Consumer Council

Submission on the enhancement of the Copyright Ordinance regarding protection for artificial intelligence technology development

4 September 2024

1.   The Consumer Council (the Council) is pleased to submit its views from a consumer protection perspective with respect to a public consultation on the enhancement of the Copyright Ordinance (CO) issued by the Commerce and Economic Development Bureau (CEDB) and the Intellectual Property Department (IPD), regarding the protection for artificial intelligence (AI) technology development.

The Council's Views

2.   Given the rapid advancement of AI technology, the Council agrees with the necessity of reviewing the CO to ensure Hong Kong's current copyright regime keeps up with the times. Without proper regulations, individuals could easily be embroiled and exposed to potential risks when it comes to AI. For example, copyright owners may be discontented with their works being used in unauthorised AI model training, and unwitting consumers may utilise AI-generated works that involve usage of copyrighted materials without the consent of the original creators/copyright owners.

3.   In the following parts, the Council puts forward its responses to questions in the consultation document from different aspects that concern consumer interests.

Copyright protection of AI-generated works

4.   According to paragraph 2.4 of the consultation document, under the CO, the copyright subsists in four types of original works (i.e. literary, dramatic, musical and artistic works (LDMA works)) and sound recordings, films, broadcasts, cable programmes and the typographical arrangement of published editions (non-LDMA works). The Council seeks to review the current regulatory scope to allow more flexibility to cater for all types of creations in today's society.

5.   The rapid technological advancement nowadays has enormously fostered creativity and thus diversified the spectrum of creations. To take non-fungible tokens (NFTs) and works in the virtual space (such as the metaverse) as examples, NFTs and virtual works can be produced by humans or computers,

yet it is unclear whether they are covered in the existing CO, and even if so, whether they are categorised as LDMA works or non-LDMA works.

6.   Referencing the copyright law in the Europe Union (EU), other kinds of works such as computer programmes and databases are also within the regulatory framework. However, there might still exist doubt to consumers in Hong Kong. Therefore, the Council deems that it is of vital importance to regularly review the coverage of the current CO to keep up with the times as well as the standards and latest developments of their counterparts.

7.   In addition, as stated in paragraphs 2.4 to 2.10 of the consultation document, the CO currently identifies two types of LDMA works, i.e. those created by a human author (ordinary LDMA works), and those that are computer-generated in circumstances such that there is no human author (CG LDMA works).

8.   However, the emergence of generative AI has not only rapidly fostered the production of CG LDMA works but also enabled recreations on top of these works (i.e. secondary creations or derivative works), causing controversy over the right to use and the right to own. For instance, there could be cases where the individual creates a comic by adding dialogues to a series of images produced by AI and claims the ownership of the whole work. For this kind of recreation, it might be difficult to define whether it should be an ordinary LDMA work or CG LDMA work, and which party should have the ownership and the right to use. The Council expresses concern over the exceptions to the dichotomy and considers that the above scenario should be taken into account when making any enhancement or clarification of the CO.

9.   Moreover, under the current CO, the duration of copyright is often set at 50 years from which the work was made or the author's life plus 50 years after death. Despite rare discussion on this around the globe at the moment, the Council proposes that the CEDB and the IPD might take a forward-looking approach to examine whether there is a need to shorten the duration of copyright, especially for computer-generated works.

10.   The Council believes that a tiered copyright system might be considered to better honour the effort of human-created or computer-generated works. Nevertheless, the Council emphasises that shortening the copyright duration does not mean sacrificing the interests of copyright owners; it is essential to maintain effective protection of copyright owners in the meantime.

Copyright infringement liability for AI-generated works

11.   The Council shares the same view in paragraph 3.16 of the consultation document that laying down rigid rules that assign infringement liability to specific person(s) across the board would fail to account for the unique factual context of each infringement, thereby compromising fairness across a diverse range of situations. Still, the Council is concerned that the consumer-user might be unwittingly held liable for copyright infringement because the AI system had already infringed the copyright at the model training stage even when his/her prompts are appropriate. The Council emphasises that the liability of copyright infringement should not be unfairly transferred to the consumers, especially when the consumers are oblivious to the algorithm or the training of the AI system.

