

# ChatGPT等生成式人工智能使用作品的合理边界与侵权规制\*

顾男飞<sup>1</sup> 方舟之<sup>2</sup>

(1. 吉林大学法学院, 长春 130015; 2. 湘潭大学知识产权学院, 湘潭 411105)

**摘要:** 为高效释放作品的要素价值, 需厘清ChatGPT等生成式人工智能使用作品的边界, 以调和产业发展与著作权保护间的冲突。在具体论证中, 首先, 通过剖析ChatGPT出现智能跃迁的原因, 展示数据要素在生成式人工智能发展中的关键作用, 但也发现不同使用阶段的侵权风险制约作品获取。其次, 在作品合理使用层面, 基于产业政策目标, 解析相关群体的利益诉求, 提出“转换性使用”标准来划定生成式人工智能使用作品的合理边界。最后, 在侵权规制层面, 结合生成式人工智能在输入阶段和输出阶段的差异性使用目的和技术实现原理, 以“转换性使用”和“表达性使用”二分法认定著作权侵权风险; 同时, 明确作品使用的合规治理路径, 通过双向监管来高效规制风险。

**关键词:** 生成式人工智能; ChatGPT; 合理使用; 著作权侵权; 转换性使用

**中图分类号:** G923.41 **DOI:** 10.3772/j.issn.1673-2286.2023.07.001

**引文格式:** 顾男飞, 方舟之. ChatGPT等生成式人工智能使用作品的合理边界与侵权规制[J]. 数字图书馆论坛, 2023 (7) : 1-8.

以ChatGPT为代表的生成式人工智能一经推出便引发全世界激烈讨论<sup>[1]</sup>, 其强大的学习与拓展能力标志着应用于不同领域的通用人工智能成为可能。在可预见的未来, 通用人工智能将成为第四次工业革命的重要驱动力。2023年4月28日召开的中共中央政治局会议明确提出重视通用人工智能发展, 营造创新生态, 重视防范风险<sup>[2]</sup>, 从国家战略层面肯定发展生成式人工智能的必要性, 但也指出应当注意防范技术的伴生风险, 避免冲击社会秩序。与传统人工智能不同, 基于大模型技术的生成式人工智能不仅可以通过对数据的学习来提炼信息和预测趋势, 而且能生成不同于训练数据的新内容<sup>[3]</sup>。但生成式人工智能并不能凭空创造内容, 在训练过程中输入数据的质量和规模直接决定生成内容的优劣和应用场景的多寡。然而, 高质量数据往往是受保护的作品, 囿于较强的版权保护力度

和分散的著作权, 用户高效且合规地获取大批量作品面临巨大的掣肘。作品使用伴随着侵权风险, 因此在生成式人工智能产业发展中, 厘清合理使用边界以释放作品的要素价值成为关键, 并且应明确侵权认定标准, 避免过度侵蚀著作权人利益。简言之, 需有效回答生成式人工智能“为什么要用, 能不能用, 怎么用”作品这一关键问题, 但现有研究更多聚焦于著作权侵权认定, 对作品使用问题关注度较低, 论述并不充分。

## 1 生成式人工智能的作品使用需求与侵权风险

以ChatGPT为代表的生成式人工智能通过对超大规模作品进行深度学习, 有效地提升了生成内容质量,

收稿日期: 2023-05-24

\*本研究得到2021年国家社会科学基金重大项目“总体国家安全观下产业知识产权风险治理现代化研究”(编号: 21&ZD204)资助。

而且具备多轮对话能力,能结合用户提问内容和所提供材料进行有针对性的回复,具有十分优异的语言理解能力。ChatGPT出现智能跃迁的原因在于数据规模的扩大有效满足了生成式人工智能的发展需求。随着ChatGPT性能的飞跃,生成式人工智能所伴随的风险也被进一步放大。既有研究对ChatGPT进行了较多讨论,不仅涉及算法歧视和工作替代等社会风险,也涉及技术垄断等经济风险,可讨论更多聚焦于对著作权侵权等知识产权风险的规制。诚然,需重点关注生成式人工智能产业发展中的侵权风险,有效降低新兴技术对社会发展的负面冲击<sup>[4]</sup>,但既有研究在一定程度上忽略了人工智能的产业发展需要。在算力和算法水平无代差的国际竞争中,大规模作品的使用是生成式人工智能出现智能跃迁的基石,也是人工智能产业的发展基础。为弥合我国与其他国家间的产业代差,避免再遭遇类似芯片产业的困境,应有效释放作品的要素价值。

