

人工智能文生视频大模型的作品风险、 著作权归属及有效治理

丛立先 李泳霖

（华东政法大学知识产权学院，上海 201620）

摘要：以 Sora 为代表的文生视频大模型的出现，赋予作品风险、著作权归属和著作权风险治理新的讨论意蕴。笔者认为文生视频大模型作为人类精神产品创作的工具，其生成的内容构成作品。基于文生视频大模型的技术特点，其引发的作品风险不能一概而论，应基于对作品的不同使用行为加以识别和判断。在著作权归属领域，应坚持视听作品的私权属性，在秉持以意思自治约定归属优先的同时，在个案中结合实质贡献、投资激励和利益平衡原则确定著作权归属。在著作权风险治理领域，应通过著作权规则完善、著作权审判指导、著作权执法监管和大模型行业自治，共同作用于文生视频大模型的合规、高效与可持续发展。

关键词：Sora；人工智能；文生视频；作品风险；著作权归属；风险治理

中图分类号：TP18；D9 **文献标识码：**A **文章编号：**1005-9245（2024）06-0101-11

从 ChatGPT 的文生文，到 Midjourney 的文生图，再到 Runway、Pika、Sora 的文生视频，人工智能的工具能力被不断延展。相较此前出现的人工智能，以 Sora 为代表的文生视频大模型具有的可视性、高仿真性和时空延展性，使其生成内容不断接近人类创作水平，提高了人类对其创作的依赖程度，甚至超越了人类对既有视听作品的认知。Sora 之所以具备上述能力，一方面得益于对既有作品的深度学习，另一方面依赖于算力和算法的加持。可以预见，基于出色的视频生成能力，以 Sora 为代表的文生视频大模型将被应用于更多下游产业。为进一步提升文生视频大模型生成内容的著作权价值，解决大模型在下游产业中可能出现的著作权归属纠纷，有必要在认清著作权风险的基础上，厘清其生成作品的著作权归属，并通过著作权规则完善、著作权司法定性、著作权监管介入和大模型行

业自治，追求大模型发展的良法善用和算法善治，进而实现“优质作品学习—高质量视频生成—弱风险使用”的良性循环。

一、人工智能文生视频大模型的著作权风险

与人类通过“数据—信息—知识—作品”这一传统路径获取知识并创作作品相似，文生视频大模型也离不开对作品的使用。基于作品使用行为的差异，不同文生视频大模型面临差异化的著作权风险。

（一）文生视频大模型学习过程中的著作权风险

通常情况下，未经许可使用他人享有著作权的作品，在不构成合理使用或法定许可等法定免责事由的情形下，可能构成著作权侵权行为。总体而

收稿日期：2024-03-25

基金项目：本文系研究阐释党的二十大精神国家社科基金重大项目“全球知识产权治理面临的主要挑战和中国方案研究”（23ZDA124）的阶段性成果。

作者简介：丛立先，华东政法大学知识产权学院教授、博士生导师；李泳霖，华东政法大学知识产权学院博士研究生。

言，大部分文生视频大模型在训练过程中都会不可避免地使用他人享有著作权的作品。例如，Sora在训练模型过程中会将原始视频输入降低视觉维度的网络，并压缩到低维潜在空间^①。腾讯科技（深圳）有限公司在申请的名为“视频生成方法、装置、电子设备及存储介质”的专利中显示，要将筛选的图片素材、文本改写和提炼的文本素材以及文本匹配音乐素材输入视频合成引擎^②。度小满科技（北京）有限公司申请的名为“文生视频生成方法、装置、电子设备及可读存储介质”的方法专利，同样需要输入与待生成视频对应的文本信息^③。由此可知，上述模型生成视频的技术环节均包括输入原始文本、图片或视频等作品，该行为实际上在大模型深度学习和使用作品前，已将在先作品通过数字化方式制作一份或多份，可能构成《中华人民共和国著作权法》（以下简称《著作权法》）意义上的复制行为。在此基础上，不同模型对作品的后续使用方式存在差别。例如，Sora从压缩输入视频中提取一系列时空补丁作为转换标记。由于这一行为仅提取原作品中的信息，并未直接使用原作品，因此，既不构成著作权法意义上的作品使用，也不会造成著作权风险。但Sora模型除可根据纯文本生成视频外，还有三项功能，即文字与图片组合生成视频，拓展、延长已有视频以及视频融合。上述行为本质上不改变作品内容，只将图片作品转变为视听作品；或虽未改变视听作品这一类型，但针对在先视听作品进行改写或扩展，此时只要创作出具有独创性的新作品，就构成对原作品的改编。与此同时，如果将输入的图片、音乐、文本等素材加以非线性编辑，并通过二次编码生成视频^④，或将多个

主体分别对应的主体文本表述以及主体图像混合为文本与组合图像^⑤，同样可能构成改编行为，等等。

在既有《著作权法》规则下，由于机器并非自然人，且机器投资者通常不属于国家设立的教育、科研公共事业单位（例如，经相关主管部门、教育部批准设立的学校、全额财政拨款的国家科研机构等）^⑥，因此，机器学习一般不构成“个人学习、研究”或“科学研究”^⑦，加之机器学习并未包含在有限的法定许可范围内，难以满足合理使用或法定许可的要求^⑧，故而可能被认定为著作权侵权行为。具言之，大模型投资者将面临两种选择。一种是致力于实现使用在先作品的著作权合法性；另一种是仅使用不构成作品的数据或放弃对大模型领域的深耕，意味着在当前数字环境下难以适用先授权后使用的常规模式。在人工智能引发全球范围内的社会生产力变革和全球价值链重构的背景下^⑨，无论缩小数据使用范围还是放弃大模型投资，都可能导致我国失去人工智能技术发展的历史机遇，甚至受到发达国家人工智能产业的不利冲击。笔者围绕文生视频大模型学习过程中的著作权风险进行制度设计，旨在弱化文生视频大模型著作权规则障碍，对于提升我国人工智能产业在全球价值链中的竞争力具有重要意义。

（二）文生视频大模型生成内容的作品认定

工具指生产劳动中使用的器具，或用以达到某种目的的事物^⑩。技术的发展过程是把人的工作逐渐外包给机器的过程。计算机技术的出现和应用将人类社会推向数字经济时代，生产工具也由此进入智能化时代^⑪。美国版权作品新技术应用国家委员会将计算机同照相机、打印机以及其他创作工具进行

① Video generation models as world simulators, <https://openai.com/research/video-generation-models-as-world-simulators#fn-20>.

