

# 公众咨询意见

## 针对中国香港知识产权署《版权与人工智能公众咨询》

1. “健全的版权制度是推动创意、创新科技产业以至整个数码和知识型经济发展的基石”。根据 2024 年 7 月 16 日立法会工商及创新科技事务委员会的讨论文件（简称“**讨论文件**”），现政府拟完善《版权条例》对人工智能技术发展（尤其是生成式人工智能）所提供的保障，并展开公众咨询。
2. 现行香港第 528 章《版权条例》已经包含了一些适用于电脑产生作品的条款，但是，这些规定如何适用于人工智能相关的话题并不明确，条例的模糊给该领域从业者尤其是生成式人工智能技术开发者、服务提供者带来了困惑。各司法管辖区采取了不同的法律规定和市场实践进一步加深了该困惑，有必要借此修订机会予以澄清，既能形成友好的营商环境，又能促进人工智能技术在港的健康发展。
3. 借此契机，本意见函结合现行香港《版权条例》，针对香港知识产权署提出的《版权与人工智能的公众咨询（2024 年）》提出几点意见，以期促进人工智能技术在港的发展及相关应用市场的繁荣：

### A. 生成式人工智能生成的作品（人工智能生成作品）的版权保护

4. 关于人工智能生成作品的作者问题。根据现行《版权条例》第 198（1）条的定义，“电脑产生（computer-generated）作品”是指“在没有人类作者的情况下由电脑产生的”作品。进一步，第 11（3）条规定，“如文学作品、戏剧作品、音乐作品或艺术作品是电脑产生的，作出创作该作品所需的安排的人视为作者。”“公众咨询”认为，现行《版权条例》给电脑产生作品提供了比较清晰的规范指引，但本意见函认为，电脑产生作品的规定在英国出现于 1988 年、在本港出现

于 1997 年，彼时尚未出现具有创作能力的人工智能，无法预测当今存在的新问题，该规则未必适用于人工智能生成作品，实践中也出现了困惑。尤其是“作出创作该作品所需的安排的人视为作者”在新语境下的模糊定义，导致了交易安排的困难，有必要沿袭保护“电脑产生作品”的传统，明确“人工智能生成作品”的定义及权利归属。

5. 关于电脑产生作品的定义及保护，现行《版权条例》与《英国版权法》（CDPA1988）类似，可以借助英国判例理解“作出创作该作品所需的安排的人视为作者”的含义。Nova Productions Ltd. v. Mazooma Games Ltd.& Ors Rev 案是英国至今唯一就“计算机生成之作品”进行裁判的判例，该案强调游戏画面组成元素和相应计算机程序的设计过程，并据此认定程序设计员是计算机生成之游戏画面“创作所必需的安排之承担人”。<sup>1</sup>也就是说，“创作该作品所需的安排”是带有创作性的难以替代的安排，比如程式设计。中国内地亦有秉持相同思路的案例：对于 Dreamwriter 机器人自动撰写的文章，法院认定其著作权属于原告公司，因为文章数据的输入、处理、触发条件的设定、文章框架模板及语料的选择、智能校验算法模型的训练等均由原告公司的主创团队完成选择与安排。<sup>2</sup>

6. “公众咨询文件”的注释 7 提到：“人工智能生成作品是指由生成式人工智能系统按用户的指令创作和生成的作品，过程中不涉及人类作者”。从这个论断可以判断，《版权条例》本意是遵循现行“电脑产生作品”判例对“创作该作品所需的安排”的理解，认为商业化权利（财产性权利）属于设计、训练人工智能的人，从而保护设计、训练人工智能所投入的大量资金成本，而提供指令的用户不属于此处的“作出安排的人”。

7. 然而，虽然人工智能生成作品属于“在没有人类作者的情况下由电脑产生的作品”，但是“作出创作该作品所需的安排的人”的表述并不清晰，该模糊的表述使得人工智能设计者、训练者不确定是否享有相关权利，更无法在与人工智能服务提供者的交易中合理评估合同标的价值。例如在用户利用人工智能生成图片的情形中，用户使用的大模型或者模板图由人工智能设计者提供，设计者的

---

<sup>1</sup> 2006 EWHC 24(Ch)(20 January 2006), para. 105.