## 1.1 生成式人工智能的作品使用需求

数据之于生成式人工智能,就如字词之于一部名篇巨著,对于算法训练和性能提升不可或缺且影响巨大,其中优质数据更是不可替代的。人工智能的技术发展取决于三要素:数据、算法和算力。在世界范围内,算法大多已经公开,算力需求能通过购买显卡和云服务加以满足,制约产业发展的关键就是数据要素。对生成式人工智能而言,充分且优质的数据是算法进行训练的重要基础,特别是对具有一定独创性且属智力成果的作品,人工智能可通过学习大量作品以创造出更为优质的内容。比如,2017年微软小冰通过学习上百位诗人所创作的现代诗,而后通过10 000次以上的训练逐渐掌握创作现代诗歌的能力,出版了第一部完全由人工智能创作的诗集《阳光失了玻璃窗》<sup>[5]</sup>。就生成式人工智能训练过程而言,算法会预先对于既有文本语料进行无监督学习,内容包括文本语法、语义、惯用法和上下文信息等,再借助Transformer模型构建高度类似人类表达逻辑和方式的语言算法模型<sup>[6]</sup>。从训练过程可发现,数据规模和质量直接影响生成内容的质量,但中文领域的优质数据缺失现象严重。比如就可供直接爬取的网页数据而言,在规模排名世界前100万的互联网网站中,中文网站仅占1.3%,但是英语网站占比为59.3%<sup>[7]</sup>。如果能打破版权壁垒,充分利用我国大规模且高质量的作品,将在很大程度上解决我国优质训练数据不足的

问题,有效满足人工智能发展的数据需求。

## 1.2 作品使用不同阶段的侵权风险

生成式人工智能训练过程中对作品的不同使用方式会产生差异结果,不能一刀切地认定为侵权,需结合作品使用的不同阶段进行综合研判。虽然生成式人工智能具有一定的创造性,但是其创作过程高度依赖输入的数据,对于高质量作品的使用会更为频繁。也基于这一原因,著名语言学家Noam Chomsky直接称ChatGPT基本上是高科技的剽窃<sup>[8]</sup>,但这样简单的定性判断并不准确,特别是当前生成式人工智能使用的具体算法与模型较为复杂,用户无法知晓哪部作品以何种方式被使用,甚至开发人员也不能进行清晰说明。虽然上述“剽窃”论断并未得到证成,但就生成式人工智能在收集作品训练的输入阶段和内容生成的输出阶段而言,侵害著作权人署名权、修改权、保护作品完整权、复制权、改编权以及汇编权等风险依旧存在,且这种著作权侵权风险在某种程度上是不可避免的。

一方面,在输入阶段,生成式人工智能对需要使用的作品先进行复制,而后通过处理以获得“计算机可读”的标准格式,这属于我国《著作权法》中数字化复制作品的情形,将对复制权产生直接侵害。另一方面,在输出阶段,生成式人工智能会通过对于原作品的学习,结合使用者的需求输出对应内容,虽然不会直接输出原始作品,但也会综合诸多作品来进行再创作:可能是截取原作品片段,也可能是结合使用者需求对多份作品进行修改与重新加工,这时著作权人的修改权、汇编权以及改编权受到直接侵害。同时,由于生成式人工智能的输出内容并不会署上原作者姓名,也存在侵害著作权人署名权的风险。

## 2 作品使用的利益冲突与合理边界

作品作为重要的数据要素对于生成式人工智能的发展具有重要价值。考虑到社会利益以及产业政策目标,应允许作品的有序使用以满足人工智能产业的发展需求。但有序并非无条件,因此需解析生成式人工智能开发者与著作权人间的利益冲突,以弥合“促进创新和文化繁荣”与“保护作者著作权”的立法目标抵牾。在此基础上对合理使用制度加以改造,明确“转换性使用”标准,划定作品使用的合理边界,有效满足我国

人工智能产业的作品使用需求。

## 2.1 生成式人工智能产业政策的目标厘定

生成式人工智能不止于聊天,还将与工业4.0结合成为智能终端的大脑,共同助推我国产业升级。为此,应发挥我国新型举国体制的优势,以产业政策促进数据开放,其中当然也包括宝贵的作品。习近平总书记指出:“加快发展新一代人工智能是我们赢得全球科技竞争主动权的重要战略抓手,是推动我国科技跨越发展、产业优化升级、生产力整体跃升的重要战略资源。”<sup>[9]</sup>就生成式人工智能的数据需求而言,如果对用于人工智能训练的作品都需要获得许可,成本畸高,企业仅凭自身力量难以继,需要政府的积极介入。