②④《视频生成方法、装置、电子设备及存储介质》，公开号CN117453952A, <https://pss-system.cponline.cnipa.gov.cn/retrieveList?prevPageTit=changgui>。

③《文生视频生成方法、装置、电子设备及可读存储介质》，公开号CN117061763A, <https://pss-system.cponline.cnipa.gov.cn/documents/detail?prevPageTit=changgui>。

⑤《一种定制化多主体文生视频方法、装置、设备及介质》，公开号CN117496020A, <https://pss-system.cponline.cnipa.gov.cn/documents/detail?prevPageTit=changgui>。

⑥ 李明德、管育鹰、唐广良：《著作权法专家建议稿说明》，北京：法律出版社，2012年版，第252页。

⑦ 焦和平：《人工智能创作中数据获取与利用的著作权风险及化解路径》，《当代法学》，2022年第4期。

⑧ 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》，《山东大学学报（哲学社会科学版）》，2023年第4期。

⑨ 何宇、陈珍珍、张建华：《人工智能技术应用与全球价值链竞争》，《中国工业经济》，2021年第10期。

⑩《现代汉语词典》，成都：四川辞书出版社，2017年版，第281页。

⑪ 赵宸宇、王文春、李雪松：《数字化转型如何影响企业全要素生产率》，《财贸经济》，2021年第7期。

对比,认为照片的作者是使用照相机的人,计算机“创作作品”的作者就是使用计算机的人^①。计算机“创作作品”指利用信息技术在计算机上创作并在显示屏上显示,或以其他拷贝形式存在的作品^②,其中,计算机程序是被动性协助创作工具^③。文生视频大模型作为人工智能的最新成果,能够在人类的指令下辅助人类高效率生成视频,属于人类创作作品的工具,只是相较石块和工业社会的机器更具智慧属性^④。在“春风图案”^⑤中,法院明确提出“人们利用人工智能模型生成图片时,……本质上,仍然是人利用工具进行创作”^⑥。尽管本案中人工智能生成的内容是图片,但视频正是由一帧一帧连续的图片组成,因此,“春风图案”的裁判结论对文生视频大模型的著作权归属具有重要的参考价值。

作为人类创作视听作品的工具,文生视频大模型生成的视听作品融入人类的智力活动,属于人类的智力成果。以 Sora 为例,当前, Sora 官网上展示的提示词包括两类,一类提示词较为详细,例如,“镜头跟在一辆带着黑色车顶架的白色复古 SUV 后面,车辆在陡峭的山坡上沿着松树环绕的陡峭土路加速行驶,灰尘从轮胎上扬起,被阳光照射的越野车在土路上加速行驶,在现场投下温暖的光芒。这条土路弯弯曲曲地延伸到远处,看不到其他的车辆。道路两旁的树木都是红杉,点缀着一片片绿色植物。从后面可以看到这辆车轻松地沿着弯道行驶,看起来就像是在崎岖的地形上行驶。土路本身被陡峭的山丘和山脉包围,上面是晴朗的蓝天和稀疏的云层”。这一段提示词对应的视频通过大模型语义分析后,根据提示词中包含的独创性内容生成,是使用者智力投入的成果。另一类提示词的内容较简短,例如,“淘金热时期的加州

历史镜头”。根据该提示词生成的画面中,出现了小溪、人、马、建筑、远山、杂草以及形态各异的村落等,而提示词并未直接给出这些元素。原因在于 Sora 先利用 GPT 将简短的用户提示转换为较长的详细字幕,之后发送至视频模型,进而视频模型按照较长的详细字幕生成高质量视频^⑦。至少从 GPT-3 开始,训练者就使用基于人类反馈强化学习(RLHF)的训练^⑧,在训练过程中,GPT 为每个 Prompt(提示)生成多个输出,标注者将其从最佳到最差进行排序,进而建立新的标签数据集。这些新数据用于训练奖励模型,并根据该模型调整输出结果。这意味着在被延展的文本中融入了人类的主观偏好,按照具备人类偏好的提示词生成的视频,属于人类的智力成果。有观点认为,人工智能完成特定任务的使用者不能凭借其自由意志决定人工智能生成的内容^⑨,并提出创作行为必须对应唯一结果,但实验表明,同样的提示词和参数生成的结果有所不同,因此,该观点认为使用者的行为不是创作,生成内容也不是使用者智力投入的结果。笔者认为,创作行为并非必须对应唯一结果,针对同一事物,站在不同角度理解会得出不同结论。此外,使用者在使用文生视频大模型时,可通过指令不断调试结果,进而使大模型输出使用者预期的内容,即文生视频内容仍可凭借人类的自由意志决定。

文生视频也存在著作权保护例外,即如果文生视频内容是单纯的事实消息、历法、通用数表、通用表格和公式以及具有唯一性的独创性思想表达,将无法成为著作权法保护的作品。单纯事实消息指对时间、地点、人物、起因、经过、结果等客观事实的单纯叙述,不包括作者的主观感受、思想情感或修辞、评论^⑩。《著作权法》不保护纯粹

①③ National Commission on New Technology Uses of Copyrighted Works, Final Report on New Technology Uses of Copyrighted Works, 1979: 43-45, 44.

② Jonathan C. Jackson. Legal Aspects of Computer Art, Rutgers Computer & Technology Law Journal, 1993(19): 495-516.

④ 袁曾:《人工智能有限法律人格审视》,《东方法学》,2017年第5期。

⑤ “春风图案”即李某诉刘某侵害作品署名权、信息网络传播权纠纷案,该案判决第一次确认了作者利用人工智能生成的图片享有著作权。

⑥ 北京互联网法院(2023)京0491民初11279号民事判决书。

⑦ Video generation models as world simulators, <https://openai.com/research/video-generation-models-as-world-simulators>.

⑧ Long Ouyang, et al. Training language models to follow instructions with human feedback, <https://arxiv.org/pdf/2203.02155v1.pdf>.

