

September 14, 2024



Commerce and Economic Development Bureau (Division 3)
23/F, West Wing
Central Government Offices
2 Tim Mei Avenue
Tamar, Hong Kong

**The American Chamber
of Commerce in Hong Kong**
22/F, Hong Kong Diamond Exchange Building
8-10 Duddell Street, Central
Hong Kong

Dear Sir / Madam,

Response to the Public Consultation on Enhancing the Copyright Ordinance regarding Protection for Artificial Intelligence Technology Development

I am writing on behalf of the American Chamber of Commerce in Hong Kong (“AmCham”) to submit our response to the recent public consultation regarding the enhancement of the Copyright Ordinance.

AmCham comprises a diverse membership, and in the past month, we have gathered a range of perspectives from relevant members on this important issue. We recognize the significance of considering whether or not updates to the Copyright Ordinance are required in light of the rapid advancements in artificial intelligence and the challenges posed by generative AI. However, it is essential to note that our members hold very different views on various aspects of the proposed enhancements. We try to consolidate these opposing views in the enclosed two documents for your reference and consideration.

Given these at times divergent viewpoints, we would like to make clear that the Chamber does **not** take an official position on this matter. Instead, the letter serves as an overview of the diverse opinions and concerns expressed by relevant AmCham members regarding the enhancement of the Copyright Ordinance.

We appreciate the opportunity to contribute to this consultation and encourage the HKSAR Government to consider the multifaceted nature of these issues as you move forward. Our members are committed to engaging in constructive dialogue and remain open to further discussions on this difficult but essential topic.

Thank you for your attention to our response. We look forward to seeing how the consultation outcomes will shape and balance copyright protection whilst fostering innovation in Hong Kong.

Sincerely,

Eden Woon
AmCham President

Encl.

1. Consolidated views of members from the innovation & technology community
2. Consolidated views of members from the creative & media industry

Consolidated views of members from the innovation & technology community

1. Support a broad and flexible TDM exception

- (a) Members support an innovation friendly approach to regulating AI, which includes the introduction of a broad and flexible text and data mining (“TDM”) exception in the Copyright Ordinance (“CO”), which could help in driving the development of artificial intelligence (“AI”) in and transforming industries across Hong Kong. To reach their full potential, AI models, especially foundational large language models, require access to immense, diverse datasets to be effectively trained. Without a broad and flexible TDM exception, AI developers would face significant barriers and transaction costs in obtaining licenses from multiple copyright holders, hampering their ability to build cutting-edge AI systems. Moreover, overly restrictive copyright regimes that empower rightsholders to arbitrarily block or monetize training data access will distort the competitive landscape, entrenching the position of large incumbents and stifling the emergence of innovative AI startups. Therefore, a broad and flexible TDM exception that allows the use of copyrighted works for both commercial and non-commercial TDM activities would enhance the legal certainty and flexibility needed to utilize a wide range of data sources, which is crucial for securing Hong Kong’s position as a leading international innovation and technology hub.
- (b) A broad and flexible TDM exception is also consistent with the goal and rationale of copyright protection. Copyright law is fundamentally focused on protecting the unique creative expression of authors, not restricting access to the underlying facts, data or ideas contained within copyrighted works. The purpose of copyright is to incentivize the production and dissemination of original expressive content, not to grant rightsholders monopolistic control over raw information or concepts. A broad and flexible TDM exception upholds this principle by allowing the use of copyrighted materials for TDM activities, which does not involve enjoyment of the copyrighted work, without infringing on the exclusive rights of the copyright holder. By facilitating broad access to diverse data sources, which may include copyrighted works, the TDM exception enables AI developers to uncover valuable insights, trends and patterns – the type of non-expressive information that copyright law was never intended to monopolize. This allows for the development of innovative AI applications that serve the public interest, while still preserving economic incentives for authors to create original, expressive works. Maintaining this balance is crucial for promoting technological progress and ensuring a vibrant, balanced copyright system that fulfills its purpose.
- (c) To maintain its position as a leading innovation and technology hub within Asia, Hong Kong should consider adopting a broad and flexible TDM exception similar to the approaches taken in peer jurisdictions like Singapore and Japan. As summarized in the Consultation Paper, both Singapore and Japan have adopted TDM exceptions which permit both commercial and non-commercial computational data analysis and processing of copyrighted work, with very limited conditions. In particular, Singapore’s TDM exception cannot be overridden by contract terms, ensuring AI developers and researchers have legal certainty to access and utilize diverse data sources needed to train advanced models efficiently. If Hong Kong instead adopts a restrictive approach, AI model creators may move AI training workloads to overseas where they face fewer hurdles. This will create not only a missed opportunity for Hong Kong to promote AI development within the jurisdiction, but also give rise to a risk that models will not adequately take into account Hong Kong context and culture (due to insufficient open training materials available in Hong Kong). Therefore, aligning Hong Kong's approach with regional best practices would cement its attractiveness as a hub for AI innovation and sustain its competitiveness in the fast-evolving global technology landscape.