作品同样属于数据要素,在生成式人工智能的发展中不可或缺,当然也应有有序开放。特别是在对开放数据运动的探索与反思基础上,包括我国在内的不少国家都提出要优先开放高价值数据集<sup>[10]</sup>。比如,2022年8月发布的《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》明确提出集聚人工智能场景数据资源,在确保数据安全的前提下提供数据开放服务。2023年2月发布的《浙江省人民政府办公厅关于培育发展未来产业的指导意见》更是明确指出应当提升数据要素对未来技术研发、应用场景创新的支撑作用。由此可见,为更好满足社会需求,我国的产业政策鼓励利用高质量数据,以作品的有序开放促进生成式人工智能产业的发展。

## 2.2 相关主体利益诉求的解析与平衡

作品的有序使用能高效促进中文领域生成式人工智能的发展,但在我国“保护优先”的著作权制度安排中,此举将严重侵害著作权人权益。为此,应结合著作权法立法目的和公共利益所需,解析相关主体的利益诉求并调和存在的利益冲突,实现产业发展与著作权保护的平衡。一方面,对于社会整体利益而言,作品之所以成为受法律保护的财产,是因为“促进创新和文化繁荣”的公共政策需要。在大数据以及算法的加持下,生成式人工智能的创作能力远远超过人类,其能提高创作效率,降低创作门槛,使更多民众参与创作,有效促进社会主义文化和科学事业的发展。但在训练过程中,作品作为优质数据,在相当程度上将决定生成式人

工智能性能和生成内容价值。缺乏足够的训练作品将会导致生成式人工智能生成内容的同质性和低劣化,并严重制约技术创新与文化繁荣。另一方面,对生成式人工智能开发者而言,成本过高的授权许可和法定许可制度并不适用于大批量的作品授权,而且有悖于市场公平竞争原则。同时,大企业能依托自身业务来收集既有作品,或者通过提供服务来换取作品授权,以规避著作权侵权风险,但是中小企业欠缺获取作品授权的资金,也无法以服务换取授权。由此,两者技术差距越来越大,市场竞争受限,甚至将倒逼中小企业通过侵权行为来获取训练所需的作品,致使双输。

著作权法不仅保护原创性内容的生成,也促进作品创作之后的使用,以平衡著作权人的专有权与社会公众使用和享受作品的利益<sup>[11]</sup>。考虑到社会利益和人工智能产业的发展需求,作品应有条件地开放给生成式人工智能开发者加以使用,著作权人利益也需得到充分保障,以实现利益的平衡。一方面,人工智能是战略性新兴产业,应基于我国的产业政策目标,营造有利于生成式人工智能产业发展的法治环境,允许作品的有条件开放。国外对人工智能训练规定了较为宽松的使用制度,比如2019年欧盟颁布的《数字单一市场版权指令》增加文本与数据挖掘例外条款<sup>[12]</sup>。企业为降低合规成本可能倾向于使用其他国家作品,我国大规模作品所具备的使用价值当然也就无法被充分利用,这进一步制约国内人工智能产业的发展。另一方面,在作品使用过程中应有效保障著作权人的合法权益。特别是要明确原作者的署名权,为后续修改权、保护作品完整权、复制权、改编权和汇编权等权利的行使,及财产利益的分配打下扎实基础。同时,通过明确署名能够有效规制部分开发者为降低侵权风险而隐匿作品来源的行为,避免对著作权人合法权益的进一步侵害。

## 2.3 作品使用的合理边界区分

为更好释放作品的数据要素价值,应当调整我国的合理使用制度来有效助推我国生成式人工智能的产业发展。如果说,在人工智能尚不发达的阶段,还有学者认为开放获取仅存在于法律制度之外的伦理层面,那么如今有必要重新考虑这一安排<sup>[13]</sup>,以进一步扭转“重保护轻共享”的传统立法范式。有学者指出,人工智能对数据的利用应该建立在合理使用的基础上,适用“合理学习”(Fair Learning)原则<sup>[14]</sup>,但合理使用制

度变革面临着较大阻碍：由于严重依赖司法裁量权，这一制度的适用结果往往是不可预测的<sup>[15]</sup>，模糊的标准亟待明确。尤其是在我国“作者中心主义”的制度安排中，生成式人工智能对作品的使用并不能完全适用于三步检验法，在实际认定中面临较大阻碍。一方面，我国《著作权法》第24条并未将文本与数据挖掘等人工智能发展所必需的情形纳入合理使用范畴。由于生成式人工智能需要大规模与高频率地利用作品，难以将其直接认定为“特殊情形”。另一方面，并非所有作品都有文学价值且需全文阅读。生成式人工智能通过对作品的整合与梳理，直接输出对应的结果，这可能降低使用者阅读原作品的可能性，与作品的正常使用抵触，无法完全满足“不得不合理地损害著作权人的合法权益”的要求。纵观当今各国立法和司法实践，证成合理使用的主要理由有以下3类：①非商业性使用；②少量使用；③公共利益需要<sup>[16]</sup>。然而，生成式人工智能并不能完全满足既有规定的要求。

