

Government of Canada: Consultation paper: Consultation on Copyright in the Age of Generative Artificial Intelligence

Survey Questions:

I. Text and Data Mining (TDM)

a. What would more clarity around copyright and TDM in Canada mean for the AI industry and the creative industry?

Due to the potential for generative artificial intelligence (GAI) to train on, and output, data that is covered by intellectual property (IP) protections, ACT | The App Association agrees that GAI platforms and their users require clear policy guidance on how Canada will evaluate liability associated with text and data mining (TDM) to train GAI platforms and with the use of its potentially IP infringing outputs by a platform user.

For small business software developers, such as App Association members, GAI platforms are advanced technical tools that are invaluable to their creative and innovative processes that save costs and time by streamlining repeatable tasks and optimizing limited resources. Software developers have been using AI tools that heavily rely on human intervention to deliver a desired output for years, but AI tools are reducing, and are very likely to continue to reduce, the need for human instruction and intervention. For example, software developers use AI to improve the software coding process; support training a new generation of strong software developers; complete repeatable tasks; detect common mistakes, issues, and risks in the software development process that would otherwise require manual interventions; and run quality assurance checks that reduce the chance of human bias and error. GAI tools are different from traditional AI tools in that they have an independent process that mimics cognition to develop new outputs. For software developers, GAI platforms not only seamlessly predict and complete lines of code, but they produce outcomes by training on various resources, including internet archives, data laundering sources, and even platform user inputs.

Much like a human brain, AI systems train on data to understand patterns and create rules that help them make decisions. Like a human brain, GAI might output, in part or in whole, an image, writing, wordmark, or other IP protectable work that it was trained on. Where a GAI system does produce an infringing work, App Association members are concerned about the liability of a GAI user that unintentionally incorporates the infringing output in their final product. The law is still developing on this issue, and the outcome of ongoing proceedings in Canada (and in other important jurisdictions) is likely to influence the speed of innovation across sectors increasingly looking to operate as efficiently as possible. For example, if it is revealed that GAI platforms are, in some cases, effectively copy-and-pasting data, the effect can be catastrophic because of the amount of work that has already been conducted using GAI systems.

App Association members operate with minimal resources and are most acutely harmed by unpredictable copyright outcomes related to liability. The interdependent relationship between a GAI platform and its users is important, and today GAI platforms are essential for small business technology developers to compete. As this type of relationship grows, clear, reliable, and strong guidance for GAI platforms and its users to advance an understanding of the legal bounds of TDM activity will be imperative to protecting IP, as well as supporting investment and

innovation. Since the law for GAI and copyright is still developing, we urge the Government of Canada (GOC) to take a fact-finding role at this time, and to make its findings public to inform the debate about future policy changes needed to address GAI and copyright, should any be needed. Canadian policy changes should be based on well-demonstrated systemic problems (and not edge use cases or hypotheticals that exemplify possible uses and capabilities of AI outside what we presently understand). As Canadian courts examine and define the boundaries of using AI, we urge the GOC to seek industry input on an ongoing basis to inform the development of detailed and guidance, and potential changes in policy, related to copyright and the use of GAI.

b. Are TDM activities being conducted in Canada? Why is it the case or not?

Yes. In general, researchers and companies have been using TDM practices to analyze large data for years and, when done in good faith, these practices are an economical and efficient way to leverage information. TDM activities related to GAI technology are occurring both in Canada and around the world, and we support GOC evaluating and publicizing its findings as to the extent and nature of TDM activities it is observing in Canada to inform the public debate.

c. Are rights holders facing challenges in licensing their works for TDM activities? If so, what is the nature and extent of those challenges?

Currently, there is no standard licensing practice for TDM activities as related to AI, and specifically, GAI. Yet, IP rights holders have certain tools at their disposal to protect their public-facing IP-protected works, including ones that themselves use AI. For example, rights holders may deploy exclusion protocols (ex. robots.txt) that notify search engine crawlers and miners that they are prohibited from accessing certain data.

d. What kind of copyright licenses for TDM activities are available, and do these licenses meet the needs of those conducting TDM activities?

