

Canadian Civil Liberties Association

Daniel Konikoff (Privacy, Technology & Surveillance) & Laura Rourke (Pro Bono Students Canada, University of New Brunswick)

### Consultation on Copyright in the Age of Generative Artificial Intelligence

In October 2023, the Government of Canada, as part of a broader effort to modernize Canada’s legislative frameworks to keep pace with technological developments, launched a consultation on the implications of generative artificial intelligence (AI) for copyright. The consultation solicits views on four key areas: 1) the technical aspects of AI; 2) how copyright frameworks apply to text and data mining (TDM) activities; 3) how copyright frameworks should apply to AI-assisted and AI-generated content, and; 4) concerns that AI raises for copyright infringement and liability.

The following responses were submitted to the Ministry of Innovation, Science, and Economic Development (ISED) through an online consultation portal. Our submission addresses some of the concerns that generative AI raises for artists and creatives—concerns that begin with impacts on intellectual property and copyright, but which trickle down to impacts on artists and creatives’ livelihoods, as well.

#### *Text and Data Mining*

TDM is a necessary step in training generative AI models. But in the case of generative tools that spit out poetry, music, and illustrations, the “T” and “D” that need to be “M’d” are actually human-made works of art, work that has sprung from the boundless creativity of human minds. There is already something dispiriting about the way generative AI can flatten pre-existing creative works into “text” or “data” for harvesting, and just because these models need lots of data does not mean that that data should be mined with such little regard for the creatives whose work fuels it.

Standardizing TDM usage rights is thus a crucial step toward respecting copyright holders of creative works. These platforms should obtain licenses, pay licensing fees, and bear primary liability for copyright infringement. If platforms lack these licenses, or if no prior authorization is granted, then platforms should be liable for the unauthorized acts of communicating these copyright-protected works to the public, including making them available to the public. Further, standardization of usage rights can create helpful benchmarks for transparency and accountability in how companies approach TDM. This means establishing effective complaint and redress mechanisms along with human oversight of what may otherwise be an automated process. Establishing strict standards of disclosure for users and effective complaint processes for copyright holders can protect individuals when TDM is occurring, and if copyright holders’ work is being used unlawfully.

One such model Canada can look to for protecting rights of copyright holders during TDM is that of English-Corpora.org. In this model, the accessible content has limitations to how it can be used when downloaded. To quote McCracken and Raub (2023), “the vendor manages the limitation of copyright by removing 5% of all the content. Doing this through removing the last 10 of every 200 words, the vendor has created a collection that essentially has no resale value but is still fully valid for linguistic analysis and

research.”<sup>1</sup> English-Corpora's model is appealing for many reasons: it creates safe access to content that does not risk delivering the full licensed text collection; vendors would not need to create their own portal to control access to the data; and patrons would have the tools to work on more than one data set and establish reproducibility. Overall, the goal of whatever regulatory scheme is implemented should further protect the rights of the creator and owner of the copyright.

We acknowledge that in some cases, those who wish to train generative AI models may find it necessary to seek exceptions to copyright protections. However, these exceptions should be limited in scope to protect copyright holders' rights and interests. This is consistent with TDM regimes around the world. Though they vary by jurisdiction, countries like Japan, Singapore, Estonia, and Germany allow copyright exceptions for non-commercial uses and research purposes. Both exceptions balance the interests of users and copyright holders: they understand the benefit and value that generative AI research may create for society, while also ensuring that copyright holders derive financial reward from the use of their work in commercial contexts. To balance the moral rights of the owner and the interests of the developer, ISED should implement regulations that require the developer to recognize the use of the owners' works in non-commercial and research applications.

Another option that could be implemented into the law is a fair use and fair dealing exception that is followed in the United States. This is where the exceptions to the copyright law regarding TDM are determined by a test of proportionality and includes assessing the purpose and character of use, the nature of the copyright, the amount and substantiality of the portion taken during the TDM process, and the effect of the use on the potential market. However, this still needs to consider whether granting this exception to commercial uses would grant a disproportionate amount of control to the technology sector, giving them inordinate power over the creators of copyrighted works. There needs to be attention given to the economic rights of the owner of the copyright so that the law does not allow too many exceptions that the owner can no longer derive financial reward from the use of their work. Moral rights are also an important consideration in TDM and to balance the rights of the owner with the user of their work, regulation should be implemented that requires the user to recognize the use of the owner's work.

### *Authorship and Ownership*

When it comes to generative AI, there is uncertainty surrounding authorship and ownership for copyright purposes. This uncertainty stems from the separation between types of AI-generated works, and from the difficulty of trying to fit these works into current models of authorship and ownership.

