Maureen Bocka
713 Research paper
Fall 2019

Intersections between Digital and Print Preservation

Preservation is a wide field to look into. Print preservation policies have been impacted due to decisions made about digital preservation. This is evident in many libraries, but especially in academic libraries. This essay will look at how digital preservation and print preservation impact each other based on preservation policies and collection development methods. This topic is of interest because there has been discussion of preservation, but it is interesting how different aspects of preservation play off each other. Some of the issues that librarians face when choosing what to preserve are limited storage space, limited equipment, finite collections, and deciding what needs to be digitized.

One of the problems that preservation specialists face is having limited storage space. This is true for digital and print collections. No library has unlimited space for all the works that have been created to show how human study has progressed. As a result, libraries have to figure out what to do with items that are popular enough to survive the weeding process, but not popular enough to be requested often. This can also make it harder to have collections that are static as collections are being constantly adjusted to get more space.

One tactic that is used by academic libraries is to reformat journals. Journals are often focused on because once a decision is made for them to go to a digital format, the library gets to have "more space per decision" (Horava, et al. 2017, p.3). This is compared to the amount of space that would be made if individual books were weeded. Doing this can free up several bookcases worth of space. This is helpful for growing collections or making space to use for another purpose. Some libraries have a last copy preservation policy. That means if a book is the

last copy that exists in the library/ state then that book gets sent to a special storage facility for long-term storage. Moving to digital journals also means that the digital preservation could potentially be done by a third party (Meddings, 2011, p.55). This would take some of the responsibility away from the library and allow the library to focus on other aspects of preservation.

Another way to combat limited storage is for libraries to band together and form a network that collectively own materials and share storage space. This allows for libraries to get books that are low circulating off their shelves, and yet remain available while being stored in ideal conditions. Horava et al. (2017) found that this system allows for libraries to save money on preservation by only having to preserve a single collection of low functioning books (p. 6). This also helps libraries by condensing storage and allowing libraries to have more usable space.

Information is still being preserved and buys the library time to either change to a new format or eventually weed the low circulating books. However, in order to do this successfully, libraries have to come to a common understanding of what preservation means and what benefits the libraries want to get out of the partnership. A great deal of planning has to come about for any partnership to take off and it is a lot of work. The rewarding thing is that after all this work the libraries will be able to have more usable space in their libraries and will save money in the long run when it comes to storage.

Horava et al. (2017, p. 6) found out that one of the struggles that came about from creating partnerships between universities was figuring out which books were duplicates. This was due to the fact that 4 different ILS systems at the 5 partnering universities (Horava et al. 2017, p. 6). Another issue was finding the funding to take on such a big project. As the participating universities are located in Canada the only funding they received was from the

province (Horava et al. 2017, p. 6). There was also no additional staff hired to help take on the responsibilities that come from creating partnerships.

In addition to having limited storage space, libraries also have finite collections. This does not seem like a problem at first, but there is more information than ever. Black, 2016 makes the point that collections have become representative of the work that is being published and cannot be comprehensive (p. 3). Only a small percentage of all information produced can be preserved. This is especially true now that anyone can become a producer of information. This puts more pressure on librarians to be picky about what they select for the collection. Especially when there is limited space for the collection to grow in.

Schonfeld (2011. p. 141) brings up the issue that administrators are putting more pressure on libraries to reduce the size of their print collections. However, it is often difficult to determine how to slowly eliminate print journals in a way that makes sense to the patrons (Schonfeld, 2011, p. 142). Some patrons and staff may be uneasy about having the print versions of the journal go away. This transition is made easier when there is a timeline when the print journals will be phased out and the digital format will take over. It is easier to consider preservation when the publisher provides digital versions of materials such as journal. If the library has to manually digitize them it takes longer to make the transition from one format to the next. Making it an even longer time to reduce pressure on storage capacity.

Making the transition from print to digital brings up preservation issues about the future of the items in the collection. One future issue that libraries have to be ready to face is changes in access agreements from the vendor or publisher. Some libraries join an agreement to keep one entire set of a journal collection in case this should happen (Schonfeld, 2011, p. 142). This is also illustrated by (Horava et al. 2017). Other libraries do not hold onto the print copies and go

straight to a digital collection (Schonfeld, 2011, p. 143). This puts the library at risk in case something changes to the digital access agreement.

There are plenty of preservation reasons why a library would want to hold onto print copies. It is helpful for fixing scanning errors, fixing errors in preservation, and preserving elements of the print copy that are not correctly represented in the digital version (Schonfeld, 2011, p. 143). It could potentially be very valuable to have a print copy in case the collection has to be redigitized due to changes in hardware or software that could not be solved with migration. If the library can easily have access to the print version of the journal by interlibrary loan then the same preservation needs can be met.