⑨ 王迁:《再论人工智能生成的内容在著作权法中的定性》,《政法论坛》,2023年第4期。

⑩ 蒋强:《著作权侵权案件中时事新闻的认定:新闻报道著作权侵权纠纷案评析》,《科技与法律》,2011年第3期。

消息或繁杂事实，也不保护对上述消息或事实的单纯报道，因为这些材料不具备可以被称为作品的必要条件。当大模型生成的视频是根据单纯事实消息创作的新闻作品时，应受《著作权法》保护。此外，思想的唯一表达即使构成视听作品也不应给予著作权保护，例如，如果不同的人使用 Sora 预测同一线虫未来 10 秒的运动轨迹将产生相同的视频，则不应将这一视频的著作权归属于某一特定主体。

（三）视频生成后的著作权风险

任何作品的价值都在于有效利用^①，倘若未经许可使用他人享有著作权的文生视频，就会产生著作权风险。首先，生成的视听作品可能基于被直接使用行为面临著作权风险。例如，在“菲林诉百度案”^②“腾讯诉盈讯案”^③以及“奥特曼案”^④中，被告均未经许可使用原告享有著作权的文字作品或美术作品，从而被认定为侵权。尽管涉案作品并非视听作品，但本质上都属于在未经许可的情况下使用他人享有著作权的作品，对降低和应对视频生成后的著作权风险具有重要价值。又如，Sora 网站公布了提示词以及生成的视频，如果自媒体创作人未经著作权人许可，将英文提示词译成中文，并将其与生成的视频上传至第三方平台，此时可能侵犯文字作品的改编权以及视听作品的信息网络传播权等权利。再如，主播通过直播的方式展示其使用大模型生成的视频，此时虽然社会公众无法在选定的时间和地点获得文生视频，但主播的行为属于以无线方式公开传播视听作品，可能侵犯相关视听作品的广播权。其次，生成的视听作品可能基于切条、图解、混剪等方式被使用，此时是否存在著作权风险不能一概而论。如果将大模型生成的电影作品切条传播，或截取电影作品中的图片并播放，可能侵犯该电影作品的复制权、信息网络传播权等权利。例如，由 50 位艺术家完全使用文生视频大模型生成的电影《Our T2 Remark》，就可能面临这一侵权风险。但如果通过混剪、解说等方式使用文生视频，可能构成合理使用，不存在著作权侵权风险。在合理使用的判

断上，如果该行为在满足合理使用一般条款，即三步检验法的情形下，同时符合法定可以不经著作权人许可、不向其支付报酬使用作品的十二种具体行为^⑤，将毫无疑问构成合理使用。值得注意的是，即使不属于上述法定的十二种具体情形，也可根据合理使用一般条款与兜底条款进行判断。在兜底条款的解释上，可适用中华人民共和国最高人民法院于 2011 年出台的《关于充分发挥知识产权审判职能作用 推动社会主义文化大发展大繁荣和促进经济自主协调发展若干问题的意见》（以下简称《意见》）第 8 条的规定，“在促进技术创新和商业发展确有必要的情形下，考虑作品使用行为的性质和目的、被使用作品的性质、被使用部分的数量和质量、使用对作品潜在市场或价值的影响等因素，如果该使用行为既不与作品的正常使用相冲突，也不至于不合理地损害作者的正当利益，可以认定为合理使用”。此外，如果文生视频的使用者利用大模型生成违禁作品，该违禁作品的著作权行使应受到严格限制。

除文生视频的使用者可能面临著作权风险外，基于大模型生成的视听作品大多通过网络平台传播，此时，网络平台在不同情形下面临不同的著作权风险。第一种情况是网络平台直接提供侵权的视听作品，可能面临直接侵权的著作权风险。但网络平台更多情形下属于第二种或第三种情况，即帮助、诱导用户提供侵权视听作品，或以网络服务提供者的身份提供合法的自动接入、存储、搜索链接等技术服务。对于前者而言，网络平台可能面临间接侵权的著作权风险；对于后者而言，应为其提供避风港规则予以免责。如果网络平台既提供侵权的视听作品，又提供技术服务，则不享受避风港规则使其免责。

二、人工智能文生视频大模型生成作品的著作权归属

作品著作权的归属及行使规则通常取决于作品性质^⑥。文生视频在构成视听作品的情况下，作为

① 丛立先：《论短视频作品的权属与利用》，《出版发行研究》，2019 年第 4 期。

② 北京互联网法院（2018）京 0491 民初 239 号民事判决书。

③ 广东省深圳市南山区人民法院（2019）粤 0305 民初 14010 号民事判决书。

④ 广州互联网法院（2024）粤 0192 民初 113 号民事判决书。

⑤ 《中华人民共和国著作权法》第 24 条。

⑥ 卢海君、申耘宇：《电影作品的属性研究——以我国〈著作权法〉第 15 条修改为背景》，《重庆理工大学学报（社会科学）》，2011 年第 3 期。

私权的著作权归属应坚持意思自治的基本原则^①,在此基础上,衡量不同视频生成中的智力投入,兼顾著作权归属结果对投资者利益、社会利益和国家利益的影响。

(一) 文生视频著作权归属的复杂性

人工智能生成作品的著作权归属不仅是保护生成作品的制度起点,而且不同的著作权归属会影响其生成作品的利益分配^②。作为人工智能的最新产品,文生视频著作权的可能归属主体至少包括大模型的使用者、大模型的训练者和大模型的投资者。其中,大模型的使用者和训练者对视频的生成可能作出实质贡献,大模型生成视频以人类输入的指令为依据,即使用者输入的指令越详细具体,视频内容与文本内容的匹配程度就越高。当前的文生视频大模型尚不具备脱离人类指令自动生成视频的功能。美国联邦最高法院在 *Burrow—Giles Lithographic Co.v.Sarony* 案中提出“如果不是”规则^③,在文生视频大模型应用过程中,如果没有使用者发出的指令,大模型就无法生成视频。此外,如前文所述,在用户输入的提示词过短、视频由被延展后的提示词生成的情况下,训练者同样对文生视频作出了实质贡献。因此,使用者和训练者应居于文生视频著作权归属的重要地位。