<sup>2</sup> (2019) 粤 0305 民初 14010 号民事判决书

选择与最终形成的图片风格息息相关，但用户如果输入了比较充分和明确的提示词甚至提供了图片作为底稿，是否可以认为用户的选择与最终形成的图片内容亦存在某种联系、某种安排？人工智能生成作品的“作出创作该作品所需的安排的人”，似乎也可以包括用户，是否人工智能设计者和用户应共同成为版权的第一拥有人？

8. 希望《版权条例》考虑到人工智能生成作品的上述特点，对作者身份及版权归属予以明确，该行动有助于减少法律的不确定性，增强投资者对人工智能技术研发的信心，同时有助于评估人工智能的交易价值，确保开发者、提供者能够合法地享有其创新成果的经济回报。同时，为了平衡用户使用人工智能的目的及合法权益，应当为用户提供使用生成作品的低程度保护，包括允许用户能够以非商业用途无偿使用人工智能生成作品，并同等享有《版权条例》第 96 条“免被作品虚假署名的权利”。

9. **因此，本意见函建议：**摒弃“作出创作该作品所需的安排的人”的模糊表述，明确电脑产生的作品包括人工智能生成作品，设计、训练人工智能的人视为作者；操作人工智能而生成作品的人有权以非商业用途传播人工智能生成作品并享有免被作品虚假署名的权利。如此明确，亦有助于人工智能服务提供者评估用户协议的有效性，评估是否能够声明人工智能生成作品的版权拥有权，同时保持文本在法律上的有效性。

## **B. 人工智能生成作品侵犯版权的法律责任**

10. “公众咨询文件”的 3.6 (a) 与 (b) 强调，根据现行规定：假如人工智能系统开发者/程式编制员/操作员可以控制该人工智能系统并有能力防止该项侵权（例如通过采取可行的系统措施或保障措施），则有理由认为该人工智能系统开发者/程式编制员/操作员为侵权承担主要的法律责任。如最终使用者向人工智能系统下达的指令特意指示、提议、示意或间接指示该人工智能系统制作某项版权作品的侵权复制品，这亦可支持向最终使用者提出侵权申索。该陈述体现出，人工智能服务提供者在《版权条例》中不属于“服务提供者”，不可能依据第 IIIA

分部“法律责任的限制”得到责任的豁免。可是，此结论不符合我国保护人工智能服务发展的方向，不利于本港发展区域知识产权贸易中心。

11. 现行《版权条例》第 65A (2) 条定义，“服务提供者 (service provider) 指借电子设备或网络 (或同时借两者)，提供任何连线服务或为任何连线服务操作设施的人”。目前，《版权条例》对 65A 条的“连线服务”不能确认是否包含人工智能服务提供者，但生成式人工智能服务提供者实际并不直接影响、控制生成作品的内容，用户既可以借助生成式人工智能服务生成不侵权作品，也可能借助实施侵权行为，生成式人工智能服务提供者本质上可以属于服务提供者。中国内地的判例已经显示，通过可编程接口的方式接入第三方服务商系统的方式，面向用户提供生成式人工智能服务的，属于生成式人工智能服务提供者。<sup>3</sup>

12. 服务提供者承担侵权责任，依据的是过错归责原则。而生成式人工智能服务提供者难以精准控制模型训练语料的来源，特别是涉及不知名作品时，用户通过人工智能生成服务生成侵权作品的可能性被大大增加，而生成式人工智能服务提供者对此侵害结果通常不具备主观过错。所以，《版权条例》不应仅仅允许人工智能服务提供者在承担责任后向终端用户追索赔偿，而应当允许服务提供者根据指称侵权通知及时地移除、断开侵权内容后 (参考《版权条例》第 88A 条至第 88J 条的规定) 不再承担侵权责任。如果人工智能服务提供者提供了版权投诉渠道、提供了及时屏蔽侵权行为的措施的机制，且没有证据证明其存在侵权过失，却会因为个别用户的不当行为而需要承担侵权责任，相当于要求人工智能服务提供者承担事前的、还量化的版权审查义务，该义务过于严苛，成本高昂且不利于人工智能应用场景的创新和发展。

13. **因此，本意见函建议：**将生成式人工智能服务提供者纳入第 IIIA 分部的服务提供者的范围，给予其适用“避风港原则”即法律责任的限制的身份，不仅能够合理地减轻其在版权保护方面的负担，还能为人工智能技术发展提供更为宽松的法律环境。同时，服务提供者须提供便捷的版权投诉渠道和快速的判断、移除机制，便于权利人发送侵权通知并顺利维权。

