

Preventing Digital Preservation: Copyright Legislation (C-11) and TPMs

Joshua Chalifour

www.chalifour.org

www.chroknowlogy.ca

29 November 2012

Preventing Digital Preservation: Copyright Legislation (C-11) and TPMs

Recent copyright legislation prevents archives from legitimately fulfilling key requirements for the long term preservation and provision of access to digital fonds. Bill C-11 (An Act to amend the Copyright Act)² changed many elements of copyright law but the area posing the greatest problems to archival practices is the portion that prohibits circumventing technological protection measures (TPMs).

The problems surface in the confluence of TPM usage with the present copyright framework's legal application to archival processes. TPMs in and of themselves are not necessarily detrimental to archival processes. Some forms of TPM may even be useful toward ensuring authenticity and identifying provenance. However, when considering strategies for long term digital preservation as applied in trusted digital repositories (TDRs) or when considering appraisal criteria for the acquisition of digital fonds, the law turns TPMs into barriers against preservation.

In the following, I will first examine the TPM environment with the changes in copyright law brought by Bill C-11. Second, I will assess the threat to long term digital preservation with TDRs. Third, I will identify some of the effects on appraisal.

I hope to show that portions of Bill C-11, introduced by the current Conservative government, prevent archivists from legally accomplishing their objectives.

The TPM Environment and Copyright Change from C-11

A TPM is any of a group of means used to discourage people from accessing or copying content stored in a digital medium; in fact sometimes TPMs are called digital locks. A TPM may discourage access through tracking and the threat of legal repercussions. Alternatively it may restrict access to only those people privileged with the appropriate means. What is appropriate is typically determined through an agreement with a rights-holder (such rights are bestowed on the basis of copyright law).

A TPM, as defined in Canadian copyright law (C-11), means

...any effective technology, device or component that, in the ordinary course of its operation,

(a) controls access to a work, to a performer's performance fixed in a sound recording or to a sound recording and whose use is authorized by the copyright owner; or

(b) restricts the doing — with respect to a work, to a performer's performance fixed in a sound recording or to a sound recording — of any act referred to in section 3, 15 or 18 and any act for which remuneration is payable under section 19.³

While this law takes a rather general approach as far as what qualifies as a TPM, TPMs usually come in the form of software programs or physical devices.

2 An Act to amend the Copyright Act. Statutes of Canada. 2012, c 20.

3 Section 47, An Act to amend the Copyright Act. Statutes of Canada. 2012, c 20.

The law speaks directly to both “control” and to “restrictions”, implying that someone or some organization will be the locus of that control and thus determine the scope of the restriction. We need to consider the nature of this control and related restrictions because it is critical to understanding associated archival problems.

Watermarks are a form of TPM, which tend to include metadata and other identifying information about the source of the content to which they're applied. A watermark might be something that is not perceptible to the person using or accessing the content but allows a rights holder to track and identify the content. TPMs like watermarks (which merely discourage use, and mostly in a commercial milieu) do not present archives with the problems that other forms of TPMs, which prevent or prohibit use do.

Encryption is a form of TPM that is of great concern with respect to long-term digital preservation and copyright law. Encrypted content ceases to be meaningful or even recognizable as content to people—upon being encrypted, it's arranged in such a way that it can only be rearranged into something meaningful if someone has the necessary code or other means to decrypt the content. Encryption can prevent use.

Other problematic forms of TPM may not encrypt the content but use techniques that simply don't permit the content to be accessed without some sort of an electronic or physical key. This might be applied through a proprietary application using storage formats that do not conform to open, public standards. In a broad sense, proprietary software applications themselves could be a form of TPM because their proprietary nature prevents them (and thus the content's accessibility) from being used on systems their developers didn't design them to operate on. Someone would require the rights to use not just the proprietary program but also the operating system on which it runs. There would be at least two levels of rights necessary to access the content, in addition to any content rights-holder permissions.

A significant way in which TPMs under Bill C-11 interfere with archival practices is in the digital backup process. I've mentioned several forms of TPM, so I would like to identify what the bill disallows with respect to these and then explain what that means to archival practices.

A portion of Bill C-11 covers making backup copies. Making backup copies is necessary to digital preservation practices. As I'll explain, it's a standard part of TDR strategies without which, problems arise for ensuring the integrity, availability, and long-term preservation of both digital content and digital surrogates of physical archives. Section 29.24 states:

It is not an infringement of copyright in a work or other subject-matter for a person who owns — or has a licence to use — a copy of the work or subject-matter... to reproduce the source copy if...