12.   Besides, the Council proposes that when formulating the way forward, more scenarios or cases should be given to demonstrate different parties' liability under different circumstances for the lay public's reference. It is believed that concrete examples would foster consumers' understanding and facilitate them in making cautious decisions before using AI-generated works.

13.   Understanding that there exists difficulty in enforcement given the fast-evolving nature of generative AI and the complexity of data sources for AI model training, the Council suggests that the Government might make reference to the judgement of relevant court cases around the globe, and develop practical enforcement guidelines for the public's awareness.

Possible introduction of specific copyright exception

<u>Opt-out option for copyright owners</u>

14.   The Council notes the proposal of introducing an opt-out option for copyright owners to expressly reserve their rights in the text and data mining (Proposed (TDM) Exception). As the training of AI models could not be undone, which means it may not be possible to erase the data once the data has been used for training of AI model, it is therefore essential to obtain consent from the copyright owners or respect their demand for opt-out before everything goes irreversible.

15.   Such practice could keep Hong Kong on par with the development of other advanced economies. To quote from the consultation document, in the EU, an opt-out option is available to copyright owners; in execution, copyright owners may indicate their choice to reserve their rights through machine readable means in case the content made publicly available online.

Alternatively, copyright owners may opt out by sending notifications to the companies before the training of AI model.

16.   However, the execution might be challenging in reality as there could be errors in the reading process of machines, and copyright owners might not know what company (or party) would extract their data for AI training. Thus, these limitations should be considered while exploring the Proposed TDM Exception in order to enhance the feasibility. The Council suggests that emphasis on the importance of human oversight and setting of guidelines or requirements should be taken to ensure that AI companies should frequently check the accuracy of machines in reading data, particularly in identifying copyright owners' opt-out preference. This is very crucial in strengthening copyright protection while fostering AI development.

Other comments

<u>Cross-border issues</u>

17.   Cross-border copyright issues could be complex. The origin of data, the storage of data, the training, and the development of AI models could be across various jurisdictions in the world. For example, in January 2023[1], Getty Images (a visual media company) alleged that Stability AI (an AI company) had "scraped" 12 million images from various websites operated by Getty Images without its consent, and unlawfully utilised those images to train and develop Stable Diffusion (a text-to-image model) under the Copyright Designs and Patents Act 1988 in the United Kingdom. Nevertheless, Stability AI claimed that the training and development of Stable Diffusion took place in the United States[2].

18.   The case is still ongoing, but it would indeed impose a profound impact on future AI development. If the AI model creator is convicted of copyright infringement, end-users of the model may subsequently receive the same allegation made by copyright owners, and the cross-border nature of cases like this might complicate the interpretation of consumer-users' liability.

19.   Meanwhile, different jurisdictions have different regulatory approach towards copyright issues, which also increased the difficulty in enforcement. For instance, the EU AI Act (the Act)[3], which entered into force in August 2024, stated that companies not complying with the rules will be fined. Fines could go up to 7% of the global annual turnover for violations of banned AI

---

[1] https://newsroom.gettyimages.com/en/getty-images/getty-images-statement
[2] https://www.penningtonslaw.com/news-publications/latest-news/2024/generative-ai-in-the-courts-getty-images-v-stability-ai
[3] https://eur-lex.europa.eu/eli/reg/2024/1689/oj

applications. Yet, how this could be related to the copyright issues across the world is still unclear. The Council considers close collaboration with international counterparts is indispensable when considering the enhancement of the CO.