作为创作工具,大模型之所以能够根据使用者的提示词生成视频,是因为数据、算力和算法的加持。一方面,使用者受技术能力限制,不仅难以实现海量数据的收集和挖掘,而且难以承担创作作品需要的巨额资金和投资风险^④,只能依赖大模型投资者的资金保障和组织;另一方面,文生视频大模型的开发是整个团队运用各种先进设备和技术集体劳动的成果,从结果看,如果不考量投资者的利益回报,就不会有更多的资本投入^⑤。因此,尽管

《著作权法》授予的专有权并非建立在作品的商业化基础之上^⑥,但大模型的出现强化了投资或资本在文生视频领域的地位,影响视听作品权利归属的制度构造。有学者认为,保护作者和激励创作只是支撑著作权制度正当性的符号性表达,《著作权法》的真正目的是保护作品的投资者^⑦,但投资者在文生视频中发挥的作用不能一概而论。视听作品中的电影作品、电视剧作品的著作权之所以由制作者享有,是因为制作者要确保投资者对所投资电影或电视剧作品的预期收益,进而产生集中权利。大模型投资者的投资对象是大模型本身,而非特定的视听作品。换言之,某些视频的生成可能需要耗费较大算力,但其也可能仅需小部分算力即可生成,甚至使用者无需借助大模型也可获得相同或类似的视频。因此,对投资者利益的考量应在个案中结合生成视频本身确定,同时,兼顾使用者、训练者以及投资者的利益平衡。利益平衡指在各方面因素发生冲突时,通过法律进行调节,使相关方利益能够共存和相容,并在此基础上达到优化、合理状态^⑧。著作权法作为一种平衡的制度设计^⑨,不仅要在宏观层面充分考量文生视频著作权归属对整体作品市场的影响,而且要从微观层面分析各主体对创作物的整体贡献,以合理的制度设计达到利益平衡的效果^⑩。1976年,美国在完善雇佣作品条款时,创作者一方提出,创作者相较于投资者处于弱势地位,投资者获得初始权利将使个体创作者处于不利地位^⑪。因此,文生视频著作权归属的理想目标是投资者因投资激励推动大模型迭代优化,进而提升生成视频的质量,使用者因个人利益的满足而使用大模型创作更多作品。

(二) 意思自治在文生视频著作权归属中的应用 强化以意思自治为核心的私法系统,是政府在

① 丛立先、李泳霖:《聊天机器人生成内容的版权风险及其治理——以ChatGPT的应用场景为视角》,《中国出版》,2023年第5期。

② C.R.Davies.An evolutionary step in intellectual property rights-artificial intelligence and intellectual property, *Computer Law &Security Review*, 2011(27): 610.

③ *Burrow—Giles Lithographic Co.v.Sarony*, 111 U.S.53(1884).

④ 梁志文、李忠诚:《论算法创作》,《华东政法大学学报》,2019年第6期。

⑤ 李芳芳:《人工智能创作物的著作权保护研究》,《出版广角》,2018年第9期。

⑥ *Stewart v. Abend*, 495 U.S.207(1990).

⑦⑩ 熊琦:《著作权法中投资者视为作者的制度安排》,《法学》,2010年第9期。

⑧ 冯晓青:《知识产权法利益平衡理论》,北京:中国政法大学出版社,2006年版,第11-12页。

⑨ Richard Stallman. Reevaluating Copyright: The Public Must Prevail, *Oregon Law Review*, 1996(75): 291-298.

⑪ 陈全真:《人工智能创作物的著作权归属:投资者对创作者的超越》,《哈尔滨工业大学学报(社会科学版)》,2019年第6期。

经济政策和法律制度中的必然选择^①，尤其是知识产权的商法特质，决定了知识产权合同的意思自治将成为主流^②。在市场经济背景下，文生视频大模型生成的视听作品将迅速进入交易市场，此时投资者与使用者自由处分自身权利的规定不仅必要，而且处于关键地位，知识产权的转让、许可等合同也需更多的意思自治^③。意思自治即私法自治，指私法主体有权按照个人意志实施私法行为，他人不得干预；私法主体仅对基于自由表达的真实意思实施的私法行为负责；在不违反法律规定的前提下，私法主体自愿达成的协议优先于私法适用^④。视听作品享有的著作权为私权，对著作权的约定应是文生视频著作权归属的首要依据。例如，2023年11月14日，OpenAI在更新的使用条款中规定，用户拥有输入内容的所有权以及产出内容的所有权利及利益^⑤。Sora作为OpenAI开发的人工智能产品之一，上述约定应适用于其生成视频的著作权。通过诸如此类的约定归属，不仅有利于大模型投资者与用户形成权利义务预期，维护法律关系的稳定性，而且有利于节约交易成本和解决争议。

由于文生视频大模型的投资者与多数使用者在研发、使用和转化能力方面力量对比悬殊，致使并非所有著作权归属都具有法律效力。意思自治的前提是意志自由，只有当事人处于平等地位，契约的规定才不会发生倾斜，契约的内容才是当事人的合意^⑥。形式平等无法确保实质平等，有效的意思自治不仅要求当事人之间满足形式平等，而且应通过实质平等实现形式平等的具体化和实质化^⑦。这要求著作权归属的约定应实现著作权利益的合理分配以及司法裁判的公正。从大模型投资者与使用者的地位看，投资者掌握资金、资源的绝对优势，容易形成投资者相较于使用者更具

有优势的局面。尤其当使用者在输入的提示信息上作出较大智力投入的情况下，约定由投资者享有著作权仅具有形式意义，这一分配结果背离了意思自治的实质价值，甚至可能导致著作权走向垄断，最终妨碍竞争、阻碍科技进步与创新。例如，2024年1月12日，文生视频大模型Pika更新的使用条款规定，Mellis及其供应商或许可人拥有艺术品、动画、音乐作品、视听效果等内容的所有权利、所有权和权益^⑧，这一约定忽视了使用者的贡献，应得到必要的审查与矫正。此外，在追求使用者、训练者与投资者之间著作权归属的实质平等时，应注意两个问题：一是兼顾社会公共利益和国家利益，防止片面地以效益替代公平的价值取向；二是协议内容不应属于约定无效等情形，也不能与法律的强制性规定相冲突^⑨。