- (d) Section 38 of the CO already provides a fair dealing exception for research activities, which AI developers could potentially rely on to conduct TDM activities using copyrighted works. TDM is fundamentally a research tool, employing computational analysis to identify valuable insights and patterns within large data sources. Given that Section 38's fair dealing exception could be interpreted to cover commercial research activities, AI developers in Hong Kong may already have been able to leverage this provision to carry out commercial TDM activities on copyrighted works. As the existing fair dealing exception in Hong Kong has a broad scope with limited conditions, it is recommended that any new TDM-specific exception follow a similar approach and should not be more restrictive than the existing fair dealing exception. This would help ensure regulatory consistency, where the same risk and activity is dealt with in a consistent manner. It would also avoid creating legal uncertainty that could disrupt ongoing commercial TDM activities that may have been relying on the fair dealing exception. Aligning the TDM exception with the existing fair dealing framework would provide AI developers the necessary legal clarity and flexibility to fully leverage copyrighted data sources for research and innovation.

2. Limit the conditions to the TDM exception

- (a) The Hong Kong Government should limit the conditions for triggering the TDM exception. In particular, the TDM exception should not allow copyright owners to reserve their rights by exercising opt out or implementing licensing schemes. As submitted above, a broad and flexible TDM exception not only promotes the development of AI in Hong Kong, but is also consistent with the goal and rationale of copyright protection. Allowing copyright owners to make TDM activities unauthorized by opting out or implementing a licensing scheme would undermine the very purpose of a TDM exception by empowering copyright owners to arbitrarily block access to valuable data sources. Rather than promoting innovation and creativity, licensing schemes of copyrighted works for TDM activities would only give copyright owners monopolistic control over their work and generate windfall profits to copyright owners (with licensing price arbitrarily set by copyright owners), which imposes significant transaction costs and administrative burdens on AI developers, particularly smaller startups and researchers, which lead to a distortion of the competitive landscape and stifling the emergence of innovative AI startups.¹
- (b) Allowing copyright owners to reserve their rights by exercising opt out or implementing licensing schemes would also fragmentize the comprehensive datasets needed to develop high-performing AI models, as not all AI developers would have equitable access to all available data sources for training their AI models. AI developers would be forced to patch together incomplete, biased datasets. This fragmentation undermines the ability to uncover the rich insights and patterns hidden within diverse, large-scale data sources - the very purpose of TDM. This can lead to suboptimal model performance, propagation of biases, and missed opportunities for breakthrough discoveries and innovations.
- (c) Footnote 70 on page 31 of the Consultation Paper says that “the computational data analysis and processing activities conducted for training AI models in the model market may be of a commercial nature, and may copy and *store* the whole of a copyright work” (emphasis added). Members would like to clarify that AI models do not actually store the training data used to create them. This is elaborated in a discussion paper issued by The Hamburg Commissioner for Data protection and

¹ Martens, B. (2024) “Economic arguments in favor of reducing copyright protection for generative AI inputs and outputs,” Working Paper 09/2024, Bruegel

Freedom of Information of Germany recently². During the training process, highly abstracted mathematical representations called embeddings are calculated from the original text data is transformed. These embeddings capture patterns and relationships between linguistic elements, but do not retain the specific content from the original training data. The trained model only stores these statistical representations, not the raw training data itself. It reinforces the argument that the use of copyrighted works for TDM activities do not involve the enjoyment of the author's expression in the copyrighted works, and so no condition should be imposed on the TDM exception.

