March 12, 2021

Via email: copyright-consultation-droitdauteur@canada.ca

Re: Comments in Response to "Consultation paper on how to implement an extended general term of copyright protection in Canada"

Adobe Inc. ("Adobe") submits the following comments in response to the public consultation, launched by the Honourable François-Philippe Champagne, Minister of Innovation, Science and Industry, and the Honourable Steven Guilbeault, Minister of Canadian Heritage, to consider whether to adopt accompanying measures to mitigate the potential implications of the extension of Canada's copyright term under the Canada-United States-Mexico Agreement (CUSMA) from 50 to 70 years.

I. BACKGROUND

Adobe has been setting the standard for digital experiences since the company's founding in December 1982 and is a global leader in digital experience and digital media solutions. Adobe offers products and services that allow creatives to create, deploy, and optimize digital content. We pride ourselves on providing our users, who range from emerging artists to global brands, with the tools they need to design and deliver exceptional digital experiences.

In addition to our innovative products and services, Adobe is collaborating with leaders across a variety of industries in the Content Authenticity Initiative ("CAI").[1] The mission of the CAI is to develop an open standard that will provide a secure layer of tamper-evident attribution data to photos, including the author's name, location, and edit history.[2] While the main goal of the CAI is to provide an open standard to create tools so that consumers can better understand the content they view online and have greater confidence to evaluate its authenticity, attribution matters for another critical reason: providing credit to creators for their work. Think of it as a simple equation: Exposure (for a creative work) plus attribution (so people know who created it) equals opportunity (for more collaborations or jobs). With a tamper-evident way to seamlessly attach a creator's name to their creative work, it can go viral, and the creator will still get credit.

Strong copyright laws are important to Adobe and our larger communities, which includes the creatives who use our tools to produce copyrightable works. Adobe supports the harmonization of copyright laws. We believe that harmonization of Canada's term with that of the US, the UK, the EU, and other countries will have many benefits, including facilitating trade by reducing barriers. However, extending the term of copyright makes it even more important to ensure that appropriate and clear exceptions are in place to permit uses that are beneficial for creativity and society, so as not to stifle innovation. One of these uses that is of concern to Adobe and our users is the ability to use data, databases, and data sets (which may include copyrighted materials) in the development of artificial intelligence ("AI") systems.

As a highly innovative company, Adobe has embraced AI to enhance user experiences and empower our customers to create and innovate. Adobe's business is composed of three cloud-based solutions – Adobe Creative Cloud, Adobe Document Cloud, and Adobe Experience Cloud – and all of them use AI to deliver

[1] https://contentauthenticity.org

[2] https://blog.adobe.com/en/publish/2019/11/04/content-authenticity-initiative.html#gs.v5fpha

intelligent features across Adobe products.[3] As a core element of our platform, we use AI, machine learning, and deep learning capabilities to tackle today's complex experience challenges. Adobe has many examples of innovations powered by AI that are used by creative professionals, knowledge workers, and enterprise customers. Adobe's AI offering, which we refer to as "Adobe Sensei," powers intelligent features within each of our cloud offerings to dramatically improve the design and delivery of digital experiences, blending the art of content with the science of data.

In the creative space, Adobe is focused on using AI to deliver useful tools to the creative professionals who use our products every day. For instance, when graphic designers use an Adobe tool to create an image in a creative work, they can use AI-assisted search in Adobe Stock to find stock images on criteria like emotions (search for "happy") or abstract concepts (search for "jump" or "love").[4] Using similar AI tools, filmmakers can review footage and have edits suggested to them, and other creative professionals can have Adobe Photoshop customized to their skill level and work focus.[5] For this reason, Adobe maintains a unique position – we use AI to create tools for creatives who then use those tools to produce copyrightable works. Adobe understands the needs of both the creative community and the technology community and knows the importance of striking the right balance that will protect both creativity and innovation.

AI opens creative possibilities and brings significant financial benefit, but truly realizing these possibilities and benefits requires reasonably unrestrained access to large quantities of data[6] to train AI systems that enhance products across our business and elevate our users' ability to more efficiently innovate and unleash their creativity. Such training may involve the making of a temporary copy of the training data, which may include copyrighted materials. Uncertainty around the legality of using copyrighted materials to train machine learning models may stifle creativity and innovation and make it difficult for AI to thrive. Extending the copyright term serves to exacerbate this issue since it would reduce the amount of material that is in the public domain that could be used for training. Therefore, we urge, as part of the accompanying measures adopted to mitigate the potential impacts of the extended term, that the law be clarified to provide better guidance and a clear path for AI researchers to use copyrighted materials for the limited purpose of training AI models, without the fear of liability for copyright infringement.

