Geneva, 8 September 2024

Ref: Public Consultation by the Hong Kong Commerce and Economic Development Bureau Intellectual Property Department on "Copyright and Artificial Intelligence"

Dear Madams and Sirs,

This submission is made on behalf of the publishing industry by the Association of American Publishers, the Association of Canadian Publishers, Börsenverein des Deutschen Buchhandels (the German Publishers and Booksellers Association), the Federation of European Publishers, the International Publishers Association, the Singapore Book Publishers Association, the Sindicato Nacional dos Editores de Livros (the Brazilian Book Publishers Association), STM - the International Association of Scientific, Technical and Medical Publishers and the UK Publishers Association. Our membership includes numerous publishing companies native to or operating in Hong Kong, fostering jobs, contributing to the creative economy and supporting Hong Kong's educational, scientific and cultural excellency.

The copyright framework is and always has been the foundation of the publishing industry. Adequate copyright protection entails a combination of enforceable exclusive rights and carefully calibrated exceptions and limitations, which must be objectively justified and narrowly defined upon the Berne Convention's three-step test. The integrity of copyright protection is a key condition for publishers' investments. Enforceable exclusive rights are fundamental to incentivize authors, publishers, and other copyright owners to create, invest in, and make available to the public original and valuable works of human authorship. Through these works, publishers drive inspiration, entertainment, education, and significantly contribute to both local and global economies.

The time-tested principles of the international copyright treaties must continue to drive copyright related policy developments, including on AI. This framework sustains copyright owners' ability to license and enforce their exclusive rights, including in the context of AI, while establishing that AI developers must seek licenses to use copyrighted works. Along these lines, publishers advocate that policy reviews should abstain from introducing legislative changes that could encroach on present and future market developments based on licensing, including by setting out exceptions & limitations for TDM and other AI uses, and should instead foster transparency, accountability, and ethical development of AI systems in full respect of copyright and human creation.

We thank you for the opportunity to participate in the above captioned public consultation and respectfully offer our views on the issues raised in Chapter 4.

Chapter 4 Possible Introduction of Specific Copyright Exception

Q: What further justifications and information can be adduced to support (or roll back) the idea of introducing the Proposed TDM Exception into the CO with a view to incentivising the use and development of AI technology and pursuing overall benefits?

We respectfully submit that plans to introduce a new Proposed TDM Exception should be abandoned.

Businesses relying on text-and-data mining (TDM), including highly profitable Generative AI companies, depend on creative works of authorship to fuel and train their tools and products. Some of these companies claim that they do not have to seek permission or pay for the works they are so freely usurping for their advantage, but this is false.

It has become commonplace for some generative AI companies to highlight their innovation and investments as a reason for governments to give them broad copyright exceptions or regulatory loopholes when it comes to existing copyright requirements. But these companies owe their considerable success to the prior innovations and

investments of others—the intellectual and creative investments of authors and the financial investments of publishers.

Works of authorship—which include scores of fiction and nonfiction such as award-winning novels, renowned children's books, ground-breaking scholarship, biographies, scientific publications, political memoirs, and more— are not free for the taking. Rather, these works are economically incentivized and legally protected by both national copyright laws and international treaties and have been for centuries. This legal protection is neither debatable nor dispensable and there is no good public policy reason to weaken it for the convenience and objectives of technology companies, some of which are the most dominant corporations the world has ever known.

Legislators should abstain from introducing legislative changes that could encroach on present and future market developments. New exceptions & limitations for TDM and other AI uses deprive copyright owners from their rights to license their works. Instead, policy measures should continue to ensure that right holders can effectively and adequately license and enforce their rights.

The three-step test set out in the Berne Convention and in the TRIPS Agreement must continue to guide legislators in establishing exceptions & limitations, including in the context of AI. Systematic unauthorized use of copyrighted works to train and/or produce outputs of LLMs and other TDM applications, including generative AI, is at odds with the three-step test. It is a generalized free use that undermines the legitimate interests of authors and publishers, preventing them from exercising their intellectual property rights to exploit their works through licensing.

