

·图书馆与图书馆事业·

文生视频类人工智能在图书领域的应用场景及有效治理*

李 涛

(1.中南财经政法大学法学院 湖北武汉 430073)

摘 要:文生视频类生成式人工智能作为一种“新质生产力”,在图书出版领域中具有广阔的应用前景。相较于 ChatGPT,文生视频类生成式人工智能 Sora 可赋能图书馆建设可视化数字资源、虚拟空间建设、智慧服务模式创新与用户个性化定制服务创新等。但与此同时,它也使图书馆面临侵犯人身权、知识产权、数据信息、算法安全以及虚假信息等风险的挑战。鉴于此,为消弭这些法律风险与科技隐患,应加强对人工智能 Sora 在图书领域应用风险的有效治理,明确人工智能生成物的著作权归属,并在此基础上进一步强化算法监管,提升图书馆的自我合规能力,使文生视频类生成式人工智能技术更好地服务于图书行业的可持续发展,共同营造一个健康有序的人机共存环境。

关键词:文生视频;Sora;生成式人工智能;图书出版;治理

中图分类号:G250.7;TP391.1

文献标识码:A

DOI:10.11968/tsyqb.1003-6938.2024049

Text-to-Video Artificial Intelligence in the Field of Book Application Scenarios and Effective Governance

Abstract As a kind of "New quality productivity", text-to-video generative artificial intelligence has a broad application prospect in the field of book publishing. Compared with CHATGPT, text-to-video generative artificial intelligence Sora enables the construction of library visual digital resources, virtual space construction, intelligent service model innovation and user personalized service innovation. But at the same time, it also makes the library face the challenges of infringement of personal rights, intellectual property rights, data information, algorithm security and false information. In view of this, in order to eliminate these legal and technological risks, we should strengthen the effective management of the risks of the application of artificial intelligence Sora in the field of books, and clarify the ownership of the copyright of artificial intelligence products, and on this basis to further strengthen the algorithm supervision and enhance library's self-compliance ability so as to make text-to-video generative artificial intelligence technology better serve the sustainable development of the book industry, and jointly create a healthy and orderly human-computer coexistence environment.

Key words text-to-video; Sora; generative artificial intelligence; book publishing; governance

文生视频类生成式人工智能 Sora 的出现,意味着人工智能自身的理解能力与内容表达能力实现了新的飞跃。2024 年 2 月,美国 OpenAI 公司对外公布了一款新型文生视频生成式人工智能 Sora,这是继 ChatGPT 之后又一技术革新。“Sora”一词取自日文“そら”(即天空),旨在表达“无限的创造潜力”之意。根据 OpenAI 所披露的技术报告显示,Sora 是一款可根据文字描述自动生成高清、连贯、仿真度极高的视频大模型,它彰显对物理世界的认知和理解,故

被誉为“物理世界的模拟器”。

2024 年全国两会期间,政府工作报告提出“加快发展新质生产力”。从其概念来看,新质生产力是与数字化生产要素相适应的生产力的质的跃进^[1]。在数字化时代,生成式人工智能作为一种新质生产力,它具有提高经济发展效率、优化升级经济结构、增强经济韧性等显著优势,能充分发挥科技创新的主导作用,以科技创新推动产业创新。与 ChatGPT、元宇宙等技术一样,Sora 作为一种新质生产力可在

* 本文系教育部人文社科重点研究基地重大项目“人工智能时代知识产权制度发展研究”(项目编号:22JD820030)研究成果之一。

收稿日期:2024-08-07;责任编辑:刘婷

多个行业领域中被广泛应用。如在影视制作领域,Sora 能提高影视创作的效率、降低制作成本,开辟叙事艺术表达新形式^[2];在教育领域,Sora 有助于实现“教”与“学”的可视化,打造情景化、沉浸式教育,提高学生探索意识与自主学习能力^[3];在网络游戏领域,其能帮助游戏开发者创建多元类型游戏场景和富有表现力的人物角色^[4]。基于此,本文尝试探讨生成式人工智能 Sora 在图书行业中的应用场景,并对其可能存在的法律风险进行回应,期冀为我国图书出版领域的数字化、智慧化建设及发展提供些许建议。

1 人工智能 Sora 的技术解析与工作机制

与以往发布的人工智能大模型相比,Sora 已然突破了过去的技术局限,其在深度学习和多模态处理能力上呈现明显优势,尤其在语言识别、机器视觉、图像生成等方面拥有强大功能。理解人工智能 Sora 的技术构成及其运作机制,对认识 Sora 的应用前景、风险防控、有效治理等方面具有重要启示性意义。