- (d) This group of members therefore recommend the Hong Kong Government to adopt the same flexible approach to TDM exception as Singapore, which adopts a broad and flexible TDM exception without allowing copyright owners to override by contractual terms or opt-out conditions that would allow copyright owners to selectively restrict access to their works. This could help in establishing a favorable legal regime for the development of Hong Kong into a major AI and technology hub within the region.
- (e) If any conditions need to be included within a TDM exception, they should be narrowly crafted to avoid undermining the core purpose of the TDM exception. As an example, Singapore's TDM exception only applies if users have "lawful access" to the copyrighted works. However, even this seemingly straightforward condition has faced reconsideration, with the Singapore government recently launching a public consultation on whether prohibitions on circumventing technological access controls would impair or adversely affect the dealing with copyrighted works that would be non-infringing based on permitted uses.³ Any additional conditions, such as an opt-out right for copyright holders, must be designed with similar caution. If an opt-out mechanism is included, the intention to opt out should be readily identifiable in a clear, standardized machine-readable manner. This would minimize the administrative burden and legal uncertainty AI developers face in ascertaining the copyright status of each work they seek to analyze. It may also be worth examining limitations on how the works are used beyond TDM, such as limiting further distribution of the works. However, any such conditions should be mindful that TDM researchers commonly share their data mining corpus to allow for reproducibility of results. Overall, the guiding principle should be to craft TDM exception conditions that unambiguously enable access to comprehensive datasets, rather than imposing hurdles that fragment the available training data and distort the competitive landscape.
- (f) Regardless of the Hong Kong government's approach to the TDM exception, AI developers could still choose to voluntarily exclude content from data sets. For instance, it is common practice for content owners to include text files in their web sites identifying content that they do not wish to be collected by automated web scraping tools. Many AI developers voluntarily comply with such notifications when compiling data sets and can continue to do so, regardless of whether a TDM exception exists. Commercial decisions, such as opt-out and licensing schemes to remain voluntary is crucial for maintaining a balanced ecosystem between content creators and data miners.

² Hamburg Commissioner for Data Protection and Freedom of Information. (2023). Discussion paper: Large language models and personal data.

³ Ministry of Law and the Intellectual Property Office of Singapore. (2024). 2024 Public Consultation on Prescribed Exceptions In Part 6, Division 1 of The Copyright Regulations 2021 (https://www.mlaw.gov.sg/files/2024_Public_Consultation_on_Prescribed_Exceptions_in_Part_6_Division_1_of_the_Copyright_Regulations_2021.pdf)

3. Allocate responsibilities on copyright infringement by AI-generated works

- (a) Some members agree with the general proposition in the Consultation Paper that liability for copyright infringement by AI-generated works should be determined on a case-by-case basis. Who should be liable for such infringement is indeed fact-dependent. Nevertheless, when determining the responsibility of a copyright infringement by AI-generated works, the Hong Kong Government should take into account the fact that an AI system typically generates output independently based on user input without the direct involvement or knowledge of the AI developer. It is possible that some users could manipulate an AI service to generate content that is similar to copyright protected content. If a model produces such content only through prompts that violate the terms of service of the model, then it should not be attributed to the model's developers. Under the current CO, however, there remains a possibility that the AI developer could be held liable/jointly liable for copyright infringement or for authorizing an infringing act on the basis that the AI developer has (through the AI system) arguably collaborated with the AI user, or has control over the AI system whose output reproduced in whole or substantial part of a copyrighted work. To give legal certainty to AI developers for using copyrighted work to train their AI models, the Hong Kong Government is recommended to consider clarifying in the CO that only the party with the intention to infringe copyright should be held liable. In addition, the Hong Kong Government may consider including a safe harbor provision in the CO to limit AI developers' liability in circumstances where users prompt or cause the AI systems to reproduce existing works in whole or substantial part, provided AI developers take reasonable steps to limit or stop copyright infringement when notified (for example, notice and take down of infringing content).