为平衡产业发展和著作权保护的需求，可适用“转换性使用”标准来清晰划定作品使用的合理边界，救赎人工智能发展担负的侵权“原罪”，促进作品的有序、规范与大规模使用。“转换性使用”标准最先由美国法院Leval法官在著作权合理使用的判断标准中提出，在美国司法实践中被准许适用，其内涵也越发明确<sup>[17]</sup>。在“转换性使用”标准的使用中，关键在于判断新作品是否增加了不同特征的新内容，使用新的表达方式、意义或者信息。新作品越具有变动性，就越可能构成合理使用<sup>[18]</sup>。随后该标准被引入我国司法实践与学界讨论，为合理使用制度提供了重要参考，有效地弥补了《著作权法》的封闭性缺陷，比如北京、上海、广东等地人民法院审理的案件中，已有30多份裁判文书中出现“转换性使用”的内容<sup>[19]</sup>。但由于我国并没有明确规定该标准的适用要件，其司法适用并不统一，为此应当结合生成式人工智能的技术实现原理来明确“转换性使用”标准的内涵与外延，有效规制新技术背景下的著作权侵权风险。当前“转换性使用”标准主要可分为两类，分别是内容转换与目的转换<sup>[20]</sup>，生成式人工智能对作品的使用对二者均有涉及：对于前者，应当增加新内涵或改变原作品内容，特别是重视加入独创性表达或意义，而不是简单复制或重组文字，否则将陷入“表达性使用”的侵权范围；对于后者，由于对作品的使用目的在于训练通用模型，而不是简单复制并使用原作品，原作品也不会因此而丧失其目标受众。此外，偏向前者的“转换

性使用”能通过保护原作品的署名权来扩大原作品影响力，并能降低创作门槛来促进文化繁荣；偏向后者的“转换性使用”还能创造出新的市场与客户群体，利于激励市场竞争和技术创新。简言之，使用“转换性使用”标准能更好划定合理使用边界，促进作品的有序使用和高效流通。

### 3 作品使用的侵权认定与合规治理

合理使用并非无限制使用。为有效规制不同使用阶段存在的著作权侵权风险，应结合输入阶段和输出阶段的差异化使用目的和技术实现原理等，明确生成式人工智能对作品“转换性使用”和“表达性使用”下的著作权侵权标准。同时，为促进技术创新，应有效平衡产业发展需要和侵权风险。可参照竞争合规制度<sup>[21]</sup>，明确作品使用的著作权合规路径，通过生成式人工智能开发者的自我管理与风险防控，加强对侵权风险的治理。

#### 3.1 输入阶段“转换性使用”的侵权豁免

生成式人工智能在输入阶段的数字化复制行为并不属于简单复制，而是涵盖多重目标，不能被一刀切地认定为复制权侵权，若符合“转换性使用”标准则可豁免对应的侵权责任。已公示的《生成式人工智能服务管理办法（征求意见稿）》第7条规定：训练数据不仅需要满足多样性要求，也不能够含有侵犯知识产权的内容。但这对数据使用提出的较严苛监管要求在当前阶段是否妥当，还需要进一步的检视<sup>[22]</sup>。具体而言，从技术实现原理可知，生成式人工智能对作品的使用不是简单复制。虽然有学者指出ChatGPT不会直接复制文字内容，而是会重新进行文字组合、遣词造句，进而表达相同的观点<sup>[23]</sup>，但这并不完全符合技术实现原理。当前生成式人工智能的技术路线主要有两条，分别是以OpenAI研发的ChatGPT为代表的自回归预训练语言模型和以谷歌研发的Bert为代表的自编码语言模型，底层原理均为Transformer模型，该模型相较于传统的神经网络，能够提取出更高级、更抽象的语义信息，理解能力更强<sup>[24]</sup>。具体来看，前者根据上下文的内容预测当前词，学习训练文本中长距离词语间的依赖关系，语言生成能力更强；后者根据上下文的语义来预测中间词，能更好地理解两个句子间的关系，以适应问答以及推理等理解类的任务<sup>[25]</sup>。特别需注意到，作品作为先验知识

被整合到预训练模型中,用以提升模型性能。在美国司法实践中,是否构成“转换性使用”的关键在于使用目的是否与原作品的创作目的一致:当目的不一致时,可被认定为合理使用<sup>[26]</sup>。结合上述分析可知,生成式人工智能输入阶段的作品复制行为兼具多重目的,构成“转换性使用”时应被认定为合理使用并加以豁免。