For the first issue, it would be helpful to break up the legal definition for copyright purposes into two categories. One category would be 'works created solely by AI systems that have no human contribution.' Under the current copyright laws, these works would not be protected by copyright. The second category would be 'human works that were made with the assistance of AI.' If following approaches taken by other jurisdictions, these works could be protected by copyright, if it can be proven that a level of 'sweat of the brow' and 'modicum of creativity' are demonstrated in the works by a human being. This means that it would need to be determined in law what level of effort of skill, labour, and

---

<sup>1</sup> McCracken, P. & Raub, E. (2023). Licensing Challenges Associated with Text and Data Mining: How Do We Get Our Patrons What They Need? *Journal of Librarianship and Scholarly Communication*, 11(1), eP15530. <https://doi.org/10.31274/jlsc.15530>

creative pursuit is necessary to fulfill these two categories, in a sufficient manner to allow for copyright protection (This approach is addressed within *Nova Productions v. Mazooma Games [2007] EWCA Civ 219*).<sup>2</sup>

The concept of originality has also been considered internationally when addressing how copyright protection is created for a work. Originality is an important factor for copyrighted material and this concept should continue to be considered when works of AI are being analyzed because without ensuring that originality is present, works of AI would merely be a compilation of publicly available or already copyrighted materials. The case of *Eastern Book Company v. D.B. Modak (2008) 1 SCC 1*, addresses the importance of originality as a guiding principle<sup>3</sup>. This case states that to grant copyright, it would need to be proven that there was expertise<sup>4</sup>. The case of *Rupendra Kashyap v. Jiwan Publishing House Pvt. Ltd*<sup>5</sup>. This court describes how originality could be broken down into the following factors to determine whether its work is original: whether the expression and idea are inherently linked; whether the author applied expertise and effort; whether the least possible level of imagination is present in the work; and whether the resultant work is a product of only work and skill. Both courts in Canada and India have been clear on their view of copyright protection: originality is essential to grant this protection. Without it, there is nothing that the copyright can attach to, and this would be the same result for AI-generated works. That said, the issue of authorship and ownership of AI systems and works still exists and needs to be addressed to answer the question of who will hold the copyright if originality is found in AI-generated works, and consequently who will be held liable for potential issues of infringement.

Although many cases support this traditional view that copyright protection shall only be granted to human-made works, and that our legal ideas of authorship can only exist within this framework, this does not mean that things cannot change. Canada's living tree approach to our legal system supports reconceptualizing definitions of authorship within the realm of copyright to better suit the technological present and the current needs and interests of society.

Redefining authorship to allow for AI-generated works to fit under current copyright frameworks is another approach that can be taken by the Canadian government. For this, Canada could apply the concept of the "work made for hire" doctrine. This doctrine states that "the individual who created the work is always considered to be the author of the work, regardless of whether the work was created in an employment or independent contractor relationship," but that the producer's employer is the first owner of the copyright of works created by an employee in the course of employment (Moskal, 2021).<sup>6</sup> This approach could change how ownership was determined as it would classify AI as employees of whatever company was utilizing the system. Although this doctrine slightly changes the traditional method from granting the ownership of the copyright to the author of the works, to instead granting the ownership of the copyright to the employer of that individual, the work for hire doctrine allows for a more viable solution to issues of infringement and liability that arise with AI-generated works or AI systems. The validity of the work for hire doctrine is supported by *Beloff v. Pressdram (1973) 1 All E.R. 241 (Ch. D)*. It is a solution to honouring the ideas of originality, creativity, and labour that exist in traditional copyright

---

<sup>2</sup> This approach is addressed within *Nova Productions v. Mazooma Games [2007] EWCA Civ 219*.

<sup>3</sup> *Eastern Book Company v DB Modak (2008) 1 SCC 1*.

<sup>4</sup> *Eastern Book Company v DB Modak (2008) 1 SCC 1*.

<sup>5</sup> *Rupendra Kashyap v Jiwan Publishing House Pvt Ltd (1994) 28 DRJ 286*.

<sup>6</sup> <https://edwardslaw.ca/blog/work-made-for-hire-explained/#:~:text=Unlike%20the%20U.S.%2C%20in%20Canada,employment%20or%20independent%20contractor%20relationship.>

law, while also allowing the legislation surrounding copyright to evolve with the creation of new technology to promote societal evolution and creation.

Although this doctrine has benefits that support the creation of new technologies and provides an answer to the ownership of copyright if it is granted, there are issues with this work-made-for-hire approach to ownership. For one, it could over-reward users, programmers, and companies. It could also allow companies to own every piece of work that the AI program could produce, which has the potential to lead to access inequality issues, where copyright is being obtained at unprecedented rates and access to autonomous AI systems becomes impossible. For these reasons, there would still need to be limits placed upon what can be copyrightable works under this new definition of ownership. A solution would be to still separate AI works into the two above-noted categories, where copyright protection and ownership was only granted over the AI works that were created by humans with the assistance of these AI systems and not solely AI-generated works. This approach has been contested because of the potential issues flagged above, but it is evident that the other option of immediate entrance into the public domain could create issues down the line with liability and infringement. If there is no classified owner over AI systems, then there will be no one to hold liable for infringement over other copyrighted works.