(c) the person, in order to make the reproduction, did not circumvent, as defined in section 41, a technological protection measure, as defined in that section, or cause one to be circumvented; and

(d) the person does not give any of the reproductions away.⁴

Presumably an archive would own or have a licence to use the work in its holdings (this could be

⁴ Section 29.24, An Act to amend the Copyright Act. Statutes of Canada. 2012, c 20.

specified in donor agreements and its relationship with the institutions it serves). Although there is an exemption for archives to make copies in the case of obsolescence or technological unavailability⁵, it does not exempt circumventing TPMs to make those copies. Subsection (c) of 29.24 effectively prohibited the circumvention of a TPM for the backup copy process.

Of course, an archive's purpose is not just for preserving its holdings but to enable access to them. Archives will necessarily have to circumvent TPMs to make copies of their digital holdings (as I will show when I discuss TDRs). Providing access is just the opposite of subsection (d).

Given the present state of common digital technologies, it's increasingly likely that some form of TPM might encumber records that could be of value to an archive. Here is an example.

As published material, e-books (or other digital documents available from online journals, library electronic lending systems, and commercial providers such as companies like Amazon) are in-and-of-themselves probably not of much interest to archives. But the rise of their use changes something valuable for archives, and the associated technical problems for archives are not isolated to e-books.

Whereas in the past, institutional archives like those serving a university might include the notes of professors in their holdings, these notes may now be produced just in a digital-native format. The reading applications used for e-books enable people to annotate or otherwise insert content of their own in relation to the material of the e-book. Like the professor's notes of the past, these digital annotations could be valuable to an appropriate archive.

Assuming the professor worked with a TPM-encumbered document, the software tool used to read and annotate it might store the notes separately from the source PDF, relying on a relationship between the two files to present them together. Alternatively, the application may have altered the PDF and saved it containing the notes.

Here are the problems with this situation. The professor could donate (or someone else could on his behalf) his notes to an archive. Without the means to legitimately access the notes or if the means no longer exist, performing backup procedures or gaining access for migration processes that are needed for long term preservation, would require circumventing the TPM.

Because archives may not circumvent the TPM under C-11, they may not be capable of properly storing the notes. Even if an archive decided that separating the notes from the corresponding source material wouldn't render them meaningless, it's possible that the archive would need to find a technical means to bypass the permission restriction of either or both the document and reading application.

Additionally, the instability of commercial entities' (e.g. software vendors) business processes or of their ongoing viability becomes a major threat to assuring future generations of the ability to access TPM-encumbered content. The Association of Research Libraries states:

No one really likes technical protection measures, especially as they are employed by copyright owners to "safeguard" digital contents. Members of the information-seeking public justly complain that the use of TPMs may interfere with the exercise of their rights to

5 Section 28 (c), An Act to amend the Copyright Act. Statutes of Canada. 2012, c 20.

access and use copyrighted content fairly.⁶

The exercise of rights has important implications with respect to long term digital preservation. To understand this complaint we should consider why TPMs interfere with our exercise of rights such as the fair dealing⁷ right to make copies, in an archival context.

Next, I would like to examine archival goals and strategies for the long term preservation of digital content, particularly in terms of the role of TDRs. This will enable a better understanding of how TPMs and current copyright law come into conflict with the exercise of our rights. I will address how backups and migration practices form key parts of a TDR strategy.

Long Term Digital Preservation and the Role of TDRs

Archives face inexorable problems with the duty to preserve massive quantities of information, stored on frail digital media. There is at once, the archives' own mandate and its capability to fulfill that mandate. To understand what an archive must be capable of doing, here is a definition of digital preservation.

...the managed activities necessary: (1) For the long term maintenance of a byte stream (including metadata) sufficient to reproduce a suitable facsimile of the original document and (2) For the continued accessibility of the document contents through time and changing technology.⁸

This definition calls the byte stream a reproduction, suggesting that the digital preservation is of some physical record. However, this is true even for digital-native (originally produced in a digital format) content because on each instance of accessing that content, a computer copies the byte stream into memory, thereby reproducing it. Regardless, the archive must support accessibility while mitigating the threat of technological change over time.