结合差异化的使用目的,可调整现行合理使用制度中的适用情形来满足生成式人工智能输入阶段中的作品使用需要,但应遵守“转换性使用”标准,否则将被认定为著作权侵权。首先,明确“文本与数据挖掘”属于合理使用情形。文本数据存储、挖掘为后续的机器学习、内容生成提供了材料支撑,是人工智能时代科研和商业开发的重要基础,能够有效促进公共知识的积累,并增强对创新的激励,可被认定为合法<sup>[27]</sup>。应采取开放、包容和灵活的立法取向,除引入三步检测法之外,还应将“文本与数据挖掘”增列为一项新的合理使用情形,正面回应人工智能的作品利用问题<sup>[28]</sup>。新加坡政府也采取了这一例外情形,指出人工智能对作品的使用并不会影响作品的“表达性”,也就不会对著作权人的名誉和商业利益产生不利影响,为此可加以豁免,并支持开发者的数据利用<sup>[29]</sup>。其次,应允许企业出于商业性目的进行使用。公共利益在合理使用判断中应被纳入考量<sup>[30]</sup>,当商业性使用有利于公共利益时,应当予以肯定。我国的百度、华为、京东等15家商业公司正在研发新一代人工智能创新平台,可见制度调整有其现实依据。对生成式人工智能的发展而言,商业公司出于商业目的使用作品,能够促进我国新一代人工智能创新平台的建设,有利于公共利益的实现,应认定为合理使用。最后,为规制“与作品的正常使用相抵触”和“损害著作权人的合法权益”的侵权行为,应完善配套规则,要求企业在生成内容中明确标记原作品的权利主体,并可通过数字水印等技术永久嵌入标记,维护著作权人权益。当生成式人工智能输入阶段的复制行为满足上述要求时,可以认定为合理使用,反之则应被归入侵权情形加以法律规制。

### 3.2 输出阶段“表达性使用”的侵权判定

生成式人工智能对作品的侵权通过生成内容而外显,而基于使用目的和技术实现原理等因素,应重视对生成内容进行专业判定,当其满足“转换性使用”标准时应认定为不构成著作权侵权。当前学界讨论更多集中

于对复制作品的侵权判定,忽略了机器学习的核心并非“复制”,而是“学习”和“生成”,即能快速、大量地生成具备输入作品特征值的新内容。

一方面,当生成式人工智能可以满足内容转换要求时可构成“转换性使用”,由此被认定为合理使用。开发者对作品中的表达性内容并不感兴趣,重点关注的是作品蕴含的可用于模型训练的功能性内容,同时也不会向公众重新分发或者传播原作品<sup>[31]</sup>,因此输出的差异内容并不会侵害著作权。生成式人工智能基于大量优质且差异化的作品来训练出优秀的通用模型,且伴随训练数据、提问方式和反馈内容等因素的改变,其内容生成路径出现对应变化。特别是联网的大模型,比如微软联合OpenAI推出的NewBing,能结合最新网络数据和此前的用户交互历史来输出差异化的答案,使输出内容因人而异,以更好满足不同使用者的需求。输出阶段的直接复制即“表达性使用”行为正逐渐减少,这主要是因为生成式人工智能在输出时会进行过滤,避免输出内容与原作品一致,且随着训练数据集的扩大以及多元化,生成全新作品的可能性变大<sup>[32]</sup>,生成式人工智能将逐步实现“转换性使用”。但为避免侵权豁免的泛滥,应严格把握内容转换的认定标准,即输出阶段的生成内容应当增加新内涵,特别是应加入独创性的表述或意义,否则将因被认定为简单的“表达性使用”而被判定为著作权侵权,并受相应规制。

另一方面,“表达性使用”与“转换性使用”相对,当生成式人工智能输出阶段的生成内容仅为对作品的简单复制或重组时,应被判定为著作权侵权。具体而言,部分生成式人工智能在内容生成阶段直接复制或重组不同作品并输出,并未赋予原作品新的内涵,也并未产生新价值,而仅改编再现或简单复制原作品的独创性表述,在很大程度上构成对原作品的实质替代,也属于著作权侵权。为明确侵权边界,有学者提出了“逐字相似”标准来衡量生成内容是否是对原作品的简单复制<sup>[33]</sup>,然而这一标准过度追求形式化表征,可操作性较弱,在人工智能时代难以有效适用。美国司法实践中还存在“客观外在检验—主观内在检验”“普通观察者测试”“抽象—过滤—比较测试”等不同标准<sup>[34]</sup>,虽然彼此间并不统一,但“实质相似”标准基本得到了一致认可,在我国司法实践和学术界也得到了肯定。考虑到此前的作品输入,在侵权判定中可采取“接触+实质相似”的判定标准:如果人工智能输出的内容与原作品存在实质相似,而且事前未获得著作权人的许可,则可