---

<sup>3</sup> (2024) 粤 0192 民初 113 号民事判决书。

### C. 拟引入的特定版权豁免

14. 本意见函支持引入“拟议文本即数据开采豁免”且支持该豁免可适用于商业用途，其论据已经由“公众咨询文件”充分阐述。本意见函进一步从国际义务角度进行强调。

15. 在“拟议文本及数据开采豁免”不违反我国的国际义务，暂无任何国际公约明确不得引入可涵盖商业用途的“拟议文本即数据开采豁免”。《伯尔尼公约》以及《与贸易有关的知识产权协议》(TRIPs)第13条规定了各缔约国设置版权特定豁免的测试方法：“成员应将专有权的限制或例外局限于某些特定的情形，并且不得与作品的正常使用冲突，也不得不合理地损害权利人的合法利益”。《世界知识产权组织版权条约》第10条第2款强调，各缔约方在适用《伯尔尼公约》时，应对任何限制或例外限于符合“三步测试法”的特殊情况。“拟议文本即数据开采豁免”即能符合“三步测试法”的要求：第一步，该情形仅限于文本及数据开采，并非普适的情形；第二步，该情形未与作品正常使用相冲突，“拟议文本即数据开采豁免”针对的是“复制”版权作品进入模型训练，且该训练过程仅涉及“暂时复制”——大模型对版权作品的训练结果，以训练数据集而不以版权作品本身的形式进行保存，该利用方式与作品正常使用（欣赏作品所表达的思想或感情）相去甚远；第三步，该情形并未不合理地损害权利人的合法利益，只要大模型不输出侵害他人版权的内容，版权作品的主要利用价值（传播价值）就没有受到影响，虽然版权作品本可能因为被用于训练而得到经济效益，但该经济效益一方面是版权拥有人期待的次要价值，另一方面可能要低于从模型训练者处收取效益的事务成本，理性的版权拥有人一般不会以销售“训练素材”为收益点。

16. 根据“公众咨询文件”4.8和4.9，一些主要的司法辖区已经在各自的版权法中引入特定的版权豁免，以涵盖相类似的文本及数据开采活动，尤其是日本和新加坡所订文本及数据开采豁免适用范围广泛，如本港在“拟议文本即数据开采豁免”行动迟缓，将导致优秀的人工智能开发企业流向其他环境宽松的国家或地区，从而失去人工智能发展的先机、丧失经济增长的机会。

17. 另一方面，“公众咨询文件”第 4.17 考虑向版权拥有人提供足够保障，以维持适当的权益平衡，并拟对“拟议文本及数据开采豁免”施加条件，但例举的施加条件的方式可操作性较差，例如使用者必须合法取用版权作品、版权作品有特许计划可供应用，或版权拥有人已经表明保留其权利（即“选择退出”选项）。开发、训练或提升人工智能模型所需的数据往往达到数千万甚至上亿级别的数据记录，可能需要借助爬虫从互联网上自动抓取数据，依靠这种“规模效应”人工智能模型才能出现深层次的理解、推理和创造力，在这种情况下，要一一确保训练数据的使用符合版权法基本是不可能的，对应的，要一一确保取用合法、一一形成特许计划进行应用或一一确认版权拥有人是否已经“选择退出”是非常艰难的，尤其是对于中小型企业 and 初创公司，这种成本可能是无法承受的负担，将大大损害“拟议文本及数据开采豁免”的意义。

18. 为了引入“拟议文本及数据开采豁免”的同时向版权拥有人提供足够保障，需要讨论建立使人工智能开发者能够自动化识别、操作的规则。例如借鉴“开源协议”设置明确、渐变的特许计划并便于人工智能模型在训练时自动处理，或者要求权利人的事先声明须符合机器可自动识别的要求（类似于 robots 协议），令人工智能开发者在采集数据时自动跳过带有此类声明的作品。不过，如果大多数权利人随手预留退出的声明“拟议文本及数据开采豁免”也会失去意义，无法真正发挥促进技术创新的作用。豁免条款的目的在于提供一个宽松的法律环境，鼓励数据驱动的研究和开发，若设置过多限制性措施，将背离这一初衷，导致豁免政策无法达到预期效果。是否有必要设置此类的保障，如果不设置此类保障会给版权拥有人造成何种困扰，还是有待商榷的。