Adobe is interested in these topics because we view the world through a lens of global digital citizenship – the responsible use of technology for the good of creatives, our customers, and society.[7] Our comments will discuss (1) why AI is critically important in today's world, (2) why specific guidance regarding the ability to use copyrighted materials for training AI is crucial to provide the certainty needed for AI to flourish, and (3) why allowing the use of copyrighted materials for training AI models serves the goals of copyright, promotes innovation, and significantly benefits society.

---

[3] https://theblog.adobe.com/amplifying-human-creativity-with-artificial-intelligence/
[4] https://theblog.adobe.com/adobe-sensei-ai-ml-image-search-discovery/
[5] https://theblog.adobe.com/demystifying-and-democratizing-artificial-intelligence/
[6] Note that when we use the term "data" we mean to include copyrighted materials, such as images, that may be included in datasets used for training algorithms to create machine learning models.
[7] https://blog.adobe.com/en/publish/2021/02/17/adobe-unveils-new-ai-ethics-principles-commitment-responsible-digital-citizenship.html#gs.v5g9wp

II. COMMENTS

A. Artificial intelligence is critically important in today's world.

Canada, along with other global leaders, has recognized the importance of AI. In the 2017 Federal Budget, the Government of Canada recognized that "Canadian talent and ideas are in high demand around the world – but activity needs to remain in Canada to harness the benefits from artificial intelligence."[8] Canada is a leader in the AI field. Montréal has the highest concentration of researchers and students of deep learning in the world, and Toronto has the highest concentration of AI start-ups in the world. In addition, there has been significant investment in AI by foreign companies in recent years. AI affects all sectors of business and society and is quickly becoming the way we understand and process the vast amounts of information that is available today. However, AI is highly dependent on access to large quantities of high-quality data, and limiting access to data too much may cripple Canada's ability to keep up with other countries innovating in AI and deprive people of AI's vast benefits.

A number of other countries who have grasped the importance of AI and the role of data have adapted their copyright rules to facilitate the development of AI. These are often referred to as text and data mining (or TDM) exceptions. Japan recently amended its Copyright Act to add exemptions for the use of copyrighted works for machine learning. Article 30-4 allows use of a copyrighted work "… where such exploitation is not for enjoying or causing another person to enjoy the ideas or emotions expressed in such work."[9] And Article 47-4 facilitates machine learning by allowing incidental copies of copyrighted works.[10] Other countries, such as Singapore, Australia, China, and Thailand, are also looking to update their copyright laws to further facilitate machine leaning. The EU has recently adopted limited TDM exceptions and continues to explore further refinement of these exceptions.[11] As noted above, Canada is currently a leader in AI, but it could quickly fall behind if its policies hinder or limit access to data, pushing AI development to countries that have implemented data mining exceptions.

B. Although the current laws may support the legality of processing copyrighted material for the purpose of training AI models, specific guidance is crucial for providing the certainty needed for AI to flourish.

As the process of training a machine learning model can involve temporary reproductions of the training data, the question arises as to whether that temporary copying constitutes an infringement of copyright under Canadian law (which prohibits making a reproduction of an original work during the term of the copyright without the owner's consent).[12] No Canadian court has assessed whether any temporary copying for machine learning amounts to a reproduction, but if it does, then machine learning involving copyrighted materials would be infringing unless an exception applies.

There are two exceptions to infringement under Canadian copyright law that may support the right to use copyrighted material to train AI systems. The first is the fair dealing provision, and the second is the Tech

---

[8] House of Commons, Minister of Finance, *Building a Strong Middle Class* (March 2017) at 103.
[9] Copyright Act (Japan) Article 30-4. https://wipolex.wipo.int/en/text/504411
[10] Id. Article 47-4.
[11] See, e.g., Association of European Research Libraries, Europe Needs a Broad and Mandatory TDM Exception (Nov. 13, 2018) https://libereurope.eu/blog/2018/11/13/europe-needs-a-broad-mandatory-tdm-exception/
[12] *Copyright Act*, RSC 1985, c C-42, ss 3(1), 27(1).

Exception. Both provide promising arguments that would allow such training, but neither provide enough certainty for data scientists to feel comfortable relying on them.

1. Fair Dealing Exception

Even if copies made in the training process are found to implicate the reproduction right, there are strong arguments that the fair dealing provision may apply to exempt any reproductions made in the training of AI from constituting infringement. The fair dealing exception, which is provided under Section 29 of the Copyright Act, provides that "[f]air dealing for the purpose of research, private study, education, parody or satire does not infringe copyright."[13] To be covered by the fair dealing exception, a defendant must show that the dealing both falls under one of the enumerated purposes listed in the statute and is "fair."