TDM licenses exist as written agreements between private parties or as embedded codes in a website, digital work, or digital representation of a work. In practice, there is no standard TDM license or common practice. GOC should evaluate the ecosystem and, based on systemic issues and ambiguities identified, develop appropriate guidance to ensure that copyright licenses for TDM activities do not strip the ability for those activities to benefit the public good.

e. If the Government were to amend the Act to clarify the scope of permissible TDM activities, what should be its scope and safeguards? What would be the expected impact of such an exception on your industry and activities?

While it is imperative that copyrighted works are strongly protected under Canadian copyright law, it is equally important that innovation and creativity is not stunted by restrictive laws that curtail the benefits of essential and advanced technological tools, including AI tools. TDM activities are used to research, inform, educate, and, as a result, expand the ability of stakeholders to grow critical and new markets. Therefore, GOC should ensure that a “fair dealings” analysis is applied to TDM activities on a case-by-case basis. With the advent of GAI, courts have yet to carve out and define unique nuances of GAI platforms and their TDM

activities. Other long-standing industries have been using TDM tools in order to advance, discover, and build upon existing works. GOC should not alter existing, or create new, copyright laws or regulations unless systemic problems are clearly identified; instead, GOC should look to generate helpful guidance addressing TDM activities and potential copyright infringement in alignment with relevant Canadian court decisions.

f. Should there be any obligations on AI developers to keep records of or disclose what copyright-protected content was used in the training of AI systems?

Since AI systems may store mined content in their internal database, it is possible to require their developers or owners to keep records of content that is copyright protected. However, disclosure requirements may be taxing dependent on the amount of content trained on and stored in the AI system's internal database. We urge GOC to hold off on creating strict disclosure requirements for AI systems at this time; instead, GOC should partner with the private sector to encourage that AI systems should take reasonable steps to ensure that they are not infringing upon copyright protected content.

If an AI provider reasonably knows that copyrighted content was infringed by the AI or its user, the provider should take reasonable steps to remedy the situation, and such reasonable steps could be outlined in public guidance laid out by GOC. The App Association commits to collaborating with GOC to develop and socialize copyright and AI guidance for AI providers and its users.

g. What level of remuneration would be appropriate for the use of a given work in TDM activities?

N/A

h. Are there TDM approaches in other jurisdictions that could inform a Canadian consideration of this issue?

Some countries have rigid TDM restrictions that are antithetical to Canadian law, many of which create unnecessary hurdles to innovation that are not in the public interest. We urge GOC to rely on its own legal regime and the opinions of its stakeholders to determine proper approaches to TDM activities, and to take action to address demonstrated harms/problems for the Canadian economy and copyright system.

II. Authorship

a. Is the uncertainty surrounding authorship or ownership of AI-assisted and AI-generated works and other subject matter impacting the development and adoption of AI technologies? If so, how?

While Canadian copyright law does not explicitly state that AI-assisted and AI-generated works require human intervention to be copyrightable, courts have interpreted copyright "authorship" to

require a natural person exercising skills and judgement to create the work. This understanding is consistent with significant jurisdictions, including the United States, EU, and UK. While these jurisdictions specifically reject works that are wholly constructed by AI, Canadian courts have not made this distinction. In fact, the Canadian Intellectual Property Office (CIPO) has accepted AI as a co-author to an artistic work in at least one registration. “SURYAST” (Artistic) Ankit Sahni, Can 1188619 (1 December 2021) registered. While CIPO registrations are not precedential due to a lack of substantive review, this system has, at times, gone against the understanding that Canadian copyright protection should only be awarded to natural persons. GOC should work to provide stakeholders with stronger guidance on elements of copyright protection, including authorship.

b. Should the Government propose any clarification or modification of the copyright ownership and authorship regimes in light of AI-assisted or AI-generated works? If so, how?

GOC should provide stakeholders with clear guidance on what works can be registered in compliance with understandings from Canadian jurisprudence. Currently, the CIPO has accepted at least one registration for which AI was enlisted as a co-author, while judicial interpretation suggests that only natural persons can be authors of a copyrighted work. GOC is strongly advised to provide stakeholders with clarity on this matter unless and until courts state otherwise.

c. Are there approaches in other jurisdictions that could inform a Canadian consideration of this issue?

