



特别关注

合理使用制度是应对生成式人工智能机器学习的最优解？

□胡开忠 江璐迪

自2025年1月底发布以来，DeepSeek持续引发广泛热议。凭借出色的文本、图像等内容生成能力和理解推理能力，DeepSeek迅速崭露头角，成为继ChatGPT之后的又一现象级人工智能产品，被迅速应用于内容创作、在线客服、教育培训、企业运营等多个领域。

随着生成式人工智能的普及和应用，知识产权保护也迎来了前所未有的挑战。其中，位于生成式人工智能生命周期前端的机器学习的版权合法性问题尤为突出。由于生成式人工智能技术的发展离不开依赖于海量数据训练的机器学习，版权作品

作为优质的信息资源，是机器学习过程中必要的数据来源。根据著作权的事前授权原则，利用版权作品进行机器学习需要事前获得著作权人的许可。

然而，面对海量的作品使用需求，人工智能的开发者需逐一与著作权人展开漫长复杂的谈判以获取授权，将产生极高的交易成本。于是，未经著作权人许可利用作品开展机器学习，成为人工智能产业发展的普遍现象，导致了作品市场的利益失衡。在此背景下，著作权人纷纷起诉人工智能企业，以期司法实践能积极回应新技术所引发的作品利用难题，构建机器学习场景下的版权治理框架。

生成式人工智能机器学习行为的特殊之处

从司法实践的诉讼情况来看，无论是北京互联网法院正在审理的“全国首例涉及AI绘画大模型训练著作权侵权案”，还是美国的30余起机器学习相关的版权纠纷，有关机器学习的版权合法性争议主要集中在生成式人工智能领域。这一现象的根源在于生成式人工智能机器学习行为的特殊性和复杂性。

与传统的文本与数据挖掘行为不同，生成式人工智能机器学习不仅是将作品视为事实信息予以分析并提取规律，还涉及对作品元素的表达性使用。此外，这项技术对作品市场的影响颇为复杂，一方面，生成式人工智能开启了“机器创作”的新模式，其高效率的创作能力客观上为人类创作者提供了更丰富的语料素材。同时，

利用生成式人工智能从事创作活动降低了作品创作成本，提升了艺术创作效率，促进了公众自由表达的实现。另一方面，更廉价的人工智能生成物会对人类作者的作品产生一定的替代效应，损害人类作者的整体利益。长此以往，这可能导致“劣币驱逐良币”，平庸、同质化的机器生成内容取代了独特、多样的人类创作作品，阻碍文化表达的多样性发展。

生成式人工智能机器学习的定性之难也折射出一个更本质的问题，即当版权保护和技术创新的制度发生冲突时应如何协调？不同的制度选择体现了不同的价值立场。在学界，多数学者主张采用更宽松的监管政策，即适用合理使用制度为人工智能产业发

展减负。他们认为，数据训练中的作品使用行为具有“非特定性”，属于算法黑箱中过程性使用行为和非表达性使用，不构成著作权侵权。此外，基础模型作为人工智能时代的新型基础设施，在技术效果上具有普惠性，认定生成式人工智能机器学习构成合理使用，不仅在国内层面有利于中小型企业释放创新潜力，而且在国际层面也能够吸引更多人工智能研发者，助力我国人工智能产业的高质量发展。

另一部分学者则以作品市场的利益平衡为观点，主张适用法定许可制度规制生成式人工智能的机器学习行为。他们指出，合理使用制度的滥用可能进一步巩固技术研发者的市场垄断地位，加剧作品市场的利益失衡，甚至导致人

类作者因难以维持生计而退出创作领域，最终损害文化表达的多样性。而法定许可制度作为一个折中方案，以公平价值为理念，在简化作品许可使用程序的同时，保障了著作权人的经济利益，为其提供谈判筹码，兼顾了版权保护和

此外，由于算法黑箱的不可知性，在人工智能致害的场景下，侵权责任认定的因果关系难以成立，事后救济面临障碍。法定许可制度作为一种事前规制手段，要求基础模型的研发者在技术开发初期就将著作权人的利益纳入考量。这不仅有助于督促研发者采取更多的技术措施，降低输出侵权结果的概率，还能有效内化技术发展中的负面外部性。

生成式人工智能机器学习行为的合理使用分析

生成式人工智能机器学习行为的定性目前仍存在较大争议，尽管多数学者倾向于认为生成式人工智能的数据训练应构成合理使用，但这一主张仍然需要在具体实践中通过“四要素测试法”加以检验。