There are good arguments that reproductions created in the course of training an AI model would fall under "research" and therefore be covered by one of the enumerated purposes in Section 29. Although there are no Canadian cases that specifically address training machine learning models, caselaw in analogous areas can support this argument. For example, in *SOCAN v. Bell*, the Supreme Court of Canada found that a commercial internet site that allowed consumers to preview a musical work before making a purchase fell within the purpose of "research" under the fair dealing analysis and held that there must be a generous interpretation of the fair dealing purposes, including "research."[14]

And in the recent case of *Stross v. Trend Hunter*,[15] in which the defendant performed data analysis using AI to measure a user's interaction with content, the Federal Court noted that "research" is to be given a "large and liberal" meaning and found that the defendant met the low threshold necessary to demonstrate that the dealing was for research purposes. These cases provide strong arguments that the threshold purpose requirement is met for fair dealing when copyrighted materials are used merely to train an AI model. However, because no cases yet have specifically addressed the issue of fair dealing in the context of training machine learning models, there is still some uncertainty, and companies and research institutions may be hesitant to rely on fair dealing, at least without specific and clear guidance.

Moreover, there is even more uncertainty when looking at the issue of whether the dealing is "fair." The Supreme Court of Canada has set out six factors that guide the assessment of whether the dealing is fair: (1) the purpose of the dealing, (2) the character of the dealing, (3) the amount of the dealing, (4) alternatives to the dealing, (5) the nature of the work, and (6) the effect of the dealing on the work. While Adobe believes there are strong arguments that the use of copyrighted materials to train machine learning models meets the fair dealing test, the lack of relevant case law and the complex and subjective factual nature of the inquiry makes it difficult to rely on fair dealing.

With the uncertainty inherent in applying preexisting fair dealing factual analysis to new technology such as AI, fear of infringement claims often drives companies without clear access to good data to use low quality data, or insufficient amounts of data, to train models for AI. This affects the quality of the models and the ability for all companies to compete equally. As discussed in Section C below, outdated and insufficient data also contributes to algorithmic bias. If Canada does not provide researchers and companies who work in AI

[13] Section 29 Copyright Act, RSC 1985, c C-42
[14] *Socan v. Bell*, 2012 SCC 36 at para 27.
[15] *Stross v. Trend Hunter*, 2020 FC 201.

with clear rules that permit the development of strong models, which are the building blocks of AI, the country will be at a disadvantage in the development of AI as compared to other countries that have clarified the law to clearly permit such uses.

2. The Tech Exception

In addition to fair dealing, there is another exception to infringement under Canadian copyright law that may apply to the training of machine learning models. Under the Copyright Act, temporary reproductions for technological processes are also exempt from constituting copyright infringement. This exception (referred to as the "Tech Exception") is provided under Section 30.71 of the Copyright Act and provides that it is not an infringement of copyright to make a reproduction of a work or other subject-matter if:

(a) the reproduction forms an essential part of a technological process;

(b) the reproduction's only purpose is to facilitate a use that is not an infringement of copyright; and

(c) the reproduction exists only for the duration of the technological process.

There are good arguments that the Tech Exception could provide a valuable exception to copyright infringement that applies to the type of copying made in training machine learning models. However, since coming into force in 2012, the Tech Exception has received no judicial treatment and has only been analyzed by the Copyright Board in a single decision unrelated to the AI context. Accordingly, there is little clarity on the scope of this exception. Some of the holdings in the aforementioned Board decision could support the application of the Tech Exception to cover machine learning. For instance, the Board gave examples of what it considered to be a purpose that facilitates a use that is not an infringement of copyright, which included buffer copies in a photocopier or printer, or the reproductions made in volatile memory of a computer. This suggests that the temporary copies made to train machine learning models may likely be protected by the Tech Exception (at least if the copy was not stored after the training).[16] But with so little to go on, it is still difficult to know with any certainty how these requirements may ultimately be interpreted in the context of machine learning, and therefore, it does not provide the clarity needed for AI to thrive.

Although we believe both the fair dealing provision and the Tech Exception should cover the use of copyrighted materials to train machine learning models, uncertainly remains. Fair dealing is ultimately a fact-driven analysis, and the case-by-case nature of this analysis provides little comfort to AI developers looking for clear guardrails. Similarly, with respect to the Tech Exception, the lack of clear precedent applying the exception to AI systems and the lack of judicial treatment create considerable uncertainty.

C. Providing clear guidance that allows the use of copyrighted materials for training AI models serves the goals of copyright, promotes innovation, and significantly benefits society.