4. Reconsider copyrightability of purely AI-generated works

- (a) Some members urge the Hong Kong Government to carefully reconsider whether purely AI-generated works should be eligible for copyright protection. While the Consultation Paper proposes that the existing "computer-generated works" provisions in the CO could potentially cover AI-generated literary, dramatic, musical, and artistic ("LDMA") works, there are significant uncertainties around how the current copyright regime could apply in such cases. According to the CO, all LDMA works must satisfy the originality requirement for copyright to subsist, but it is highly unclear how a work generated entirely by an AI system, without any direct human authorship, could meet this criterion. Additionally, under the CO, the "author" of a computer-generated work is defined as the person "by whom the arrangements necessary for the creation of the work are undertaken". However, with purely AI-generated works, there may be multiple parties involved – from the providers of training data, to the AI developer, deployer, and user – making it ambiguous who the true "author" is. As generative AI capabilities advance exponentially, leading to a proliferation of purely AI-generated LDMA works, it is imperative that the Hong Kong Government provides clear legal certainty for all stakeholders. And it is recommended that purely AI-generated works should not be granted copyright protection for the reasons below.
- (b) From a policy perspective, purely AI-generated work should not be granted copyright protection as the core rationale behind copyright protection is to incentivize human creativity and expression, not to grant monopolistic control over non-expressive content generated by machines. Extending copyright to purely AI-generated works would run counter to the fundamental purpose of the copyright system. Generative AI models do not require the promise of exclusive rights to spur their development – their outputs are produced through automated computational processes, without any meaningful creative contribution from human authors. From an economic perspective, there is also no need to protect purely AI-generated work as the marginal cost of producing such outputs is very low, close to the marginal cost of reproduction, which eliminates the rationale for copyright as an

incentive mechanism. The approach taken in Japan's copyright framework provides a compelling model. In Japan, materials autonomously generated by AI (i.e., material that is generated by AI without any instructions from humans or only by giving simple instructions as prompt) are not considered as creatively produced expressions of thoughts or sentiments and are therefore not considered copyrighted works. Therefore, the recommended approach is to maintain copyright's focus on original human expression, while allowing the unimpeded use of AI to generate non-expressive content that serves the broader public good.

- (c) The Consultation Paper distinguishes between AI-generated works created and generated by generative AI without a human author based on users' prompts and works created by human authors who utilize AI systems as a tool to aid their creative processes ("**AI-assisted works**"). AI-assisted LDMA works are considered as ordinary LDMA works (rather than computer-generated LDMA works) under the CO. However, the current CO lacks clear guidelines on how to reliably differentiate these two categories of works. This ambiguity risks introducing significant legal uncertainty around which copyright regime should apply to a given work. The Hong Kong Government is therefore recommended to look to the Japanese framework, which sets out factors for determining whether AI has been used as a genuine creative tool, versus cases of purely autonomous machine generation. Specifically, the Japanese approach evaluates whether the human user had a demonstrable "creative intention" and made a substantive "creative contribution" beyond simply providing basic prompts or instructions.

