

第二部分 | 从演算法到著作权：人工智能 时代下的版权反思

2025 年 2 月 13 日

3 输入端视角：数据与版权侵权

在[本文的上篇](#)中，我们深入探讨了 AI 生成作品著作权保护的「产出面向」，剖析在这个快速演变的领域中，关于拥有权与作者归属的复杂议题。现在，我们将焦点转向问题的「输入面向」：用于训练 AI 模型的大型数据集往往涉及受著作权保护材料的复制。由于专有 AI 系统的不透明性，许多版权持有者面临难以确认哪些数据被取用的困境。当前全球多宗诉讼案件更凸显一个关键问题：使用受著作权保护的作品来训练 AI 模型，是否必须事先取得版权拥有人的明确授权？

3.1 香港拟引入的版权豁免条款

香港现行的《版权条例》并无针对计算数据分析与处理的版权豁免。因此，涉及 AI 模型开发与训练的活动可能面临版权侵权风险。为此，政府拟引入特定版权豁免，容许合理使用版权作品作电脑数据分析和处理，涵盖非商业和商业用途(「**文本及数据开采豁免**」)。

为平衡促进 AI 技术创新与保护拥有人权益，政府提出多项保障措施：

- (a) 合法存取：须合法取用版权作品
- (b) 退出机制：如版权拥有人明确表示选择保留其权利，则相关文本及数据开采活动属未经授权
- (c) 使用限制：不得后续使用基于文本及数据开采豁免所复制的版权作品
- (d) 附加条件：规定使用者备存及披露版权作品来源的纪录、保密数据储存、限制保存时间/目的等



3.2 创新与版权的平衡

香港的建议框架借镜欧盟、日本与新加坡做法，这些地区均已建立适用商业与非商业用途的文本及数据开采豁免。尤值关注的是，欧盟的豁免架构赋予拥有人退出权，相较之下，英国则将豁免范围严格限于非商业性研究用途。

尽管英国政府近期曾考虑将文本及数据开采豁免扩展至商业用途，但因版权拥有人反对而撤回提案。关键异议在于此类豁免可能产生双重负面效应：其一将侵蚀拥有人透过授权金获利的既有商业模式，其二可能颠覆现行文本及数据开采的市场授权机制（如文本及数据开采活动之授权协议）。其咨询文件指出，透过适当保障版权拥有人的措施可缓解这些潜在弊端，并强调任何文本及数据开采豁免必须在权利保护与产业发展间取得平衡。促进数据开采与分析能让创作者运用现有作品透过 AI 产出新颖原创内容，有望推动教育、医疗、娱乐等跨领域创新。

更值得深入探讨的是，退出机制的具体实施层面仍存在诸多待解难题。就版权拥有人的角度而言，核心困境体现在两个维度：其一为权利行使方式的不确定性（究竟应透过缔结补充协议或发布单方权利保留声明）；其二为权利主张时效性的模糊（特别是在 AI 企业已完成模型训练的既成事实情况下）。而对 AI 开发商来说，真正的技术挑战在于如何建立有效的作品过滤系统，确保能精准辨识并筛除已声明退出文本及数据开采豁免授权范围的特定权利作品。

4 结论：寻找数位时代的新平衡

经由本系列上下两篇的分析，我们已深入剖析 AI 技术发展所衍生的多重法律困境及其对现行版权体系的结构性冲击。当前拥有人与 AI 产业间的权益拉锯虽仍存在显著张力，但建构具包容性的制度平衡机制，实为调和各方利益、最大化社会创新效益的关键所在。就香港的特殊定位而言，实行包含商业用途的广泛性文本及数据开采豁免，辅以完善的作品来源识别系统与拥有人退出权保留机制，可视为最具实践智慧的立法方向。



AI 与版权法的交汇既带来挑战也孕育机遇。透过重新思考现有框架并促进与持份者的实质对话，方能在保障创作权益与促进技术创新之间，开辟出一条兼容并蓄的法治路径。此种新型态的法律生态系统，不仅应为人類与机器的协同创作预留发展空间，更须确立符合 AI 时代特征的权利分配与责任归属原则。

撰稿人



朱庆华律师

高级合伙人

+852 2532 5490

gchu@fclklaw.com.hk



邢珈晞律师

律师

+852 2532 5441

aying@fclklaw.com.hk

执业范围

知识产权

免责声明

本文包含的资讯仅供一般指导，不应依赖或替代具体建议。对于因依赖这些材料中包含的任何资讯而可能造成的任何损失，我们不承担任何责任。对于任何此类资讯的准确性、有效性、及时性或完整性，不提供任何明示或暗示的陈述或保证。本网站特此完全保留与本文内容相关的所有专有权利。