Providing guidance that establishes a clear right to use copyrighted materials to train machine learning models is consistent with the goals of the Copyright Act, which are to promote the encouragement and dissemination of works of the arts and intellect, and to provide a just reward for the creator.[17] This guidance

[16] *Ibid.* at para 189.
[17] *CCH v LSUC*, 2004 SCC 13 at para 23.

could either take the form of adopting text and data mining laws, which clearly cover the copies made when training machine learning models, or clarifying existing law, such as the applicability of the fair dealing and tech exceptions or clarifying that the temporary copying needed to facilitate such training is not a "reproduction" under the Copyright Act. Such steps would encourage further innovation in and around the science of AI and would promote the encouragement and dissemination of new works of art by providing tools, such as those made available by Adobe, that assist artists of various types to innovate in their field.

Providing this clarity would also ensure sufficient access to large amounts of high-quality data for training AI, which is important to combat algorithmic bias. AI needs to ingest a lot of data. If the data is incomplete, outdated, or biased, AI can exacerbate problems of bias."[18] The growth of AI has brought an increase in examples of AI systems reflecting or exacerbating societal biases. This threatens to overshadow AI's technological gains and potential benefits. These biases range from racial bias in facial recognition technologies to the embodiment of sexist stereotypes in natural language processing.[19]

Copyright laws can exacerbate this problem because companies often feel compelled to use readily available, low quality data, or insufficient amounts of data, to avoid risks of infringement, even though that data and content may embody bias. Amanda Levendowski, a Clinical Teaching Fellow at NYU, wrote an interesting article that explores the ways in which copyright law can negatively influence the quality of AI, and how fair use (the US equivalent to fair dealing) might be part of the solution.[20] Her article notes that "Copyright law causes friction that limits access to training data and restricts who can use certain data. This friction is a significant contributor to biased AI."[21] The article goes on to say that "If we hope to create less biased commercial AI systems, using copyright-protected works as AI training data will be key."[22] If AI researchers are forced to use older, lower quality data and content for training models, this type of bias will persist. For example, public domain data is readily available, but these are generally works that are older which, as Ms. Levendowski notes in her article, "was [created] back when the 'literary cannon' was wealthier, white and more Western than it is today."[23] If the term of copyright is extended, the material available under the public domain, which could freely be used to train machine learning models, will be 20 years older, and thus more likely to exacerbate bias even further. In order to counteract this, it will be important for there to be a clear right to use copyrighted materials so that researchers aren't limited to using what's in the public domain. In light of these benefits, there should be clear guidance that the temporary processing of copyrighted materials for the purposes of training AI systems is not infringing.

Adobe vigorously supports the ability of creative professionals to protect their work through copyright and to realize economic value from their creative works. Adobe has built an industry enabling creative people to express themselves and has generated thousands of jobs and economic opportunities for creatives to produce excellent work. AI enhances our ability to deliver on the promise of creative tools and our creative

---

[18] EXEC. OFFICE OF THE PRESIDENT, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE (2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

[19] One toolkit that is used to represent how words relate to each other (i.e., Beijing is to China as Warsaw is to Poland) was found to be riddled with gender bias, including a statement that "man is to computer programmer as woman is to homemaker."

[20] Levendowski, Amanda, How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem (July 24, 2017). 93 Wash. L. Rev. 579 (2018).

[21] Id. at 589.

[22] Id. at 621.

[23] Id. at 615.

professionals appreciate the technological advantages AI can bring them. The non-expressive use of media to train our AI and, in turn, improve tools for creative professionals is complementary and supportive of their economic model.

Humans can read and consume copyrighted materials to extract factual information and gain insights and learnings from them, and this ingestion by the human brain does not constitute infringement. AI systems should be able to do the same. The fact that a temporary copy of copyrighted material is made so that a computer can do the same thing as humans do, should not change this. Where someone (or something) is not using the expressive content of a work, but is merely using it to extract factual information, and no copyrighted materials end up in the final product, there is no harm to the economic value of the work, and that use should clearly be permitted without the authority of the author of that content.

III. CONCLUSION

We hope our comments have illustrated the value and importance of AI technology, its furtherance of the goals of copyright law, the benefits to society of securing access to sufficient quantities of high-quality data for machine learning, and the limited, non-expressive and non-harmful nature of the use of copyrighted materials for this purpose. The need to clarify that the use of copyrighted materials for this purpose is permitted will be even more important with the extension of the copyright term, which will further narrow the scope of materials available in the public domain. Therefore, we urge that the law be clarified to provide guidance and a clear path for AI researchers to use data, databases, and data sets (including copyrighted materials) for the limited purpose of training AI models, without the fear of liability for copyright infringement. We believe this is essential for Canada to maintain its position as a leader in the field of AI and for the public to fully enjoy the benefits from this transformative technology.

Adobe appreciates the opportunity to share our thoughts on the importance of an approach that protects the rights of copyright owners while promoting robust AI development. We realize that we have a unique perspective in that we understand the needs of both the creative community and the technology community and the importance of striking the right balance to protect both creativity and innovation. We would welcome the opportunity to further discuss these issues with you or to serve as a resource.