg01_2_en.pdf (accessed August 19, 2010).
16. Rosenthal, David S. H., Thomas Lipkis, Thomas S. Robertson, and Seth Morabito. Transparent format migration of preserved web content. *D-Lib Magazine*, January 2005. <http://www.dlib.org/dlib/january05/rosenthal/01rosenthal.html>. (accessed September 21, 2010).
17. Schüller, Dietrich. 2009. Video archiving and the dilemma of data compression. *Internet Preservation News* 47:5-7. http://www.ifla.org/files/pac/IPN_47_web.pdf (accessed August 19, 2010).
18. Strodl, Stephan, Peter Paul Beran, and Andreas Rauber. 2009. Migrating content in WARC files. Paper presented at the 9th International Web Archiving Workshop in Corfu, Greece. http://publik.tuwien.ac.at/files/PubDat_181115.pdf (accessed August 19, 2010).
19. Van Wijk, Caroline. 2006. KB and migration test plan. National Library of the Netherlands. http://www.kb.nl/hrd/dd/dd_projecten/KB%20and%20Migration%20Test%20Plan.pdf (accessed September 21, 2010).