（三）无约定情形下文生视频的著作权归属

在无约定情形下，著作权归属应包括一般归属和特殊归属。在一般归属情形下，视听作品的著作权可能归自然人、法人或非法人组织享有。自然人包括大模型的使用者和训练者。文生视频的著作权归属应以使用者和训练者是否从事实质性智力投入为判断标准。对使用者输入简单指令生成的视频，例如，“飞越博物馆之旅，那里有许多绘画和雕塑以及各种风格的精美艺术品”，使用者在该视频中既没有贡献艺术性的实质技能或劳动，也没有为生成作品的最终产生作必要的安排，因此，其无法被视为著作权法意义上的作者^⑩。针对用户输入的简短指令，Sora可延展用户的指令，并按照延展后的指令生成视频。此时，视频中的独创性表达与使用者输入的指令关联度较弱，视频的内容主要根据被延展的提示词生成。在英国的Nova Productions Ltd v. Mazooma Games Ltd案^⑪以及美国的Williams

①④ 江平、张礼洪：《市场经济和意思自治》，《法学研究》，1993年第6期。

② 熊琦：《知识产权法与民法的体系定位》，《武汉大学学报（哲学社会科学版）》，2019年第2期。

③ R.C.Rerben.What's new in intellectual property:business is booming in copyright,trademark and patent law,ABA Journal, 1993(79):72-77.

⑤ Term of Use, <https://openai.com/policies/terms-of-use>.

⑥ 刘凯湘、张云平：《意思自治原则的变迁及其经济分析》，《中外法学》，1997年第4期。

⑦ 王元亮：《论形式平等与实质平等》，《科学社会主义》，2013年第2期。

⑧ TERMS OF SERVICE, <https://pika.art/terms-of-service>.

⑨ 丛立先：《人工智能生成内容的可版权性与版权归属》，《中国出版》，2019年第1期。

⑩ Jani Mc Cutcheon. The Vanishing Author in Computer-Generated Work: A Critical Analysis of Recent Australian Case Law, Melbourne University Law Review, 2013(36):915-969.

⑪ Nova Productions Ltd v. Mazooma Games Ltd, European Law Reports, 2006(10):1255-1260.

Electronics Inc.v.Arctic International Inc 案^①中，电脑游戏画面的著作权均归属于游戏软件的编程设计者。因此，对于使用者向大模型发出简单指令生成的视频而言，其训练者与电脑游戏画面的编程设计者相似，都付出了本质上相同或类似的智力劳动，其著作权应归属于训练者所有。当使用者与训练者都投入一定的智力活动时，可能出现使用者与训练者创作合作作品的情况。此外，如果生成的视频代表法人或非法人组织的意志，且满足法人作品的其他构成要件，那么，生成视频的著作权可能归相关法人或非法人组织所有。

文生视频作为视听作品具有多重属性，是一种多面向的作品形态^②，在构成演绎作品、合作作品、汇编作品、视听作品、职务作品和委托作品的情形下，应按照相应的著作权归属规则进行判断。如果使用者输入的文本或图片本身构成小说、戏剧等作品，即文生视频是从在先作品基础上改编而来，其生成的视频可能构成演绎作品。如果生成的视频对在先作品的编排具有独创性，则可能构成汇编作品。演绎作品和汇编作品的著作权归属仍遵循一般归属规则，即根据训练者与使用者的智力投入进行判断。对于视听作品中的电影作品和电视剧作品，由于需要占用较多算力、数据和更多投资，其著作权归属于大模型的投资者较为合理。以 Pika 为代表的大模型可生成带有音乐的视频，此时，音乐中的词作者和曲作者享有署名权，一旦其中的音乐可以单独使用，其作者就有权单独行使著作权。其他视听作品的著作权归属由当事人约定，没有约定或约定不明确的情况下，遵循一般归属规则。如果使用者使用大模型生成视频的目的在于完成法人或非法人组织的工作任务，其生成的视频可能构成职务作品，此时应区分两种情况：一是如果作者属于报社、期刊社、通讯社、广播电台、电视台的工作人员，此时法人或非法人组织享有除署名权之外的其他权利。如果属于一般职务作品，则按照一般归属规则分配权属，但法人或非法人组织有权在其业务范围内优先使用。此外，在作品完成两年内，如果

未经单位同意，作者不得许可第三人以与单位使用的相同方式使用该作品。二是假如使用者受他人委托使用大模型生成视频，在使用者付出实质性智力投入的情况下，可按照委托人和使用者的约定确定著作权归属。如果未曾约定或约定不明确，著作权归使用者所有。如果使用者没有付出实质性智力投入，则著作权归训练者所有。较为特殊的是合作作品。合作作品是两人或两人以上共同创作完成的作品，合作作品的作者必须是参加创作的人。“参加创作”指对作品的思想观点、表达形式付出创造性的智力劳动，或构思策划、或执笔操作，一般认为创作合作作品需要合作作者之间的合意^③。笔者认为，大模型的训练者希望使用者出于创作目的使用模型，模型的使用者也希望通过训练者开发的模型进行创作^④。从结果看，如果生成的视频在展现使用者实质性智力劳动的同时，在使用者智力投入的基础上进行相匹配的延展，那么，相应的视频构成合作作品，使用者与训练者构成该视频的合作作者。如果该视频可以分割使用，使用者就提示词对应的视频享有著作权，训练者就提示词延展部分对应的视频享有著作权，但各著作权人在行使著作权过程中不得侵犯视频整体的著作权。

三、人工智能文生视频大模型著作权风险的有效治理

人工智能文生视频大模型的著作权风险治理要以促进知识生产者和使用者互利为目标，在追求二者权利义务平衡的同时，实现技术革新与作品传播。要以著作权法律规则的完善为前提，在提升司法公正与执法效率的同时，发挥人工智能行业自治的积极效果。

（一）文生视频著作权法律规则的完善

从视频生成过程看，如果用于机器学习的数据库不完整，其输出的结果极有可能不正确，有计算机科学家将其描述为输入垃圾并输出垃圾^⑤。著作权作为技术革新的副产品^⑥，其法律规则的完善应

① Williams Electronics Inc.v.Arctic International Inc, 685 F.2d 870 (1982).