Although to determine incompatibility it suffices that the first step is not fulfilled, in this case a new TDM exception to enable unlicensed systematic uses of copyrighted works in such a broad remit is arguably incompatible with all three-steps.

First: Waiving copyright protection for TDM purposes cannot, by definition, be restricted to certain special cases as required by the three-step test. Given the overbroad application potential of the proposed new exception as outlined in the consultation paper, it would eliminate and prevent licensing by the legitimate copyright owners in a range of uses that is impossible to account for or predict, by any businesses that rely on or integrate TDM in their processes, including but not limited to AI and generative AI uses.

Second: The three-step test only allows uses to be exempted from copyright owners' permission *"provided that such reproduction does not conflict with a normal exploitation of the work"*. However, by enabling copyright circumvention for any TDM purposes, Hong Kong law would eliminate copyright owners' ability to license uses that will underpin most if not all future industries. This will have major implications. It's not only the immediate loss of revenue, but also the long-term consequences of publishers not being able to keep the continued investment in human creation that licensing revenues currently sustain. In addition, the unlicensed use of copyrighted works can function as an enabler of large-scale numbers of machine generated content, competing with original copyrighted works in an unfair manner. In practical terms, a permission to circumvent copyright would wipe out existing and emerging licensing markets, preventing copyright owners from exploiting their works in any context that involves TDM and further enabling the free and unlicensed production of competing content. In that sense, this exception not only conflicts, but eliminates the normal exploitation of copyrighted works through licensing. Moreover, there are no legal mechanisms that can effectively verify that uses under such an overbroad exception would be restricted to specific cases only, and that no subsequent uses would be made.

Third: The final step determines that an exception to copyright protection is only allowed if it *"does not unreasonably prejudice the legitimate interests of the author"*. As explained above, by wiping out any existing or emerging licensing markets in the TDM context, a new exception would unreasonably prejudice the legitimate interests of copyright owners, by excluding them from those markets for the profit of businesses built over their unlicensed content. Copyright owners will see their works used in an unauthorized manner and without compensation by users that will potentially re-use those works to create new competing content, including by using the original work without attribution and/or creating unauthorized derivative works. The unlicensed content produced will further flood and cannibalize the primary market for human-made copyrighted works. The effects of an exception would never balance interests, as claimed in the consultation paper. Instead, a new exception would further deepen imbalances rather than level the playing field between technology behemoths and copyright owners. This is especially impactful, given the large number of companies using datasets containing massive numbers of unlicensed copyrighted works.

Given the above, we respectfully submit that the new proposed exception is arguably incompatible with international law and with the integrity of copyright protection as set out in the WIPO Copyright Treaties and in the WTO TRIPS agreement.

Basic copyright principles also align with broader ethical considerations for society. New regulations should have a different focus, to facilitate transparency for the public, reduce misinformation, and most importantly, ensure that the world will continue to have authors, publishers, and a thriving creative community. Transparency as to which works have been used to train AI models is critical: it protects the public from harm and ensures that authors will not be stripped of attribution for their works.

Copyright is essential to reading, learning, and independent thought, and to the survival of the global publishing industry. Yet these issues are ultimately about the public interest itself. Because democracies depend on authors and publishers, we call the Government of Hong Kong governments to stand up for copyright and resist the calls to give technology companies more power over the public.

Q: How would the Proposed TDM Exception overcome the obstacles/limitations you have experienced in conducting TDM activities and facilitate the development of your business and industry?

The main obstacle/limitation faced by publishers in this context results from the systematic unlicensed use of their works and of the lack of legal action against the currently generalized copyright infringement. This cannot be overcome through a new exception. Instead, a new exception would only give an unfair competitive advantage to technology companies that did not invest in producing, editing, curating, peer-reviewing, promoting, and distributing the high-quality content that AI models have been trained on, and yet will be able to reap the benefits of others' creative work and investments.