1.1 生成式人工智能 Sora 核心技术之解析

文生视频类生成式人工智能技术要比文字或图片类人工智能技术复杂得多,这是由视频大模型的技术特征所决定。视频大模型的难点主要在于视频图像时空属性、视觉连贯性和风格多样性、动态图像处理、人机交互性、视觉仿真性等方面。为了攻克这些技术难点,OpenAI 在 Sora 的视频大模型训练和处理中,采用“Transformer”+“Patches”+“Diffusion”三位一体结构,最终生成视频。

Transformer 是一种基于自注意力机制和前馈神经网络进行编码和解码的模型,通常可被用于自然语言处理、计算机视觉和音频处理^[5]。在计算机视觉中,Vision Transformer 由嵌入层、Transformer 编码器和 MLP Head 三个模块组成^[6]。Transformer 编码器主要负责提取输入图像的全局特征,通过自注意力机制实现对图像特征的信息交互或聚合。而 Sora 就是在这个由嵌入层和 Transformer 编码器所形成的潜在空间网络中进行训练与学习。OpenAI 技术报告中还提到,Sora 使用了一种时空图像块“Patches”技术。Patches 取代了之前 ChatGPT 训练模型中的 Token,

但二者具有相同的功能,它们都是处理文本或图像块的最小单位。在 Sora 的训练模型过程中,视频图像的每一帧画面都将被切割并转化为若干个一维向量,即若干个图像块 Patches^[7]。然后通过嵌入层中的“可学习嵌入”和“位置嵌入”机制,分别对图像块 Patches 进行编码并添加图像块序列的时间信息和位置信息,最终获得一个包含了完整信息的图像块。将这些图像块输入到编码器中以后,就可以进行分类预测^[8]。而 Sora 生成视频的过程实际上就是通过前一个图像块“Patches”去预测、填充并生成下一个图像块“Patches”,然后再将一系列图像块进行排序与组合,最终形成每一帧完整的图像。最后,扩散模型 Diffusion 对 Sora 影响甚大,它是一种先进的图像生成模型,通过马尔科夫链映射到潜在空间。扩散模型的基本原理是,它通过在数据中逐步引入噪声,然后反向去除噪声,从而生成高质量图像或音频。它分为两个步骤进行:首先是正向过程,即先给图像模型逐渐添加噪声,直到图像被完全损坏;其次是逆向过程,即根据给定的高斯噪声,逆向逐步恢复图像原始数据^[9]。当 Sora 完成模型训练后,随机输入高斯噪声,就能生成图像。

1.2 生成式人工智能 Sora 的基本工作原理

Sora 作为一款文生视频大模型,它可根据一段文字描述自动生成画面高清流畅、内容丰富多样、一镜到底的视频。从其基本工作原理来看,通过指令输入与识别、对抗生成网络、生成器网络、生成视频等步骤,即可实现从文字内容到视频效果的转化。这一技术的出现,预示着视频制作和内容创作方式的革命性变化。

第一,文本指令输入与识别。使用 Sora 进行创作时,使用者首先需要向 Sora 输入一段文字指令,或多个关键词,用以描述视频内容。这些文字被视为内容描述而成为 Sora 进行作业的依据。在给定的指令条件下,Sora 通过海量素材和训练模型深度学习并理解,进一步解析指令并执行相应操作。基于算法规则,人类输入的文本指令将被标注为各种不同的词性、语法分析,这些标注信息将帮助 Sora 理解句子中的词汇含义、上下文关系和语法结构,更准确地进行指令解析任务,实现更高效的人机交互^[10]。第二,视频

对抗生成网络。对抗生成网络是一种通过对抗性训练学习生成新数据的深度生成模型,被广泛应用于图像任务中。它由生成器 G 和鉴别器 D 两个神经网络组成,能够通过不同方式操纵空间和时间坐标,来改善运动状态。这两个神经网络通过对抗的方式进行训练,不断提升自身的生成和鉴别水平,最终得到逼真的高分辨率生成图像^[11]。第三,生成器生成视频。生成器网络是训练模型的重要组成部分,它由自动编码器网络、注意力模块和记忆模块构成,在 Sora 运行中主要负责输出下一时刻的视频帧。生成器可根据文本指令生成相应的视频帧,并将不同视频帧进行拼凑和补全,最终构成完整的视频内容。第四,视频内容输出。视频输出是 Sora 文生视频大模型的最后步骤。当生成器生成视频后,Sora 会运用 ChatGPT、DALL·E 的部分功能,将视频内容与输入的文本指令内容不断进行比对和修改,确保视频内容能更忠实、全面地符合用户的预期。视频输出后,使用者可通过点击观看的方式查看 Sora 生成的视频,并对视频内容进行修改或评价。以上就是 Sora 生成视频的全过程,通过这项技术特征可知,Sora 的出现将对社会、个人带来重大影响。