使用作品的性质和目的。是否构成转换性使用在合理使用的认定中发挥着关键作用，“转换性”程度较高的作品使用行为，即便是商业性使用也可能被认定为合理使用。持有“合理使用说”的学者认为，机器学习是技术开发的必要环节，其对作品的使用是一种以实现某一技术功能的转换性使用。诚然，对于并不应用于内容生成的基础模型而言，作品数据训练的目的是为提取作品中的事实信息，这种功能性使用通常不会被认定为侵权。然而，某些创意生成的语言模型，其输出的结果可能会与训练作品构成实质性相似，或其应用的市场与原作品的市场存在重叠，此时上述观点便缺乏解释力。在司法实践中，法院可能会裁定输出的结果、模型的部署等构成著作权侵权，不受合理使用原则的保护。比如，在广州互联网法院审理的“奥特曼案”中，由于案涉

生成式人工智能输出的结果与奥特曼美术形象构成实质性相似，法院最终认定被告构成著作权侵权并要求其承担法律责任。这表明，生成式人工智能机器学习的合理使用判定的关键在于分析输出结果与版权作品是否构成实质性相似，若答案是肯定的，则需要进一步考察使用行为的性质和目的是否具有足够的“转换性”。通过个案的具体事实进行综合判断，方能勾勒出生成式人工智能机器学习的合理使用轮廓。

考察使用行为是否出于商业目的或非营利的教育目的。对该要素的分析需要结合转换性使用的程度、对原作品市场的影响等其他因素进行综合判断。一般来说，非商业性的作品使用行为与合理使用制度内在的公益价值相契合，有较大可能构成合理使用。然而，在人工智能的产业实践中，从事基础模型研发的主体多为商业性质的互联网企业或科技公司，这使生成式人工智能机器学习的合理使用认定于该要素上处于劣势。此外，美国最高法院在2023年“安妮·霍尔案”中强化了“商业性使用”对合理使用认定的不利影响，提升了该要素在判定中的重要性，进一步增加了以商业目的进行机器学习被认定为合理使用的难度。

判断是否为善意使用。生成式人工智能机器学习是否构成善意使用，与其数据来源的合法性密切相关，即是否规避或破坏著作权人所采取的技术措施。《生成式人工智能服务管理暂行办法》第七条第一项规定，开展数据训练处理活动应使用具有合法来源的数据和基础模型。因此，数据来源合法性是考察机器学习版权合法性的前



资料图片

提。例如，在美国的“纽约时报案”中，ChatGPT绕过了纽约时报所采取的技术保护措施“付费墙”，以逐字复制原作品的方式给未付费用户提供了文章《雪崩：隧道溪的灾难》(Snowfall: The Avalanche at Tunnel Creek)。这一行为将阻止机器学习构成合理使用。

被使用作品的性质。版权保护的强弱与作品的性质紧密相关，虚拟作品相较于事实作品能获得更充分的版权保护，因此，使用事实作品更容易构成合理使用，使用虚拟作品则反之。当前生成式人工智能被广泛应用于广告、游戏等创意产业，其输出结果更容易与虚拟作品或艺术作品构成实质性相似，存在较高的著作权侵权风险。

需考量被使用部分在整个原作品中所占比例和质量。使用作品的比例要与其所实现的正当目的或功能相适应。以生成式人工智能机器学习为例，作品输入和模型训练阶段的作品使用是为提取海量作品中的语言表达规律，是技术开发过程中的非表达性使用，此时即便是完全地逐字复制也能被认定为合理使用。在结果输出环节，生成式人工智能根据指令生成符合用户需求的内容，此过程涉及对原作品的表达性使用。理论上应以少量使用为限，过量地使用原作品或使用

其核心部分均不利于合理使用的认定。

使用行为对原作品潜在市场或其价值的影响。该要素的考察与使用行为的转换性密切相关，也需要区分因构成著作权侵权而引发的市场替代效应与基于非著作权侵权行为所导致的市场竞争损害。通常而言，使用行为的转换性越强，越容易被认定为著作权侵权，越有可能在“原作品的潜在市场”产生替代效应。以“下一个伦勃朗”项目为例，计算机通过学习伦勃朗的绘画风格、作品细节以及诸多伦勃朗作品片段，最终形成的人工智能系统能够根据指令“创作”出既具有伦勃朗绘画风格，又完全不同于其任何现存作品的全新数字绘画作品。然而，若伦勃朗的作品仍受版权保护，该人工智能生成的衍生作品将对伦勃朗作品市场产生替代效应，并影响其潜在市场，此时，机器学习便难以被认定为合理使用。

通过上述分析可知，并非所有生成式人工智能的机器学习行为都能被认定为合理使用。在输出端非转换性使用与转换性使用的交织，使生成式人工智能机器学习的合理使用判定处于灰色地带，需要结合个案中的具体事实进行综合判断。即便主张生成式人工智能的机器学习构成合理使用，研发者要免于著作权侵权责任，仍要确保人工智能的输出结果与现有作品不构成实质性相似。因此，合理使用制度的适用因个案中输出结果的差异具有高度的不确定性，这既无法为基础模型研发者提供明确的法律指引，也难以完全满足行业实践的现实需求。