The consultation paper mentions reasons to adopt a new TDM exception on an empirical basis, which can be better achieved through licensing. But a fair impact assessment is crucial to achieve a balanced copyright ecosystem. Exceptions & limitations must be subject to economic impact analysis regarding current and nascent market developments, to ensure they address a market failure, serve, and are suited to achieve a legitimate purpose, are objectively justified and proportionate. This must include assessing the impacts on creative industries which works are being exploited in an unlicensed manner to produce competing unverified content. Hong Kong's copyright law must continue to ensure that exceptions & limitations cannot be used to establish competing structures or systems (either commercial or not) or as a basis for digital players or any other businesses to operate as exploiters of copyrighted content which they have not invested in. If this is not safeguarded, exceptions will inevitably undermine the legal offer of books and literary works by human authors and publishers to human audiences.

The new proposed exception will create unfair competitive advantages for AI developers and other businesses relying on TDM in detriment of publishers' intellectual property rights. In addition, the proposed new exception will damage the incentives for publishers to develop AI licensing businesses. This unintended consequence challenges the consultation paper's assumption that a new exception would promote AI development and wider economic growth in Hong Kong. In fact, publishers' AI businesses play a fundamental role in increasing and consolidating region-relevant AI offerings. The exception, on the contrary, will erode incentives for publishers to invest in content development and to work with AI trainers to further develop new and reliable data sources. Given the hunger for more high quality data for AI training efforts across the world, sacrificing the incentives for publishers to keep producing new quality content and new reliable data sources does not seems to align with responsible development of AI and counters efforts to establish a solid knowledge economy to which copyright owners are so central.

AI is currently at the forefront of multiple regulatory and policy debates at national and international levels. Policy discussions seek to address risks, gaps, and shortcomings in key areas such as ethics, human rights, labor law, societal welfare, financial and economic systems, privacy rights, and the role of AI in disinformation, including debates around impacts and processes to secure information integrity. These considerations should not be ignored as possible effects of the new proposed exception. The proposed new exception would facilitate the massive use of unverified data to enable the creation of unverified content, creating rather than eliminating regulatory obstacles. Recent policy discussions at the international level caution against precipitating action, while drawing attention to the need to devise risk-based and human centric regulatory approaches. Accountability about the sources used to train the algorithms is key, facilitating the prevention of possible biases in the output.

Against this background, the quality of content used to train machines emerges as a critical issue, not only in AI related policy debates, but also as an element for the improvement and future development of these systems. Licensed quality content plays a role in promoting safe and ethical development of AI systems, a priority in international policy discussions, which cannot be achieved through the adoption of an exception.

Is copyright licensing commonly available for TDM activities? If so, in respect of which fields/industries do these licensing schemes accommodate? Do you find the licensing solution effective?

Despite dogmatic beliefs that overbroad exceptions are required for AI development, responsible AI developers are seeking licenses from publishers to ensure the accuracy and integrity of information used to train and operate AI based systems. Responsible AI companies value the quality of training data, conscious of the dangers of indiscriminate scrapping (garbage in-garbage out) and of the potential of non-curated training data undermining the value of their end products. These AI developers have been seeking licenses for different uses, from training data (to help improve relevance and performance of AI systems, for example) to related services such as developing accurate automated citation referencing, with improved speed and accuracy. Promising licensing markets are developing and will continue to develop within the existing legal framework.

Publishers maintain that the first response to access is licensing. Licensing is simple in the digital age, and there is no excuse to ignore it. Direct and voluntary collective licensing models are ubiquitous, flexible, efficient, and continue to evolve. Technology companies can respect copyright and avoid liability. There is no need to except them from the well-established rules of copyright. There is no public interest that justifies prioritizing the profit of technology businesses by removing publishers' entitlement to exercise and benefit from their intellectual property.

Publishers have always embraced technology and technology partners. They are early adopters of many AI tools and believe fundamentally that human expression and technological innovation are symbiotic. However, the legal framework is clear and remains fit for purpose. Collecting, handling, storing, and copying works of authorship to train AI models implicates the exclusive rights of authors, which cannot be ignored. Generative AI companies must license works of authorship that they do not own—and seek to use for their gain—in the manner directed by the rightsholder.