The Library and Archives Canada (LAC) mandate serves as a good example for examining where TPMs conjoined with current copyright law, cause conflict. The thrust of LAC's objectives although unique, are not dissimilar to many other archives. In the preamble of the Library and Archives Canada Act it says

WHEREAS it is necessary that

(a) the documentary heritage of Canada be preserved for the benefit of present and future generations;

(b) Canada be served by an institution that is a source of enduring knowledge accessible to all, contributing to the cultural, social and economic advancement of Canada as a free and democratic society;

6 "Briefing: Demystifying Technical Protection Measures (TPMs) in the Library." Association of Research Libraries (ARL). 25 January 2012. <http://www.arl.org/pp/ppcopyright/codefairuse/demystifying-tpms.shtml> 31 October 2012.

7 Section 21, An Act to amend the Copyright Act. Statutes of Canada. 2012, c 20.

8 Jantz, Ronald and Giarlo, Michael. "Digital Archiving and Preservation: Technologies and Processes for a Trusted Repository." Journal of Archival Organization 4:1-2 2008: 195. http://dx.doi.org/10.1300/J201v04n01_10 (30 October 2012).

(c) that institution facilitate in Canada cooperation among the communities involved in the acquisition, preservation and diffusion of knowledge; and

(d) that institution serve as the continuing memory of the government of Canada and its institutions;. . .⁹

Called a “source for enduring knowledge accessible to all,” the act that established LAC is phrased to imply continuing endurance through time (it does not limit the number of future generations). Thus, the archive itself should be capable of offering its holdings to those that would like to examine them. Operating as the *source*, the archive needs to be capable of enabling people to access its holdings.

LAC should not be prevented the *capabilities* to fulfil its mandate. Owning rights (e.g. the rights to access TPM-encumbered content) is of little use if LAC doesn't at least have the capability to carry out its mandate. LAC has little choice but to address content as preserved in our era's pervasive use of digital media. Indeed, LAC developed a digital preservation policy in which one of the principle tenets includes the

Generation of access and service copies. Originals and preservation masters will be preserved in archival storage, while web-friendly versions will be generated for public use.¹⁰

The first important element in this statement is that it explicitly calls out the practice of making copies. The second: that it distinguishes between originals, masters for archival storage, and other versions that the public will access. We should keep in mind that the public expects to engage with archival holdings through electronic means, generally the World Wide Web. The archive must make copies not just for the sake of preservation but for use.

Copying is a strategic and technical requirement of TDRs. Indeed, LAC's digital preservation TDR system requires multiple copies be made.

The original, unaltered bit stream that makes up a digital asset, as submitted by the content creator, will always be retained within LAC TDR as Preservation Master 0. If the original logical file format is at risk of obsolescence, it may be migrated to an alternate format, thus creating Preservation Master 1. All preservation masters exist in multiple copies.¹¹

The procedure for preserving digital assets is to take the original, as LAC receives it, and eventually to migrate it to other formats. I use the term “eventually” (though LAC's policy is written conditionally) because content stored on digital media faces the media's inevitable obsolescence. Obsolescence occurs because the devices that read the media may no longer be produced, systems that operate with those devices cease to exist, or the formats stored on the media are no longer readable by computer systems.

Migration is the periodic transfer of digital materials from one hardware/software

⁹ Library and Archives of Canada Act. Statutes of Canada. 2004, c 11.

¹⁰ Bak, Greg and Armstrong, Pam. “Points of convergence: seamless long-term access to digital publications and archival records at library and archives Canada,” Archival Science Volume 8, Issue 4, (2008) p 283, <http://link.springer.com/article/10.1007%2Fs10502-009-9091-4?LI=true> (30 October 2012).

¹¹ Bak, Greg and Armstrong, Pam. “Points of convergence: seamless long-term access to digital publications and archival records at library and archives Canada,” Archival Science Volume 8, Issue 4, (2008) p 287, <http://link.springer.com/article/10.1007%2Fs10502-009-9091-4?LI=true> (30 October 2012).

configuration to another, or from one generation of computer technology to a subsequent generation. The purpose of migration is to preserve the integrity of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology.¹²

Although the act of migrating is by definition also an act of copying, LAC goes a step further, maintaining *all* preservation masters (original or not) in multiple copies. These copies are part of TDR strategy as I'll now explain.

Trusted Digital Repositories

A TDR is an integral part of archival strategy for long term digital preservation. The Research Libraries Group (RLG) defines it thus:

A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future.¹³

Trusted digital repositories are designed in principle to have a great deal of redundancy, helping ensure the integrity of what they store. But we still have to worry about the technical components of the TDR. This results from the ungraceful deterioration of digital media as well as from mechanical and natural contingencies.