One problem that (Meddings 2011) and (Schonfeld 2011) agree on is that relying on the publisher can be questionable for long-term digital preservation and access. (Meddings 2011, p. 55) confirms that proper long-term digital preservation means materials remain available to patrons. This does not include post-cancellation access. Another problem that (Meddings, 2011, p. 58) brings up is that not being able to rely on publishers for digital preservation means that costs for the library will go up. This also means that there is a lingering issue of access if publishers change the terms of agreement. Then the library would have to figure out another way of getting access to the journals in question.

Having access to print copies or even an offline version would help ensure access if there is a cancellation of services. Meddings (2011) also makes the point that born-digital items cannot fit in a print archive due to the sheer volume of information (p. 55). (Meddings, 2011, p. 58-59) found that there is a preference for publisher and library partnerships over libraries or publishers being solely responsible for digital preservation. There is also an uncertain future about who will continue to fund digital preservation (Meddings, 2011, p. 59). This is especially true for

universities that have been faced with budget issues. This is why it is important for libraries to understand their agreements and to have an action plan should their needs change.

Digital preservation has its own set of challenges that are unique from print preservation. Long-term digital preservation has to make sure that there is infinite usability, authenticity, discoverability, and accessibility (Jan, 2019, p. 51). This is further complicated due to hardware and software becoming obsolete very quickly due to advances in technology (Jan, 2019, p. 51). This is why it is important that libraries make sure their technology stays up to date. This also makes digital collections "fragile" as they can quickly become obsolete (Lamgley, 2019, p. 93).

Not only can user agreements change, but how files are saved and read can also change. Additionally, content, context, structure, appearance and functionality of the digital item also have to be considered (Jan, 2019, p. 50). It is very easy for digital items to become separated from their original context making it harder to interpret their significance. This is part of what makes digital conservation so tricky.

Similarly, to print collections, digital collections can only preserve a small percentage of the information that is available (Jan, 2019, p. 52). Migration can be used to preserve the integrity of digital objects. Migration is defined as the, "periodic transfer of digital materials from the hardware/ software configuration to another" (Jan, 2019, p. 52). Migration prevents the loss of information by making sure that there is not going to be obsolescence. Digital materials can be reformatted to ensure they will survive new software or hardware.

Migration also helps to avoid deterioration of digital resources (Jan, 2019, p. 52). It is helpful to be skeptical of standards used to preserve digital information as they are likely to continue to change with technology. The best way to find the solution to digital preservation challenges is to compare the techniques to the problems they will solve (Jan, 2019, p. 57).

Digital preservation falls under digital stewardship which is a more holistic way of looking at preservation. Digital stewardship incorporates digital archiving, curation, and preservation (Langley, 2019, p. 94). There has to be a balance between collecting and managing digital content (Langley, 2019, p. 94). It is easy to think of curation, preservation and archiving as completely separate things, but they all work together. Thinking about how they work together makes for better collections and for better management.

It is also important to continue to track where the digital resources are, who has access to them, and how to maintain them. Another problem is that many librarians think of preservation at the end of the information object's lifecycle. Preservation needs to be thought of at the beginning of the lifecycle so that there is a plan on how to preserve the functionality of the information object (Langley, 2019, p. 97).

Another tricky aspect of digital files is that there is no such thing as an "original file" (Langley, 2019, p. 97). There is such a thing as having the "earliest" available file which is preferred due it being less likely to have inaccuracies (Langley, 2019, p. 97). This can be difficult to do especially if the library has a large number of files. This is an example of why libraries need to have systems that contain and catalog their files so they know which file is the earliest.

Not only is it important to know how to find files, but it is also important to know which file format to use for digital preservation. There is a case to be made for using the PDA/F file format as it reduces the digitization workload and it requires less post digitization processing (Han and Wan, 2018, p. 53). PDA/F allows for greater quality control and it is easy to add or update missing pages (Han and Wan, 2018, p. 53). Using PDA/F also allows for a simpler one-to-one ratio between the original print object and the digital object (Han and Wan, 2018, pg. 53).

It is also easier to create metadata that accurately represents the digitized objects. Following the one-to-one ratio is important when digitizing print objects. There is also a PDA/F validator which is important because now it is easier to make sure the PDA/F files are in the correct format.

Sustainability of digital preservation is a major question for academic libraries. Part of the uncertainty comes from a lack of institutional support. This is seen as not being given enough staff to help with preserving born-digital records (Gorzalski, 2018, p. 29). There is discussion about having a librarian who specializes in born-digital items as the preservation program grows (Gorzalski, 2018, p. 29). This seems like a good thing, but it means that in the meantime library staff will continue to struggle to find out how to juggle all the digitization tasks.

This is especially troubling when staff have limited technical knowledge further underlining the importance of needing staff that has special training (Gorzalski, 2018, p. 26). Archivists say that electronic records are among the top challenges in the profession (Gorzalski, 2018, p. 27). There is also a lack of physical infrastructure to assist with digital preservation such as digital storage capacity (Gorzalski, 2018, p. 44). This puts strain on the library staff to come up with ways to store their digital collections without having the proper tools needed. Preservation cannot be sustainable without having the support from the university to build the setting that is needed. Digital preservation is not a trend or fad and will continue to become a more permanent part of the library as time goes on.