(作者胡开忠系中南财经政法大学知识产权学院教授、博士生导师；江璐迪系中南财经政法大学知识产权学院2023级博士生)

一家之言

构建多层次、体系化的版权治理框架

□胡开忠 江璐迪

生成式人工智能机器学习的版权合法性问题关乎版权保护与技术创新之间的价值平衡。面对合理使用制度的失灵，笔者主张以利益平衡为理念，依据生成式人工智能机器学习被认定为合理使用的可能性光谱，既从著作权法的制度工具入手，又向商业实践中的自治措施延展，构建多层次、体系化的版权治理框架，以期对相关市场主体开展商业活动提供稳定的预期。

增设非营利性生成式人工智能机器学习的合理使用情形

合理使用制度的规范目标在于调和多元价值冲突，以增进社会公共福祉。非营利性的机器学习行为不仅便于公众享受技术进步带来的益处，而且在产业实践中占比较小，对著作权人的合法权益影响有限，被认定为合理使用的可能性极高。然而，在我国合理使用制度规定的12种法定情形因使用主体、数量和范围等方面的诸多限制，未能有效涵盖涉及海量作品使用的机器学习。因此，有必要单独为非营利性的机器学习设置法定情形，从而为市场主体提供明确的合规指引。

还可以考虑设立商业性生成式人工智能机器学习的法定许可情形，商业性的生成式人工智能机器学习因输出结果可能与版权作品构成实质性相似，难以受到合理使用原则的保护。这导致基础模型的研发者在实践中仍要采取技术措施，尽可能地避免输出版权侵权内容。面对合理使用制度的失灵，法定许可制度是更为适宜的制度选择，其作为事前的责任限制规则，在简化交易程序的同时，保障了著作权人的利益。

此制度的实施依赖于集体授权机制，应从技术支持和制度保障两方面着力完善我国的集体管理制度，推动该制度的运行。此外，法定许可制度能够扭转著作权人在市场中的弱势地位，激励技术研发者在市场尚未形成复杂的交易环境前，与著作权人自主协商，或通过产业实践中的自治方案降低著作权侵权风险。

借鉴遵循动态规制逻辑的“避风港原则”

在合理使用制度失灵的情形下，结合商业和技术环境的实际情况，为生成式人工智能平台设定适当的注意义务，不仅能够促进技术创新的政策纳入考量，而且能够为市场主体提供明确的合规方案。

在实践中，生成式人工智能的产业链主要分为三个层次，即基础模型层、专业模型层和服务应用层，其中服务应用层直接面向终端用户，遵循传统网络平台的运行逻辑，能从生成内容中获利并具有防范侵权的能力，要求其承担合理的注意义务具有正当性。在“奥特曼案”中，法院根据公法上的义务性规定，为生成式人工智能平台设定了相应的注意义务，包括设置举报机制、潜在风险的提示义务和显著标识义务，并要求其采取关键词过滤等技术措施来防止生成被控侵权内容。通过设定注意义务将人工智能企业合规成本控制在合理范围，有助于推动技术产业的创新发展。未来，随着技术进步和商业模式演变，注意义务将在个案中被适时调整，以平衡版权保护与技术发展的双重价值。

在人工智能时代，以“事前授权”为核心的著作权制度已难以满足海量作品使用的现实需求，这推动商业实践不断探索和创新，衍生出一系列灵活的自治手段，为生成式人工智能机器学习的版权治理提供新思路。

用户协议以合同的方式明确用户和服务提供者之间的权利和义务，在互联网环境中具有广泛的使用场景，比如在以“服务换数据”商业模式中，互联网企业以用户协议为中介，在为用户提供免费服务的同时获取用户生产的数据。再比如，OpenAI就在服务条款中推出了“版权保护盾”(Copyright Shield)，声明若用户因使用ChatGPT企业版和API而面临版权侵权诉讼时，OpenAI将介入帮助用户应诉并承担因诉讼产生的相关费用。此外，版权保险制度也能帮助人工智能企业有效应对内生性的版权侵权风险。从技术原理来说，基础模型的“记忆”现象是不可避免的，第三方保险机构的引入可以帮助企业分担潜在的版权侵权风险，有利于其提升抗风险能力和降低合规成本。同时，知识产权共享协议有助于构建“版权共享池”，支持中小企业共享更丰富的作品数据资源，推动人工智能产业创新发展。

因此，合理使用并非生成式人工智能机器学习绝对免责事由。避免生成版权侵权内容才是人工智能企业规避著作权侵权风险的核心策略。为提供更明确的合规指引并激励技术创新，应构建融合著作权制度工具与行业自治规则的版权治理框架，以实现版权保护与技术进步的协同发展。

声明：《版权监管周刊》部分插图源自网络，版权归作者所有。作者可通过左上角的联系方式与编辑部取得联系，编辑部将及时支付稿酬。