Transparency of AI models and AI content

20.   Transparency regarding the sources of training data is vital in developing AI systems, such that copyright protected materials used in training an AI model can be identified. The Council believes including requirements pertaining to transparency in relevant regulations could protect consumers' right to know. For example, the Council notices that the EU AI Act requires a general-purpose AI model to make publicly available a sufficiently detailed summary of the content (including text and data protected by copyright) utilised for training the model. The summary should include information such as main data collections or sets that went into training the model, with a narrative explanation about other data sources used.

21.   The EU AI Act further states that the European Artificial Intelligence Office should provide a template for the summary, which ought to be simple, effective, and allow the provider to provide the required summary in narrative form. Although the template is not yet available, the Council advises that keeping a close eye on its release or establishing one based on similar principles are worth attention.

22.   Moreover, the EU AI Act also requires AI developers who use an AI system to generate or manipulate content that appreciably resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic to clearly and distinguishably disclose that the content has been artificially created or manipulated by labelling the AI output accordingly and disclosing its artificial origin.

23.   The above requirements promote good market practice which could facilitate consumer-users in identifying AI-generated works and thus examining the risk of committing copyright infringement. The Council suggests that similar guidelines or requirements be developed to increase transparency for consumers' information. Besides, mandatory application of watermarks to AI-generated content is worth considering. For example, consumers should have the right to know whether an online store used pictures generated by AI as illustration when purchasing.

24.   In addition to amendments of the CO, the Council deems that the provision of a reliable and robust tool that assists consumers in identifying whether a work is AI-generated would also contribute to greater transparency.

In consideration of the emergence of generative AI in recent years, different market players have developed various AI content detector. For example, Grammarly lately released an AI detection tool to help users identify the origin of each part of a document, and distinguish which sections were created by humans and which were produced by AI; McAfee also launched a "Deepfake Detector" to flag AI-generated audio.

25.   The Council also notices that although some generative AI models provided "content sources" in the generated content, the accuracy, reliability and relevancy of such sources remain an issue.

26.   The Council recommends the Government to commit a reliable, independent third party to design a similar tool to alert consumers of AI content and whether there is potential involvement of copyright issues. Three key elements should be considered in the design of the tool, namely perpetuality (being able to stand the test of time), immutability (being unchangeable) and authenticity (involving accurate training and thus producing reliable results). The Council believes that the tool would aid consumers efficiently in avoiding copyright infringement when using AI-generated works.

27.   Establishing an economical, accessible and friendly licensing scheme for different types of users to use copyrighted works might also be considered. Such licensing scheme is supposed to provide a reasonable yet easy access to copyrighted works and ensure that the usage is authorised.

Consumer education

28.   While the novel technology constantly evolves, the public's focus is predominantly on its capabilities and limitations, often neglecting the copyright infringement issues that could arise from TDM and/or using AI-generated works at present. Therefore, it is fundamental to put in educational efforts to ensure that the public and technology industry have adequate knowledge of the relevant regulations while developing or adopting AI.

29.   Furthermore, resources should be allocated to public education to instil in consumers a mindset of harnessing the power of AI with caution, such as examining whether the sources used for TDM are copyrighted and if the AI-generated works are prone to copyright infringement. Another aspect worth covering in education is the contractual arrangements between AI system owners and consumers. It is of vital importance to educate consumers what items they should pay attention to before signing contracts and how to clarify liability, so as to mitigate the risk of copyright infringement of AI-generated works.

30.   The Council is of the view that adequate consumer education could empower consumers in self-protection and encourage acts to unleash the potential of AI without infringing copyright, thus expediting the industry development.

Conclusion

31.   The issue of copyright and AI is still a topic of ongoing debate among experts, policymakers, and stakeholders globally. The Council fully supports the CEDB and the IPD to engage all stakeholders closely with an aim to evolve the CO that can achieve a dual win in the economy and AI development for Hong Kong. It would be vital to monitor global momentum of regulatory developments for insights into catalysing technological progress while the interests of creators and consumer-users would be well protected.