We encourage GOC to monitor jurisdictions like the United States, EU, and UK that are currently gathering stakeholder input and considering further guidance for their stakeholders. We note our community’s concerns with the U.S. registration system as it pertains to AI, which is presently unworkable for its stakeholders. As it stands, the U.S. Copyright Office (USCO) requires applicants to disclaim AI-generated content that is more than *de minimis*. This standard for registration does not consider the various uses of AI across different forms of expression, nor does it ask the more important question of whether and how much human authorship was applied in the development of a work seeking copyright registration. The modified registration requirement has caused the USCO to make a series of misguided judgements on whether a work should be properly registered based on its determination of what constitutes AI-generated content. Based on recent registration decisions, the USCO is unconvinced that an individual prompting GAI to produce an output amounts to human input. This inflexible standard does not account for the technological advancements that have already taken place and will take place in the future nor does it inquire into the narrowness of the prompts and resulting AI-generated outputs. As AI advances to undertake more complex analysis, it will still never fully operate without human intervention.

If the CIPO does implement a more substantive review process for copyright registration, we urge the agency to avoid drawing similar rigid conclusions that mischaracterize the role of AI as a creator rather than a tool for creation, which requires human intervention to operate. Under a more substantive review process, the CIPO should consider the amount and level of “human authorship” rather than the amount of content generated by GAI when making registration determinations. We urge the CIPO to avoid broad tests that do not adapt with the onset of advanced and emerging technologies. Human authorship should be the threshold issue when

determining copyright registration and should be determined on a case-by-case basis. Canada's copyright registration system should not disregard the importance of allowing its stakeholders to utilize advanced tools, like GAI, that incent copyrightable intellectual and creative human expressions.

III. Liability

- a. Are there concerns about existing legal tests for demonstrating that an AI-generated work infringes copyright (e.g. AI-generated works including complete reproductions or a substantial part of the works that were used in TDM, licensed or otherwise)?**

Some stakeholders have expressed misguided concerns about the effectiveness of existing copyright laws and tests (i.e., fair dealings). These concerns are born out of growing pains caused by novel copyright disputes surrounding the use and TDM activities of AI. We urge GOC to refrain from modifying current copyright laws or developing new regulation without assessing all stakeholder concerns and allowing courts to make determinations on these issues. And in the interim, we encourage the GOC to empower and require the CIPO to provide stakeholders with a working guidance on copyright implications of AI.

- b. What are the barriers to determining whether an AI system accessed or copied a specific copyright-protected content when generating an infringing output?**

The current barrier to determining copyright infringement by AI systems is the lack of guidance framework for AI providers and its users. If GOC can provide stakeholders with a clearer understanding on how AI impacts the infringement analysis and what steps are necessary to reasonably mitigate such infringement, it will be in a better position to assess if such infringement has occurred.

In its guidance, GOC should consider how the scope of training datasets might affect the infringement analysis. Whether developing copies of training datasets for AI is fair dealings will depend on the type of AI and its training process. GAI is different from its predecessors in the way it gathers, stores, and processes data. However, courts across jurisdictions have provided significant guidance to show that the extraction and use of large datasets, including copyrighted material, for learning purposes leans towards a finding of fair dealings (or an equivalent analysis) in many circumstances. Still, a fact-sensitive component for a fair dealings analysis is imperative to maintaining a just and successful copyright system. We encourage GOC to support a flexible guidance on a fair dealings analysis that makes clarifications to the fair dealings analysis as it relates to TDM activities for AI systems.

- c. When commercialising AI applications, what measures are businesses taking to mitigate risks of liability for infringing AI-generated works?**

We are not aware of specific strategies that AI providers are utilizing in order to avoid liability for copyright infringement of AI-generated works, who are still unclear about when TDM activities

constitute infringement. For this reason, we urge GOC to provide a working guidance on for AI providers and their users to take reasonable steps to mitigate and avoid copyright infringement.

d. Should there be greater clarity on where liability lies when AI-generated works infringe existing copyright-protected works?

As discussed above, GOC should consider developing a working guidance addressing AI-generated works and Canadian copyright law. Such an AI guidance should not be final or narrow while stakeholder concerns and court disputes are still being considered.

e. Are there approaches in other jurisdictions that could inform a Canadian consideration of this issue?

Canada should develop its own liability regime based on the concerns of its stakeholders and developing jurisprudence on AI. Other countries are similarly gathering input and considering what actions to take now, which can also be informative to GOC's efforts. At this time, we would not recommend that GOC consider the liability regime of any one country, even if legal and policy mechanisms are similarly structured to Canada's copyright framework.