判定构成著作权侵权<sup>[35]</sup>，生成式人工智能的开发者应承担相应的侵权责任。但由于算法黑箱的存在和数据集不公开，著作权人很少能举证证明生成式人工智能接触了原作品，因此可采取举证责任倒置的证明责任分配方式<sup>[36]</sup>，由开发者举证证明训练过程中没有接触原作品，并对算法、训练数据以及输出内容进行检查，从而有效解决黑箱问题，更好提升人工智能的可解释性<sup>[37]</sup>。此外，生成式人工智能只是工具，并不会主动复制并生成侵权内容，特别是有时输入作品来自使用者，且使用者给出错误指示，这时如果生成内容满足“实质相似”标准，依据使用者担责规则，使用者应当承担相应责任，以促进侵权责任的合理分配。

### 3.3 作品使用的合规治理路径

当前对新技术的风险治理往往存在扩大化倾向，但治理重点不能仅限于风险治理，还应确保满足“转换性使用”标准的作品复制和内容生成能依法进行，避免刑罚滥用以及行政监管过度，以促进生成式人工智能产业的高质量发展。从以往的经验可以看出，刑法在面对技术风险时，总是难以摆脱处罚的惯性<sup>[38]</sup>：以营利为目的未经著作权人许可而复制作品或者向公众传播人工智能生成内容，违法所得数额较大，将会被认定为侵犯著作权罪。若不顾技术发展需要而固守该罪的适用标准，将可能降低科研创新积极性<sup>[39]</sup>。特别是在世界各国抢占人工智能制高点的背景下，固守我国传统的合理使用分析逻辑，仍将机器学习置于著作权规制的枷锁中，会使我国丧失发展先机<sup>[40]</sup>。应当落实过罚相当原则，落实非必要不采用的要求，避免采用一刀切方式，形成差异化、灵活化的监管尺度<sup>[41]</sup>。为此可参照刑事合规和竞争合规实践，引入著作权合规理念，完善合规治理与责任豁免制度。该制度在我国已有一定的实践基础，比如2021—2022年，上海等地先后发布《企业知识产权合规标准指引（试行）》和《企业知识产权国际合规管理规范》等文件，虽然这些文件效力层次较低且缺乏足够约束力，但也作出了有益探索。

在选择合规治理路径时，应当注重侵权风险发现与合规管理体系建设，并将风险管控能力和合规体系建设成效作为减轻乃至豁免处罚的依据，避免由著作权侵权规制导致产品被禁用与巨额罚款等严重后果，倒逼企业规范使用作品。首先，合规要求应当完整且明确，需满足《个人信息保护法》《网络安全法》《互联网

信息服务深度合成管理规定》等法律法规的要求。比如，落实透明性要求，对训练作品的来源、规模、类型等进行完整描述，并对生成的内容进行标记，明确作品来源以保护原作品著作权人的署名权等。其次，监管部门和企业间应建立有效联系。监管部门可以通过约谈以及行政指导，对于作品使用风险以及侵权案件进行警示告诫，并提出明确的监管要求；企业应专门设置合规部门或专员，自觉完善企业内部的监督机制，通过发布《依法合规经营承诺》接受社会监督。同时，要注意保存各阶段的证据，保证合规监管实践有证可查、有迹可循。最后，应把合规举措与其积极成效纳入监管和执法的考量范围<sup>[42]</sup>。当生成式人工智能开发者因著作权侵权而被处罚时，监管部门可将该企业合规管理体系建设的成效纳入执法考量，灵活应用责任豁免制度。处罚并非目的，以监管促进著作权风险治理才是目标。

## 4 结语

与过去资本追逐的元宇宙等热门概念不同，以ChatGPT为代表的生成式人工智能正在逐步渗入社会发展的进程，并成为第四次科技革命的基石。在此背景下，为获取全球科技竞争主动权，亟需调整合理使用制度来释放作品的要素价值，满足我国人工智能产业发展的作品使用需求。通过解析并平衡相关主体的利益诉求，允许作品的有条件使用更符合社会利益与著作权法立法目标，但为保障著作权人利益，应适用“转换性使用”标准来清晰划定作品的合理使用边界，避免过度使用。同时，为规制生成式人工智能在使用作品的不同阶段中存在的著作权侵权风险，应结合差异化使用目的以及技术实现原理，采用“转换性使用”和“表达性使用”二分法，来清晰判定输入阶段和输出阶段的作品使用是否构成著作权侵权，并明确合规路径来强化风险治理力度。概言之，在系统分析并且回答生成式人工智能“为什么要用，能不能用，怎么用”作品这一关键问题的基础上，防范并规制侵权风险，以期推动我国人工智能产业的规范与高质量发展。