Another approach that is taken for authorship is to grant it to the programmer, in instances where the developer of the AI system is in direct control of their development. This is supported by similar regimes in Hong Kong, India, Ireland, New Zealand, and the United Kingdom. It would need to be determined through law the specific control that a programmer has over their AI system to be able to hold authorship of works created from that system. This would take the place of the human contribution, only if there is a specific vision of the works that the AI will produce. This approach is addressed and supported through the example of "The Portrait of Edmond Belamy", which was a work created using an AI system that used a data set of 15,000 portraits and sold for massive profits at auction.<sup>7</sup> This is an example of what experts are arguing against allowing copyright protection over, as its creation is evidence of an AI system that has been "designed to access so many variations of input data that developers cannot even imagine what kind of outputs it will lead to" (Budden, 2022). The regimes that support authorship being granted to the programmer support it on the grounds that the result from the AI technology is not too remote from the work of the programmers.<sup>8</sup> It must be ensured that the pattern used by the AI system is discernible to the programmer and not just the AI system; as the basics of copyright law or creative input and labour are lost the moment that the programmer cannot foresee every possible output of the system (Chaudhary, 2022).<sup>9</sup> Granting authorship to the programmer can still fit into the work for hire doctrine if the programmer created the AI system during their employment, giving any copyright ownership to the employer of the programmer.

---

<sup>7</sup> Bradley Budden, "On the Intersection of Artificial Intelligence and Copyright Law" [2022] 47:1 Can L Libr Rev 10.

<sup>8</sup> Gyandeep Chaudhary, "Artificial Intelligence: Copyright and Authorship/Ownership Dilemma?" (2022) 13:2 Indian JL & Just 212.

<sup>9</sup> Gyandeep Chaudhary, "Artificial Intelligence: Copyright and Authorship/Ownership Dilemma?" (2022) 13:2 Indian JL & Just 212.

### *Infringement and Liability*

Currently, there is an issue with determining if an AI-generated work has infringed existing copyright, when there are autonomous AI systems that create works beyond the ideas thought of by the programmer. When AI can create and act without input or supervision of a human source, it becomes very difficult to catch infringement unless the programmer or company that is using the AI system is actively keeping record of the sources used in creating its works. One way infringement prevention has been considered internationally is through requiring the implementation of a record-keeping system for anyone involved in the development or deployment of AI systems. An approach to this kind of record keeping is outlined in article 12 of the European Union's AI Act. It states that "high risk AI systems shall be designed and developed with capabilities enabling the automatic recording of events ('logs') while the high-risk AI systems are operating. Those logging capabilities shall conform to recognized standards or common specifications" (EU AI Act, 2023). The Act also states that logging should also include that it the ability to record the period of use, note the reference database which input data has been checked by the system, the input data for which the search has led to a match, and any identification of natural persons involved in the verification of the results. A similar model to this one outlined in article 12 of the EU's AI Act would be helpful in Canada for limiting infringement and delineating clearer liability standards. Further, to help determine liability when copyright has been infringed, AI works could be divided into two categories. The first would be work created by a human and aided by AI. In this scenario, the liability would be with the creator of the works. In this type of creation, the use of the AI system by the individual is the acceptance of being liable for the content that is used to create the work. The second would be the work generated solely by AI, which is the main focus of debate with infringement and liability. This is because it is harder to track the works that are being infringed. In this case, the work-made-for-hire doctrine could be the solution for liability: it would be clear who holds liability for the AI system's work, as they would be the ones who would be directly benefiting from it. The employer would be able to implement a record-keeping system to combat infringement and liability, by ensuring that all works used are lawfully accessed.

Liability could also be limited by writing exceptions into the law, like exceptions proposed with respect to TDM. Stating specific exceptions to infringement can allow the government to balance the interests of the creators of copyrighted works and those using AI systems to create works to ensure that there is no unreasonable prejudice against the interests of the rights holder, while outlining circumstances where users of copyrighted works would not need prior authorization or payment of royalties. This approach would be based on the de minimis doctrine where the impact on the rights holder is determined to be too minimal to qualify as an infringement occurring, along with the socially beneficial nature of the work.

### *Conclusion*

Change is needed to improve copyright protections in the digital age. In this consultation, we have recommended a number of possible routes for Canada to take as it reimagines copyright in the wake of generative AI. These changes will go a long way toward protecting creatives, artists, and copyright holders against AI-generated violations.