19. **因此，本意见函提醒：**引入“拟议文本及数据开采豁免”豁免，可以减少法律障碍，鼓励企业和研究机构积极开展数据分析和人工智能模型训练。但如果《版权条例》对“拟议文本及数据开采豁免”施加过多限制，将会阻碍人工智能技术的研发进程，影响整个行业的创新能力。从这一角度出发，虽然为版权拥有人提供保障具备一定合理性，但是需要贯彻平衡利益原则，衡量提供保障的方式给人工智能从业者施加的负担（合规审核成本），是否远远超出不提供保障给版权拥有人造成的损害，如果答案为“是”，那么就不应该选择此项保障方式，

否则就损害了引入“拟议文本及数据开采豁免”的价值。

#### **D. 其他与生成式人工智能相关的、需要提请关注的课题**

20. 关于日后如果发生争议，生成式人工智能开发者是否需要披露其所使用的训练素材的问题，本意见函提醒：《版权条例》应谨慎在版权体系下为生成式人工智能开发者设置如此义务。

21. 一般情况下，人工智能生成物侵犯他人版权的情形仅发生在“输出端”，因此版权纠纷也主要关注生成式人工智能是否会输出侵害版权的内容。换言之，即使生成式人工智能使用的训练素材侵犯他人的版权，但如果开发者采取了合理的屏蔽和过滤措施，避免输出侵害版权的内容，那么版权拥有人实际难以发现其作品被用于训练并完成举证、形成争议，这种情况下也不存在对版权拥有人造成损害的结果。

22. 此外，可能有观点认为披露训练素材有助于解决版权争议中“接触可能性”认定的问题。但是，争议方其实不需要通过披露训练素材内容来证实这一点。日本文化厅自2023年6月开始对人工智能与著作权的相关问题进行审议，涉及到两种认定“接触”的方法：（1）如果版权拥有人能够证明疑似侵权者可能接触到其作品，则可以推定存在“接触”；（2）如果生成式人工智能生成了与现有作品相似的内容，通常会被推定存在“接触”。<sup>4</sup>其中，后者即属于无需披露训练素材即完成判断“接触”的方法。

23. 无论从何种角度观察，要求生成式人工智能开发者披露所使用的训练素材都不会对版权争议有实质性影响，基于版权争议而要求生成式人工智能开发者披露训练素材的论据是不充分的。中国内地的案例能够体现此问题：在版权拥有人起诉生成式人工智能服务提供者的案件中，版权拥有人要求将案涉奥特曼物料从其训练数据集中删除，而法院最终判令被告采取关键词过滤等防范措施，重点

---

<sup>4</sup> 日本文化厅：

[https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05\\_07/pdf/94011401\\_02.pdf](https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_07/pdf/94011401_02.pdf)

在于，防范程度须达到“用户正常使用与奥特曼相关的提示词，不能生成与案涉奥特曼作品实质性相似的图片”的程度。<sup>5</sup>这一案例的精神即在于，从“输出端”制止版权侵权行为，即足以解决版权争议。

24. 相反，本意见函关注到了反对“要求生成式人工智能开发者披露所使用的训练素材”的论据。在实务中，生成式人工智能模型所需的训练数据来源复杂且数量庞大，其算法决策过程也不需要完全透明，因此训练数据及相应训练的数据集结果通常会被企业作为商业秘密进行保护，具备极高的商业价值。相比之下，版权争议通常涉及的损害金额较少，因较少的损害金额而要求生成式人工智能开发者披露重要的商业机密，并承担被泄露、窃取的风险，显然不符合比例原则。

25. **因此，本意见函提醒：**只要生成式人工智能开发者采取合理措施，确保生成内容不侵犯版权拥有者的利益，双方通常不会产生纠纷。除非人工智能开发者自愿或履行在先合约安排需要向对方披露人工智能模型训练使用的素材，否则版权拥有人无权要求生成式人工智能开发者披露使用的训练素材。披露这一行动不仅对解决版权纠纷没有实际裨益，反而会增加人工智能开发者商业秘密泄露的风险，应谨慎在版权体系下为生成式人工智能开发者设置如此义务。

---

<sup>5</sup> (2024)粤0192民初113号民事判决书。