### 参考文献

- [1] 蒋雪颖, 刘欣. 生成式人工智能技术下的学术生产与出版: 变革、失范与路径[J]. 数字图书馆论坛, 2023, 19(5): 64-71.
- [2] 戴小河, 叶昊鸣, 高亢. 加快建设以实体经济为支撑的现代化产

- 业体系: 落实中央政治局会议精神做好当前经济工作系列述评之二[N]. 新华每日电讯, 2023-05-02 (1).
- [3] 陈永伟. 超越ChatGPT: 生成式AI的机遇、风险与挑战[J]. 山东大学学报(哲学社会科学版), 2023 (3): 127-143.
- [4] 刘艳红. 生成式人工智能的三大安全风险及法律规制: 以ChatGPT为例[J]. 东方法学, 2023, 92 (4): 30-43.
- [5] 张斯琦. 人工智能时代文学叙事功能与传播演变思[J]. 求是学刊, 2020, 47 (3): 148-156.
- [6] 邓建鹏, 朱恻成. ChatGPT模型的法律风险及应对之策[J]. 新疆师范大学学报(哲学社会科学版), 2023, 44 (5): 91-101, 2.
- [7] 简单说说中文互联网占世界互联网的比重! [EB/OL]. [2023-04-29]. <https://baijiahao.baidu.com/s?id=1723301712598596721&wfr=spider&for=pc>.
- [8] PERMALINK C M.Noam Chomsky on ChatGPT: it's "basically high-tech plagiarism" and "a way of avoiding learning" [EB/OL]. [2023-04-29]. <https://www.openculture.com/2023/02/noam-chomsky-on-chatgpt.html>.
- [9] 重视智能产业发展培育智能经济形态: 二〇二一智能经济高峰论坛发言摘编(一)[N]. 人民日报, 2021-07-30 (14).
- [10] 闫志开. 政府优先开放高价值数据集多维解析[J]. 数字图书馆论坛, 2023, 19 (4): 29-38.
- [11] BELL A, PARCHOMOVSKY G. Propertizing fair use[J]. *Virginia Law Review*, 2021, 107 (6): 1255-1304.
- [12] 阮开欣. 欧盟版权法下的文本与数据挖掘例外[J]. 图书馆论坛, 2019, 39 (12): 102-108.
- [13] 丛立先, 李泳霖. 生成式AI的作品认定与版权归属: 以ChatGPT的作品应用场景为例[J]. 山东大学学报(哲学社会科学版), 2023 (4): 171-181.
- [14] 於兴中, 郑戈, 丁晓东. 生成式人工智能与法律的六大议题: 以ChatGPT为例[J]. 中国法律评论, 2023 (2): 1-20.
- [15] PEDEN A. The good, the bad, and the ugly: incorporating the good or bad faith consideration into the fair use analysis[J]. *University of the Pacific Law Review*, 2023, 54 (1): 117-140.
- [16] 林秀芹. 人工智能时代著作权合理使用制度的重塑[J]. 法学研究, 2021, 43 (6): 170-185.
- [17] LEVAL P N. Toward a fair use standard[J]. *Harvard Law Review*, 1990, 103 (5): 1105-1136.
- [18] O' NEILL J. Lowering barriers to entry: YouTube, fair use, and the copyright claims board[J]. *Media & Entertainment Law Journal*, 2022, 33 (1): 176-220.
- [19] 韩伟. 数字图书馆建设中著作权侵权责任风险规避路径研究: 基于著作权案件司法审判引入转换性使用规则的视角[J]. 图书馆工作与研究, 2023 (3): 64-69.
- [20] 黄汇, 尹鹏旭. 作品转换性使用的规则重构及其适用逻辑[J]. 社会科学研究, 2021 (5): 95-104.
- [21] 王先林. 竞争法律与政策前沿问题研究[M]. 北京: 法律出版社, 2022: 106.
- [22] 徐伟. 论生成式人工智能服务提供者的法律地位及其责任: 以ChatGPT为例[J]. 法律科学(西北政法大学学报), 2023 (4): 69-80.
- [23] 王迁. ChatGPT生成的内容受著作权法保护吗?[J]. 探索与争鸣, 2023 (3): 17-20.
- [24] 张华平, 李林翰, 李春锦. ChatGPT中文性能测评与风险应对[J]. 数据分析与知识发现, 2023, 7 (3): 16-25.
- [25] 岳增营, 叶霞, 刘睿珩. 基于语言模型的预训练技术研究综述[J]. 中文信息学报, 2021, 35 (9): 15-29.
- [26] HUGHES J. The sub rosa rules of copyright fair use[J]. *Arizona Law Review*, 2022, 64 (1): 1-50.
- [27] GAO Y, KOSSOF P, DONG Y. Research on the dilemma and improvement of the copyright fair use doctrine related to machine learning in China[J]. *UIC Review of Intellectual Property Law*, 2022, 22 (1): 1-22.
- [28] 吴汉东. 人工智能生成作品的著作权法之问[J]. 中外法学, 2020, 32 (3): 653-673.
- [29] LIM D. Life after Google v. Oracle: three reflections on a theme[J]. *IP Theory*, 2022, 12 (1): 1-15.
- [30] MYERS G. Muddy Waters: fair use implications of Google LLC v. Oracle America, Inc.[J]. *Northwestern Journal of Technology and Intellectual Property*, 2022, 19 (2): 155-190.
- [31] QUANG J. Does training AI violate copyright law? [J]. *Berkeley Technology Law Journal*, 2021, 36 (4): 1407-1436.
- [32] LINDBERG V. Building and using generative models under US copyright law[J]. *Rutgers Business Law Review*, 2023, 18 (2): 1-65.
- [33] SUN H C. Redesigning copyright protection in the era of artificial intelligence[J]. *Iowa Law Review*, 2022, 107 (3): 1213-1252.
- [34] TU S S. Use of artificial intelligence to determine copyright liability for musical works[J]. *West Virginia Law Review*, 2021, 123 (3): 835-872.
- [35] 彭飞荣. 论算法创作中涉数据的著作权侵权风险及其化解[J]. 法律适用, 2023 (4): 46-55.
- [36] 顾男飞. 生成式人工智能的智能涌现、风险规制与产业调控[J]. 荆楚法学, 2023 (3): 70-83.