②③ 黄薇、王雷鸣：《中华人民共和国著作权法导读与释义》，北京：中国民主法制出版社，2021 年版，第 121-122、110 页。

④ Evan H. Farr. Copyrightability of computer-created works, Rutgers Computer & Technology Law Journal, 1989 (15): 63-80.

⑤ Amanda Levendowski. How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem, 2018 (93): 579-630.

⑥ [美]R.F.沃尔、[英]杰里米·菲利普斯：《版权与现代技术》，王捷译，《中外法学》，1984 年第 6 期。

与印刷机技术、声光电传播技术以及互联网传播技术相呼应^①。为防止过宽的著作权保护阻碍文生视频大模型进步，同时，激励自然人创作更多高质量作品，应针对大模型著作权风险的不同来源作出有针对性的规则完善，确保其作为“表达自由的引擎”^②。当前，大模型在学习阶段使用的作品主要由自然人创作，这部分作品既是大模型生成高质量视频的主要来源，又是大模型生成能力持续提升的基础。因此，针对学习在先自然人创作的作品，应建立法定许可规则，即允许大模型可以不经许可使用在先作品，但应向著作权人支付相应报酬。除学习自然人作品外，大模型还可以二次学习机器生成的作品，由于这部分作品是通过对在先作品的演绎或拆分成补丁后重新组合生成，且在法定许可规则下已对在先作品著作权人支付了报酬，故而应允许机器无需许可免费进行二次学习，即针对机器生成的作品建立合理使用规则。上述规划的建立要以机器生成的作品不存在著作权问题为前提。由于大模型生成作品的价值在于使用，过强的著作权保护会降低生成视频的社会传播^③，影响人们的社会实践^④，因此，应对大模型生成作品的著作权进行必要限制。在著作人身权上，署名权具有表明来源、人格表彰和激励创新等作用，尤其是 Sora 相较 Pika 和 Runway 而言，无论视频时长还是画质均有显著提高。加之文生视频具有高仿真性，如果使用者使用其他模型生成的错误视频被冠之 Sora 生成，或后使用者对生成视频进行歪曲和篡改，将对大模型的评价产生不利影响。因此，应为文生视频的作者保留署名权和保护作品完整权。与此同时，文生视频大模型的开发目的之一是作为工具创作和发表作品，后使用者对生成的视频加以修改和完善是使用视频的通常做法，因此，无需为文生视频的作者保

留修改权和发表权。在著作财产权上，任何一项财产权都可能弱化数字网络环境下的作品传播，文生视频的数字化特点使展览权、出租权等适用空间趋近于零。为防止对文生视频的使用带来不必要的障碍，笔者认为，仅为其保留复制权等有限财产权即可，无需为其保留完整的著作财产权，同时，应缩短文生视频的著作权保护期。此外，基于大模型对在先作品的数字化使用和生成视频的规模化传播特点，加之司法实践中已经存在使用者基于作者身份主张对机器生成内容进行著作权保护的案例，可以预见，未来可能出现更多涉及文生视频的侵权案例。为减轻基于使用作品带来的授权许可压力，对文生视频著作权的保护应回归集体管理组织的运作，通过集体管理组织收取大模型的著作权许可费并代表著作权人维权。

（二）文生视频著作权案件裁判的司法进路

当前，人工智能的生成物是否构成以及在何种条件下构成著作权法意义上的作品，如何认定权利归属，抓取数据训练大模型是否构成侵权等问题还在讨论和探索过程中^⑤。除中国对其高度关注外，美国司法机关也面临类似问题。例如，2023 年，美国作家协会以人工智能公司 OpenAI 为被告提起著作权侵权的集体诉讼^⑥，Stability AI 被盖蒂图片社以著作权侵权为由告上法庭^⑦，这两起案件目前仍在审理中。由此可知，随着人工智能产业的迅猛发展，司法案例渐次产生，其中的著作权问题迫切需要司法回应。我国司法机关应在相关案件中积极反应，勇于探索并发挥审判指导作用，为国际社会文生视频的作品认定、著作权归属以及侵权判断提供借鉴。在文生视频的作品认定领域，应依据《著作权法》既有的作品定义条款加以判断。按照《著作权法》规定，“文学、艺术和科学领域内具有独创

① 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》，《山东大学学报（哲学社会科学版）》，2023 年第 4 期。

② Amy Adler. Fair Use and the Future of Art, New York University Law Review, 2016 (91): 574.

③ A. PLANT. The Economic Aspects of Copyright in Books, Economica, 1934 (1): 167-195.

④ Wendy J. Gordon. On Owning Information: Intellectual Property and the Restitutionary Impulse, Virginia Law Review, 1992 (78): 149-282.

⑤ 《专访最高法民三庭庭长林广海：加快调研人工智能生成物著作权问题 支撑服务生成式人工智能健康发展和规范应用》，<https://mp.weixin.qq.com/s/AhgDYHiZJq2HxvDLazSUGw>。

⑥ The Authors Guild, John Grisham, Jodi Picoult, David Baldacci, George R.R. Martin, and 13 Other Authors File Class-Action Suit Against OpenAI, <https://authorsguild.org/news/ag-and-authors-file-class-action-suit-against-openai/>.

⑦ Getty lawsuit against Stability AI to go to trial in the UK, <https://www.theverge.com/2023/12/4/23988403/getty-lawsuit-stability-ai-copyright-infringement>.