All of our digital recording media require active management in order to avoid problems due to media degradation and failure. The only approach to this generic problem is for the repository manager to put in place policies for routine backups, off-site backup, and the use of mirrored sites or other types of redundancy options to ensure that there is always another digital "place" where one can find the original object. From a digital preservation perspective, redundancy of content is perhaps the most critical consideration.¹⁴

Making many copies of the digitally-stored content and continually migrating those to "upgraded" digital storage mediums ensures (hopefully) that the stored content remains accessible on whatever the present-day's computer systems are.

A TPM that causes an archive to be incapable (technically) of copying content may not be circumvented under current law. I've already identified a situation (the professor's notes) in which TPM-encumbered content would not be copyable and if it's not copyable, it's not migrateable. Such content is thus outside the scope of what archives may properly preserve as holdings in a TDR.

Let us return to considering approaches to long term digital preservation. An archive could acquire all appropriate rights and mechanisms to enable it to access TPM-encumbered content but this is not a practical solution for backup and migration processes. As new technologies, media, or

12 Andre, Pamela Q.C., et al. "Preserving Digital Information Report of the Task Force on Archiving of Digital Information," The Commission on Preservation and Access and The Research Libraries Group. (1996) p 6, <http://www.oclc.org/research/activities/digpresstudy.html> (30 October 2012).

13 Beagrie, Neil, et al. "Trusted Digital Repositories: Attributes and Responsibilities," Research Libraries Group. (2002) p 5, <http://www.oclc.org/research/activities/trustedrep.html> (30 October 2012).

14 Jantz, Ronald and Giarlo, Michael. "Digital Archiving and Preservation: Technologies and Processes for a Trusted Repository." *Journal of Archival Organization* 4:1-2 2008: 209. http://dx.doi.org/10.1300/J201v04n01_10 (30 October 2012).

infrastructure techniques are introduced to the market, there is scant likelihood that all TPM techniques will be transferable or continue to work on the necessary computer systems.

If the TPMs are dependent on a third party, which they invariably are, the third party may not continue to be able to provide the accessible rights, nor might it even continue to exist. Companies and their processes come and go. Relying on the existence of third party rights-holders is a folly that will cripple long term preservation strategies.

So long as a trusted digital repository is necessary for digital preservation practices, we must worry about anything that prevents a TDR from achieving its function. While TPMs exist in the marketplace and digital content is encumbered by them, anti-circumvention of TPMs make TDR strategies ineffective.

Copyright Law, TPMs, and Appraisal

The Canadian Council of Archives (CCA) identifies five categories of appraisal criteria. Some of these criteria mix poorly with copyright law and TPMs on the issue of determining what can be acceptable for long term digital preservation.

Criterion 3.3, Intelligibility of information¹⁵, asks whether the contents of a fonds can be read. This may have been conceived with physical fonds in mind but there's no special requirement limiting it to the tangible. I think it applies equally to intangible, digital content. This criterion for readability exists because archival holdings must be of use to people such as historians, researchers, and increasingly the general public (LAC's mandate made this explicit). Keeping something that cannot be used (is not intelligible), doesn't support archival mandates.

A TPM-encumbrance that restricts accessibility to people in general, likewise implies generally restricted readability, and is thus generally unintelligible. It only becomes intelligible by exception (when appropriate authorization is provided).

Archivists ought to eliminate for consideration, TPM-encumbered content because it fails criterion 3.3 on a general basis. This may appear extreme, especially if an archive gains the means and rights to unlock the TPM. But as I've discussed, those means surely won't persist in the future and with the anti-circumvention law, no legitimate way exists to ensure such fonds' ongoing accessibility nor intelligibility.

Such elimination results not because of some inadequacy in the content but because of the contingencies of the copyright law and the TPM's prohibitive effect. Otherwise, nothing physically, technologically, or intellectually keeps an archivist from acquiring such fonds.