- [37] LIM D. AI, equity, and the IP gap[J]. SMU Law Review, 2022, 75 (4) : 815.
- [38] 孙禹. 论网络爬虫的刑事合规[J]. 法学杂志, 2022, 43 (1) : 162-172.
- [39] 赵小锁. 人工智能时代刑法适用问题前瞻[N]. 光明日报, 2022-11-28 (15).
- [40] 高佳佳. 类型化视角下机器学习的合理使用分析[J]. 电子知识
- 产权, 2021 (5) : 18-28.
- [41] 袁博. 大数据时代个人信息保护的行政监管立场及其智慧化转型[J]. 西南民族大学学报(人文社会科学版), 2022, 43 (6) : 96-107.
- [42] 周辉. 网盘平台创新服务的著作权保护从平台责任到合规治理[J]. 中外法学, 2023, 35 (2) : 365-382.

## 作者简介

顾男飞, 男, 硕士研究生, 研究方向: 经济法、社会法。

方舟之, 男, 博士, 副教授, 通信作者, 研究方向: 知识产权法, E-mail: fangzhouzhi\_law@126.com。

Reasonable Boundary and Infringement Regulation of the Use of Works by Generative Artificial Intelligence such as ChatGPT

GU NanFei<sup>1</sup> FANG ZhouZhi<sup>2</sup>

(1. School of Law, Jilin University, Changchun 130015, P. R. China; 2. School of Intellectual Property, Xiangtan University, Xiangtan 411105, P. R. China)

Abstract: To efficiently release the value of data elements of works, the boundary of the use of works by generative artificial intelligence (AI) such as ChatGPT needs to be clarified to reconcile the conflict between industrial development and copyright protection. In the specific arguments, firstly, a dissection of the reasons for the intelligent leap in ChatGPT shows the key role of data elements in the development of generative AI, and it reveals that there are infringement risks constraining access to works at different stages. Secondly, at the level of reasonable use of works, the objectives of industrial policies are determined, and the interests of relevant groups are analyzed. In addition, the standard of “transformative use” is proposed, so as to define the reasonable boundary of the use of works by generative AI. Finally, at the level of infringement regulation, the risk of copyright infringement is identified through the dichotomy of “transformative use” and “expressive use” by considering the different purposes of use and technical realization of generative AI in the input and output stages. Meantime, this study clarifies the compliance governance path for the use of works and effectively regulates the risks through two-way supervision.

Keywords: Generative Artificial Intelligence; ChatGPT; Reasonable Use; Copyright Infringement; Transformative Use

(责任编辑: 王玮)