性并能以一定形式表现的智力成果”^①，同时，当前不属于单纯事实消息等不受《著作权法》保护的對象，应受《著作权法》保护。当前，文生视频大模型作为人类创作视听作品的工具，其生成的视频应属于人类的智力成果，对于独创性的判断应从结果出发，即如果文生视频大模型生成的作品在形式上与人类借助大模型创作的视听作品完全一致，且该作品是人类创作完成，同时被授予著作权，则满足独创性条件^②。在著作权归属领域，法院应首先审查当事人是否约定著作权归属，有约定遵从约定。如果该约定属于民事法律关系中的无效或可撤销约定，或违反法律的强制性规定，则可根据民法的公平原则、诚实信用原则、公序良俗原则以及《中华人民共和国民法典》相关条文进行规制。如果破坏市场竞争秩序，可通过《中华人民共和国反不正当竞争法》《中华人民共和国反垄断法》维护市场公平^③。如果没有约定或约定无效，应根据一般归属规则和特殊归属规则，并结合实质性贡献标准进行个案判断。在裁判结果上还应秉持后果主义标准，即是否在激励自然人创作和鼓励投资的同时，兼顾社会利益和国家利益。在著作权侵权的判断领域，首先，应判断是否构成《著作权法》意义上的使用行为。由于不同文生视频大模型的作品使用行为存在差异，应结合不同模型的技术原理，判断使用作品的行为是否为著作权法禁止，尤其在是否构成改编行为的判断上更应谨慎。一般而言，只有既使用原作品的基本内容，又在原作品内容上作了修改才构成改编。当前，Sora可根据用户输入的图像或视频，配以提示的文字生成视频，例如，Sora官网展示了提示词“一只戴着贝雷帽和穿着黑色高领毛衣的柴犬”以及一张图片，二者结合生成的视频使用了被输入图片的表达。此时，如果输入的图片属于侵权作品，生成的视频由于使用在先作品的内容创作出新的作品，则可能构成对在先作品的改编。但如果大模型仅提取原作品中的信息，此后根据用户指令将信息重新编排组合为视频，则可能不属于改编行为。其次，应判断对在先作品的使用是否构成法定许可或合理使用。当前，由于我国未设立法

定许可的兜底条款，且既有法定许可行为并不包括机器学习，所以如果大模型在训练过程中使用自然人的作品且构成侵权，应判决被告给予适当赔偿。但如果自然人或大模型使用机器生成的视频用于再创作，则应给予合理使用规则豁免。由美国法发展而来的转换性使用，通常被用作判断新技术引发的合理使用纠纷的主要标准。美国勒瓦尔法官认为，如果二次使用行为给原作品增加了新的价值，如将其作为原始素材引用，或在新信息的创作中进行转换，或增加新的审美、洞见和知识，这些行为就是合理使用制度意欲增进社会发展的使用类型^④。为实现积极的社会效果，虽然我国《著作权法》未制定机器学习的合理使用条款，但可将前文所述的一般条款和兜底条款相结合，按照《意见》第8条对兜底条款进行解释。此外，在判断两部视听作品是否构成实质性相似时，除要比对画面本身的表达外，还需判断生成视频的剧本等文字作品是否构成实质性相似。

（三）文生视频大模型平台的著作权监管

著作权人虽然可以通过提起诉讼或仲裁的方式寻求救济，但事后救济手段越来越难以应对数字化的著作权侵权行为。行政监管凭借主动、高效的监管特点可以弥补司法救济的不足，且著作权具备的促进文化发展的公益目标^⑤为著作权行政监管提供了正当性。2024年3月21日，联合国大会通过首个有关人工智能的全球决议草案，提出鼓励所有国家、国际组织、科技界、学术界、研究机构等制定框架，以支持针对AI的监管和治理。中国和欧盟分别制定《生成式人工智能服务管理暂行办法》和《人工智能法案》，用以促进生成式人工智能健康发展和规范应用。在互联网、大数据迅猛发展的背景下，我国对文生视频大模型的著作权监管应秉持促进创新与依法治理相结合的理念，致力于实现著作权平衡。事前监管主要审查大模型提供者是否履行事前说明义务。例如，欧盟制定的《人工智能法案》规定，人工智能办公室应监督通用人工智能模型提供者是否履行义务，即模型的提供者应就通用模型训练中使用的内容制定并公开足够详细的摘

① 《中华人民共和国著作权法》第3条。

② 易继明：《人工智能创作物是作品吗？》，《法律科学（西北政法大學學報）》，2017年第5期。

③ M.A.Lemley. A new balance between IP and antitrust, Southwestern Journal of Law and Trade in the Americas, 2007 (13): 237-256.

④ Pierre N. Leval. Toward a Fair Use Standard, Harvard Law Review, 1990 (103): 1111.

⑤ 吴汉东：《著作权合理使用制度研究》，北京：中国人民大学出版社，2013年版，第44页。

要，且该摘要的范围应在总体上全面，以方便包括著作权持有者在内的合法权益方行使和执行其在欧盟法律下的权利。例如，列出用于训练模型的主要公共数据库或数据档案，并对使用的其他数据来源进行叙述性解释^①。英国发布的《大语言模型和通用式 AI》报告^②也有类似规定。事中监管主要审查大模型提供者是否实施了著作权侵权行为。在侵权行为发生过程中，由于网络的匿名特征以及大模型算法黑箱的存在，导致监管部门和社会公众难以从外观上对侵权行为进行有效识别与判断。对此，可以遵循“以技治技”理念，由监管主体与大模型开发人员、著作权人以及第三方合作，围绕视听作品的使用与生成，用对抗性方法建立方便追踪、可及时干预的识别方式^③，实现动态跟踪并识别缺陷和漏洞^④，从而在发现漏洞后及时约谈相关主体，及时调整系统参数及架构，避免侵权。事后监管体现为对损害公共利益的侵权行为施以行政处罚。值得注意的是，在大模型发展过程中，如果尚不能完全了解文生视频的运作过程，那么，应对创新持更为包容的态度^⑤，不可将侵权行为泛公共利益化。因此，在大模型使用作品是否侵害公共利益的判断上，应持相对宽松的态度。公共利益指不特定个人可同时享有的一种利益，是特定社会存在和发展需要的、该社会群体中不确定的个人都可享有的社会价值^⑥。著作权领域的公共利益主要包括文化发展利益、消费者利益和公平竞争的市场秩序^⑦。从长远看，大模型合法使用在先作品对于文化发展具有积极作用，但应以向著作权人支付合理报酬为前提。同时，应通过要求大模型标明作品权利状态、指明作者等权利人身份等方式维护消费者利益。在公平竞争市场秩序的维护上，由于著作权侵权并非典型的不正当竞争行为，因此，应严格按照一般条款判断是否构成不正当竞争。此外，在互联网为大模型跨境学习数字作品提供便利的背景下，各阶段的监管都应注意加强著作权跨境执法合作。