Libraries also face having a low influence to get the university to change the way it treats preservation. It is also disheartening when there is also a lack of policies and procedures on dealing with digital preservation. Budget problems have plagued academic libraries in general, but especially in Illinois when the state went through its budget crisis. Illinois was hit with a 61%

budget cut in 2015-2016 (Gorzalski, 2018, p. 29). University libraries are only now starting to recover from the financial strain they were put under. When the main source of income gets taken away from the university, it makes it harder for the library to continue to build programs and to hire staff. This sends a message to the library that it is not a priority even though it is trying to set up procedures and programs that will benefit the university in the long term.

Another issue that preservation specialists face is building the infrastructure they need to digitize items. (Black, 2016, p. 61) points out that costs for libraries go up and budgets are stagnant. This creates a problem when a university needs to be able to digitize a large collection and they have to build a lab or studio. Zhou (2016, p. 26) discovered that it is possible to build a digitization studio on a budget. Zhou compared the way high quality digital photos from a Nikon D800 camera to scans from an overhead scanner work to preserve images.

One of the problems with using the overhead scanner was that it did not always focus correctly (Zhou, 2016, p. 27). There was also no signal to the user that the scanner was completely focused. The camera beeped or flashed when it was properly focused which reduced the number of wasted shots (Zhou, 2016, p. 27). The Nikon camera also provided clearer and more detailed shots. This makes for better preservation of the images and text. The clearer the picture the less likely it is to need to be digitized again in the future.

When using the camera and scanner to preserve text, the Nikon camera still came out ahead (Zhou, 2016, p. 28). Getting a clear shot the first time helps to reduce the number of overall photos taken. It was determined that sharpening filters could be used to increase the overall quality of the images taken with the overhead scanner, but this would be time consuming and can create haloing (Zhou, 2016, p. 36). Overall, using a Nikon camera to preserve images digitally outperformed the overhead scanner.

The price difference between the Nikon camera and the overhead scanner are large. The Nikon D800 camera costs $4,500. The overhead scanner costs $55,000 with an additional annual maintenance fee of $8,000 (Zhou, 2016, p. 27). After buying all the other equipment needed to create a digitization studio with the Nikon camera it would cost between $10,000-$15,000 to do. This would provide a dramatic savings for a library. There would also be no annual maintenance fee. This example shows that it is possible to use cheaper tools to get better results compared to using expensive equipment.

Overall, there are many problems that libraries face when dealing with preservation both print and digital. The staff may not have all the technical training, budgets have become stagnant after facing a financial crisis, there is a lack of infrastructure to support digital preservation, libraries are running out of storage space, while collections are being forced to become more representative of the work that is being created. None of these problems are new and there are some creative ways universities are solving them. Some universities are building partnerships to share storage space for the print collection. Preservation is an important function that university libraries provide. It is an important field that needs to ensure it will be able to be successful in the future.

BLACK, D. M. (2016). Pascal's Wager and the Long Conversation: Preserving the Scholarly Record in the Academic Library. *Michigan Academician*, *44*(1), 61–72. https://doi-org.dom.idm.oclc.org/10.7245/0026-2005-44.1.61

Gorzalski, M. (2018). Digital Preservation Practices among Midwestern Four-year Public Colleges and Universities. *Archival Issues: Journal of the Midwest Archives Conference*, *39*(1), 26–52. Retrieved from http://search.ebscohost.com.dom.idm.oclc.org/login.aspx?direct=true&db=lls&AN=133440634&site=eds-live&scope=site

Horava, T., Rykse, H., Smithers, A., Tillman, C., & Wyckoff, W. (2017). Making Shared Print Management Happen: A Project of Five Canadian Academic Libraries. *Serials Review*, *43*(1), 2–8. https://doi-org.dom.idm.oclc.org/10.1080/00987913.2016.1274209

Meddings, C. (2011). Digital Preservation: The Library Perspective. *Serials Librarian*, *60*(1–4), 55–60. https://doi-org.dom.idm.oclc.org/10.1080/0361526X.2011.556437

Schonfeld, R. (2011). What to Withdraw? Print Collection Management in the Wake of Digitization. *Serials Librarian*, *60*(1–4), 141–145. https://doi-org.dom.idm.oclc.org/10.1080/0361526X.2011.556453

Yan Han, & Xueheng Wan. (2018). Digitization of Text documents Using PDF/A. *Information Technology & Libraries*, *37*(1), 52–64. https://doi-org.dom.idm.oclc.org/10.6017/ITAL.V37I1.9878

Yongli Zhou. (2016). Fulfill Your Digital Preservation Goals with a Budget Studio. *Information Technology & Libraries*, *35*(1), 26–50. https://doi-org.dom.idm.oclc.org/10.6017/ital.v35i1.5704