If archivists are not allowed to circumvent TPMs, then widespread use of TPMs among digital records puts archivists at the mercy of those controlling the TPMs. We've seen how past archivists—Hilary Jenkinson in particular—took an approach in which records creators essentially selected what would be destined for preservation, while archivists focused on things like creating finding

15 Canadian Council of Archives "Guidelines for Appraisal Criteria for Non-Institutional Records" in Building a National Acquisition Strategy. Ottawa. Canadian Council of Archives, p. 59.

aids.¹⁶ Post-Jenkinson, archivists determine what is appropriate to acquire. In a sense, the combination of TPMs and anti-circumvention law brings professional archivists back to Jenkinson's days. Only digital content that creators or donors have not encumbered with TPMs, will be appropriate for preservation.

In light of these appraisal criteria, it's worth considering the ways digital content could get a TPM applied to it. I suggest that a TPM could be applied intentionally or passively by its creator or some intermediary.

In the intentional case, the creator would have been aware of applying the TPM and actively made the choice to do so. For example, a financial auditor might apply a TPM to prevent circulation of a report.

In the passive case, the author might not have had the explicit intention to apply a TPM. For example, suppose an official of the foreign service kept a digital record of her memoirs. After she dies, her husband decides that for confidentiality and safe-keeping, he'd store her memoirs using a proprietary application, which applied a form of TPM. Later, he decides to donate them to LAC with the application enabling their access. LAC's capacity for preserving these memoirs (entailing back-up and migration) becomes subservient to the application and the vendor that developed it.

LAC (or any archive) would of course have to reflect on its acquisition policy and appraise the materials accordingly. The donation might be a good choice for preservation but the likelihood of becoming inaccessible (hence unintelligible) could make it not worth acquiring.

Conclusion

I've shown why archives must be capable of copying their digital holdings to enable long term digital preservation and access strategies. This capability underpins the functioning of TDRs and informs professional archival duties such as appraisal. Yet, disallowing the circumvention of TPMs prohibits archives from copying their digital holdings.

The law gives a nod to the impermanence of businesses and eventual obsolescence of technologies. Whether or not one agrees with the overall copy restrictions in the law, the law recognizes that archives (also libraries, museums, and similar institutions) have a unique position with duties that require a very permissive degree of rights to make copies. It permits archives to ensure the future availability of our cultural heritage by exempting them from some standard copyright rules, but that is not good enough. Bill C-11 persists with the exemption for copying but adds a restriction against the act of circumventing a TPM; an act which is necessary in order to perform the copying. C-11 confounds itself.

The commercial and technological environment that puts TPMs to use, makes present copyright law problematic in its application to long term digital preservation processes. With respect to digital content, archives like LAC have been barred from fulfilling their mandates to be a source of knowledge that anyone may access.

16 Ridner, John. From Polders to Postmodernism: A Concise History of Archival Theory. Duluth, Minnesota: Litwin Books, LLC, p.147.

Bibliography

An Act to amend the Copyright Act. Statutes of Canada. 2012, c 20.

Andre, Pamela Q.C., et al. "Preserving Digital Information Report of the Task Force on Archiving of Digital Information," The Commission on Preservation and Access and The Research Libraries Group. (1996). <http://www.oclc.org/research/activities/digpresstudy.html> (30 October 2012).

Bak, Greg and Armstrong, Pam. "Points of convergence: seamless long-term access to digital publications and archival records at library and archives Canada," Archival Science Volume 8, Issue 4, (2008) pp 279-293, <http://link.springer.com/article/10.1007%2Fs10502-009-9091-4?LI=true> (30 October 2012).

Beagrie, Neil, et al. "Trusted Digital Repositories: Attributes and Responsibilities," Research Libraries Group. (2002). <http://www.oclc.org/research/activities/trustedrep.html> (30 October 2012).

"Briefing: Demystifying Technical Protection Measures (TPMs) in the Library." Association of Research Libraries (ARL). 25 January 2012. <http://www.arl.org/pp/ppcopyright/codefairuse/demystifying-tpms.shtml> (31 October 2012).

Canadian Council of Archives "Guidelines for Appraisal Criteria for Non-Institutional Records" in Building a National Acquisition Strategy. Ottawa. Canadian Council of Archives, p. 53-61.

Jantz, Ronald and Giarlo, Michael. "Digital Archiving and Preservation: Technologies and Processes for a Trusted Repository." Journal of Archival Organization 4:1-2 2008: 193-213. http://dx.doi.org/10.1300/J201v04n01_10 (30 October 2012).

Library and Archives of Canada Act. Statutes of Canada. 2004, c 11.

Ridner, John. From Polders to Postmodernism: A Concise History of Archival Theory. Duluth, Minnesota: Litwin Books, LLC, p.143-161.