Consolidated views of members from the creative & media industry

1. There is no necessity for the implementation of any text & data mining type exception

- (a) Some members disagree that any TDM exception is necessary. In this regard, it is noted that AI developers are increasingly turning to copyright licensing over unauthorized use. Such licensing not only creates certainty over the right to use copyrighted content (as opposed to the uncertainty of relying on an exception), and also provides the possibility of gaining access to content not freely available on the public internet. Thus the implementation of a TDM exception would undermine such evolving licensing practices.
- (b) In this regard, the UK had considered the need for a similar TDM exception. In 2022, the UKIPO had proposed a broad TDM exception¹ “for any purpose” with the view of making the UK “a global center for AI innovation”. It was noted that even in these circumstances, the proposed TDM exception was subject to the requirement that rightsholders would “still have safeguards to protect their content, including a requirement for lawful access”, and was limited to making a copy of the work for the purpose of carrying out computational analysis of the data recorded in the work. However, in a report published on January 17, 2023,² the House of Lords Communications and Digital Committee (the “Committee”) noted that “developing AI is important, but it should not be pursued at all costs”. The Committee went on to recommend that the UKIPO “pause its proposed changes to the text and data mining regime immediately”. The Committee also highlighted that before any such TDM regime could be implemented, it would be necessary to “conduct and publish an impact assessment on the implications for the creative industries”. This is in recognition to feedback received that “IP rights provide crucial revenue streams for those in the creative sector, from freelancers to major businesses” and that “without creators’ rights to copyright protection over the works they create there is little incentive to invest in their own future careers”.
- (c) Whilst members understand that some proponents of TDM exceptions have argued that such exceptions are needed to prevent rightsholders from maintaining control over raw information or concepts, it is felt that such arguments are inherently contradictory. If all that is being taken by AI developers is underlying facts, data or ideas contained within copyrighted works, then indeed no exception would be required as no copyright would have been infringed. As explained in a recent paper by Jacqueline C. Charlesworth³ “[c]ontrary to the suggestion that the works on which AI systems are trained are set aside after the training process, in fact they have been algorithmically incorporated into and continue to be exploited by the model ... The copying of expressive content of a work for the purpose of generating new content from the copied content capitalizes on the original expressive purpose of the work”.
- (d) The mere fact that AI developers continue to push for TDM exceptions is a recognition that the process of scraping content during the training of AI models would constitute copyright infringement in the absence of such an exception. Consequently, before any such exception is granted, as highlighted by the UK House of Lords Communications and Digital Committee, great care must be taken to ensure that the creative industries are not unduly impacted.

¹ <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation>

² <https://publications.parliament.uk/pa/ld5803/ldselect/ldcomm/125/12502.htm>

³ Generative AI’s Illusory Case for Fair Use - Jacqueline C. Charlesworth at page 4

- (e) Some members, on the other hand, are of the view that copyright law protects rightsholders by granting them the exclusive right to authorize specific uses of their works by third parties. Copyright law also includes various exceptions to that general rule, which may exist in differing forms in jurisdictions around the world. Licensing markets for training AI models have been developing. There is no need for statutory licenses or mandatory or extended collective licensing. When rightsholders license their works for training AI systems, those licenses should take the form of voluntary and direct arrangements.

2. Limitations to any TDM exceptions that may be considered

Whilst some members do not see the need for a new TDM exception in Hong Kong given that licensing markets have been developing, in the event that the Hong Kong Government were minded to consider a new exception, it should be limited by appropriate safeguards to ensure that the rights of the creative industry are appropriately protected. Such safeguards should be at the minimum include the following:

(a) Lawful content/access

The TDM exception should not authorize training AI on infringing works or pirate sites, or on works to which access has not been authorized or where it would be a breach of contract or terms of service, or where doing so would involve circumventing technological protection measures.

(b) Opt-outs

Any TDM exception must provide rightsholders the ability to easily exclude their works from training in an effective and non-burdensome manner. AI systems offered to the general public or external commercial customers must be required to comply with the choice of the rightsholder. The Hong Kong Government should encourage AI developers that use copyright works under the TDM exception to develop technical measures or standards, with input from rightsholders, to enable rightsholders to exclude their works including on a “per work” basis.

(c) Record keeping

Where AI developers rely on a TDM exception for training AI systems and services that are offered to the general public or external commercial customers, they should maintain reasonable records of training data, including copyright protected works, and make those records available for review, to provide transparency to rightsholders and users.

3. Allocate responsibilities on copyright infringement by AI-generated works

- (a) Members agree with the general proposition in the Consultation Paper that liability for copyright infringement by AI-generated works should be determined on a case-by-case basis. Who should be liable for such infringement is indeed fact-dependent.

4. Copyrightability of purely AI-generated works⁴

- (a) Generative AI may be used as an assistive tool to aid in human creativity but cannot, itself, be the creator of copyrighted works. A work that is entirely generated by AI without sufficient human

⁴ We refer to the following site that summarizes a number of the existing litigations relating to AI Platforms <https://www.techtarget.com/whatis/feature/AI-lawsuits-explained-Whos-getting-sued>

creativity or contribution should not be protected by copyright because such a work fails to meet the human authorship and originality criteria.

- (b) The underlying principles of copyrightability – including a requirement for sufficient human authorship – should be technology neutral and apply to uses of generative AI in the same way that such principles apply to uses of other technologies that assist creators in realizing their vision. The use of generative AI as an assistive tool to aid human creativity should not foreclose copyright protection for those works.