What conditions do you think the Proposed TDM Exception should be accompanied with, for the objective of striking a proper balance between the legitimate interests of copyright owners and copyright users, and serving the best interest of Hong Kong? Are there any practical difficulties in complying with the conditions?

Our view is that no exception should be introduced at this stage. Comparative law experiences, extensively illustrated in the consultation paper, demonstrate the many problems resulting from existing TDM exceptions, notably that it is impossible to assert that no subsequent uses are made in such a way that non-beneficiaries of exceptions will not in fact benefit from them and further advance their businesses on unlicensed content. Furthermore, requirements such as lawful access or use of secure network systems are not easily verifiable, especially in absence of transparency obligations and supervision mechanisms.

According to the consultation paper, Hong Kong's copyright law already allows for fair dealing for research and private study. Licensing agreements supplemented by the existing exception can promote AI development without affecting the integrity of copyright protection by continuing to enable a balanced and sustainable AI ecosystem based on licensing of quality content.

We respectfully invite the Government of Hong Kong to consider alternative solutions to support and foster licensing and to become a hub for responsible, ethical, and solid AI.

We thank you for your time and consideration and remain available for any additional information that you may require.

The Association of American Publishers (AAP) represents the leading book, journal, and education publishers in the United States on matters of law and policy, advocating for outcomes that incentivize the publication of creative expression, professional content, and learning solutions. As essential participants in local markets and the global economy, our members invest in and inspire the exchange of ideas, transforming the world we live in one word at a time.

The Association of Canadian Publishers (ACP) is the voice of Canadian independent book publishers. Our diverse membership consists of approximately 115 Canadian-owned and controlled book publishers from across the country, and range both in company size and editorial mandate.

Börsenverein des Deutschen Buchhandels (the German Publishers and Booksellers Association) represents the interests of Germany's book industry in the political and public spheres. It was founded in Leipzig in 1825 and currently has roughly 4,500 members, which include booksellers, publishers, wholesalers and other media companies. As both a trade association and a cultural organisation, the Börsenverein works to promote books and reading whilst also fostering fair copyright laws and the preservation of Germany's policy of fixed book prices. It also seeks to promote cultural diversity and uphold the right to freedom of expression.

The Federation of European Publishers (FEP) represents 29 national publishers' associations in Europe. Their members are publishers of books, learned journals and educational content, in all formats.

The International Publishers Association (IPA) is the world's largest federation of national, regional and specialist book publishers' associations. Established in 1896, our membership comprises 101 organizations from 82 countries around the world. The IPA is based in Geneva and is an accredited observer at the World Intellectual Property Organization as well as an accredited non-governmental organization enjoying consultative relations with the United Nations.

The Singapore Book Publishers Association (SBPA) represents 72 companies in our vibrant publishing industry, contributing some 1.7b S$ annually to the local economy. Our members include some of the world's largest multinational publishing companies with operations or operational headquarters in Singapore, home-grown companies with broad international reach, as well as small companies and startups.

Sindicato Nacional dos Editores de Livros (SNEL) is the representative of publishers in Brazil for over 80 years and plays a crucial role in promoting reading and books, defending the sector, and fostering artistic and intellectual production.

STM - the International Association of Scientific, Technical and Medical Publishers supports its members in their mission to advance trusted research worldwide. Our over 140 members based in over 20 countries around the world collectively publish 66% of all journal articles and tens of thousands of monographs and reference works. As academic and professional publishers, learned societies, university presses, start-ups, and established players we work together to serve society by developing standards and technology to ensure research is of high quality, trustworthy and easy to access.

The UK Publishers Association is the member organisation for UK publishing, representing companies of all sizes and specialisms. Its members produce digital and print books, audiobooks, research journals and educational resources across genres and subjects. The work of the UK Publishers Association focusses on ensuring that the value of publishing and its economic and cultural contribution are recognised. The UK Publishers Association engages with governments in respect of key policies and laws that impact and support publishing. It advocates for strong copyright laws at home and abroad. It serves its members through policy, legal and public affairs work, running industry campaigns, and through a range of key member services including export and content protection (anti-piracy) support.