（四）文生视频大模型平台的行业自治

司法与执法手段为解决著作权问题带来对抗性和不确定性，同时增加了纠纷解决成本。从文生视频技术原理看，大模型开启了视频和图片等输入并输出的通道，强化大模型对视频内容的审查义务是数字时代的必然要求。但基于大模型不同于传统内容提供者的特点，在审查内容上应有所区别。一方面，传统的内容提供者对网络平台上的内容具有较强的控制力和决定力，尤其是当内容提供者作为生产者时，可能其既是内容的传播者又是内容的著作权人。大模型作为人类创作作品的工具，其生成的视频源于使用者干预下与大模型之间的交流关系。例如，用户可以在 Sora 中输入两段关联度较低的视频，要求 Sora 生成连接两段视频的新视频。这导致大模型对作品来源合法性的控制力较弱。另一方面，大模型生成的视频主要根据用户指令，从数据库中提取补丁并组合成新的作品，不同于传统网络平台直接原样使用在先作品，其对在先作品的正常使用影响较小，对在先作者合法权益的损害较弱，因此，不应对大模型的训练与使用施加过于苛刻的义务。大模型在使用作品前对内容的审查，应区分是机器生成的作品还是非机器生成的作品。对于前者而言，可在未经著作权人许可且无需支付报酬的情况下直接使用；对于后者而言，应着重审查被使用作品的权利人是否声明不得使用，或是否应支付相应报酬。需要注意的是，作品使用前的审查不能采取结果主义，对输入视频的审查将耗费大量算力，在算力并不充足的情况下，无法从技术上完全实现零非法使用，因此，只要采取符合技术特点的内容审查，就应认定为履行了相关义务。此外，大模型的著作权问题归根结底是权利来源问题，为从源头提高作品来源的合法性，应倡导大模型投资者与著作权人之间营造互利共赢的著作权生态。一方面，大模型投资者可以通过集体管理组织授权以及签约创作者等方式建立正版数据库；另一方面，

① Artificial Intelligence Act, https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html.

② Large language models and generative AI, <https://publications.parliament.uk/pa/ld5804/ldselect/ldcomm/54/54.pdf>.

③ 朱嘉珺：《生成式人工智能虚假有害信息规制的挑战与应对——以ChatGPT的应用为引》，《比较法研究》，2023年第5期。

④ 高阳：《通用人工智能提供者内容审查注意义务的证成》，《东方法学》，2024年第1期。

⑤ 《专家谈人工智能知识产权：AI创作物是否受版权保护，如何监管》，https://mp.weixin.qq.com/s/PZW65fh4TAUw_e3SVE7BDw。

⑥ 麻宝斌：《公共利益与公共悖论》，《江苏社会科学》，2002年第1期。

⑦ 王洪友：《论版权行政执法的公共利益要件：以制度异化为视角》，《中国出版》，2020年第1期。

鼓励大模型之间建立著作权相互授权的商业模式，即不同大模型对其享有著作权的作品进行相互授权使用，避免类似 Stability AI 恶意爬取数据导致 Midjourney 服务器瘫痪事件的发生^①。在作品使用过程中，应建议避免对作品进行直接复制。具言之，建议大模型无需输入在先视频或图像等作品，而是将视频以帧为单位进行分割后，直接提取每一帧对应图片中的相关信息，分类存储到独立单元中，再根据用户指令调用相关信息组合成新的视频。对于输出的视频，一方面，可通过水印等方式对生成的视频内容设置明显标记，或给予“人工智能生成”的风险提示^②；另一方面，为防止技术审查环节出现缺漏，大模型和传播平台可以建立通知—反馈机制。相关部门在收到著作权人主张涉嫌侵权的通知后进行审查，确属侵权的，应允许权利人选择同意使用或退出机制。在著作权人同意的情况下可以支付合理报酬，如果著作权人不同意使用，则应在大模型或视频平台中删除或屏蔽。

四、结 语

以 Sora 为代表的文生视频大模型的出现，标志着视听作品生成技术取得重要飞跃，Pika 推出的伴音视频再次拓宽了人类的创作空间，将文生视频大模型的应用推向新纪元。这不仅为内容生成行业带来革命性工具或新质生产力，而且为突破时空限制提供了世界模拟器，可以预见，未来大模型或将以作品创作为基础，衍生更加多样的应用场景与业态^③，进而引发全球价值链重构。但无论人工智能技术如何发展，都应将其作为人类智能的模拟与延伸，秉持以著作权合规为技术研究与应用的发展红线，坚守以人为本这一著作权治理的基本立场。人工智能是世界数字产业建设的重要一环，开展“人工智能+”行动，推动人工智能大模型合规、高效和可持续发展，并产生“以点带面、连线成片”的影响效应是我国未来人工智能发展的重要目标。

Work Risks, Copyright Ownership and Effective Governance for Artificial Intelligence Text-to-video Large Model

CONG Li-xian LI Yong-lin

(School of Intellectual Property, East China University of Political Science and Law, Shanghai 201620)

Abstract: The text-to-video large model represented by Sora has brought about new discussion on work risks, copyright ownership and copyright risk governance. Text-to-video large model is a tool for human intellectual products creation, and the content generated by it constitutes its work. Based on the technical characteristics of text-to-video large model, the risks of work should be identified and judged based on different usages of works rather than being generalized. In the field of copyright ownership, we should adhere to the private property character of audiovisual works, take the party autonomy as the priority principle, and combine the substantial contribution, investment incentive and interest balance principles when distributing copyright ownership in cases. In the field of copyright risk governance, the copyright law improvement, copyright trial guidance, copyright law enforcement supervision and large model trade autonomy should work together to the compliance, efficiency and sustainable development of the text-to-video large model.

Key words: Sora ; Artificial Intelligence ; Text-to-video ; Work Risks ; Copyright Ownership ; Risk Governance

[责任编辑：李 蕾]

[责任校对：潘静静]

① Breaking: Stability AI Employees Banned From Midjourney, <https://twitter.com/PirateWires/status/1765484940957823486>.

② 邓建鹏、赵治松：《文生视频类人工智能的风险与三维规制：以Sora为视角》，<https://doi.org/10.14100/j.cnki.65-1039/g4.20240322.001>。

③ 吴汉东、刘鑫：《生成式人工智能的知识产权法律因应与制度创新》，《科技与法律（中英文）》，2024年第1期。