2 人工智能 Sora 在图书领域中的应用探索

不同于 ChatGPT,Sora 是生成式人工智能技术发展进程中重要的里程碑,它的出现将为内容创作者、教育服务、知识生产、娱乐产业以及普通用户带来更多可能性。在图书行业领域中,生成式人工智能 Sora 也将具有广阔的应用前景,主要集中在可视化数字资源建设、虚拟空间建设、智慧服务模式创新、用户个性化定制服务创新等方面。

2.1 可视化数字资源建设

图书馆拥有海量文献资源,是知识生产的“金矿”,但也面临文献利用低和文献分散等问题。要想让图书馆中的文献“活起来”,就必须引进生成式人工智能 Sora 这个“新质生产力”。Sora 具有强大的信息整合能力与视频内容生成能力,能够推动图书馆建立可视化数字资源知识库。具体而言表现在两个方面:其一,Sora 视频大模型依托海量数据训练和学习,可将图书馆中原本分散、碎片化的数字文献资源整

合形成互补并建立起体系化的知识库。它不仅有利于实体文献的保存与传播,还能够强化数字文献资源之间的联系与高效利用。其二,Sora 作为一个文生视频类人工智能,具有增强视觉效果的优势,通过 Transformer 计算机视觉技术可将文字内容自动转化为高清仿真视频,赋能馆藏文献以“文生图、文生视频”等多元化形式实现永续保存和传播。传统图书资源主要围绕实体文献和电子文献展开,并在此基础上进行知识生产。但这种单一资源内容难以满足图书馆自身的发展需求、以及不同用户的内容需求,也限制了知识传播的边际效应。将 Sora 与图书馆文献相结合,发挥 Sora 强大的视频生成能力,有利于把浩如烟海的图书内容转化为视频资源,使图书信息资源产生集合式效果与互换。

2.2 虚拟空间建设

虚拟空间建设是图书馆智慧化发展的新趋势,它离不开人工智能技术的加持。近年来,我国实践中元宇宙技术、VR 和 AR 等虚实技术与图书馆融合发展,正尝试构建“元宇宙图书馆”^[12]。但就目前元宇宙技术的发展现状而言,面临着内容匮乏且形式单一、设备体量大、运行成本高等问题。Sora 被誉为“物理世界的模拟器”,具有强大的模拟能力,可通过视频大模型构建三维立体式虚拟场景,推动智慧图书馆建设。将 Sora 与元宇宙技术相结合,能够为图书馆虚拟空间建设提供内容迥异、风格多样的视频内容,丰富元宇宙 AIGC 的新场景和新内容。在可预见的未来,Sora 赋能元宇宙技术后,将促使虚拟空间的建设进入快车道^[13]。随着人工智能技术的发展,在虚拟空间中,Sora 凭借自身虚拟现实和增强现实技术,能够提供高质量的视频内容和沉浸式体验感。通过 Sora 强有力的人机交互能力,能够准确收集、分析用户行为数据,为用户提供更逼真的虚拟现实体验^[14]。在图书馆虚拟空间中,通过 Sora 这一全新的数字镜像视角,图书知识中的人物角色、场景、细节将在数字孪生世界中得到展现,用户与虚拟人物角色能在虚构空间中进行人机交互,创作出更精彩的故事叙述。作为物理世界的模拟器,Sora 拥有内容、情景、条件等要素的整合能力,能为用户打造虚实融合、智慧化地沉浸式体验环境^[15]。

2.3 智慧服务模式创新

图书馆传统服务模式主要以实体文本文献为主,具有一定局限性。从知识传播的角度来看,传统的文字阅读在一定程度上消解了文字的可读性和知识的全面性。通常而言,文字阅读会增加语言歧义现象,尤其在汉语言文字中,同一个汉字或词语在内容含义上存在两种或多种解释,很难避免对内容理解时产生“似懂非懂”的结果。Sora 具备文生视频的能力,它可以使知识跳脱难以理解的“抽象、晦涩”藩篱,变得“可视”“可学”,富有趣味性。通过利用 Sora 文生视频技术可推动图书馆从提供纸质图书服务向可视化视频内容服务的转变,以更多元化的内容服务为用户提供更优质的体验。在智慧咨询服务中,Sora 的适用场景亦可表现在两个方面:其一,及时为用户答疑解惑。通过计算机视觉系统和自然语言处理技术,Sora 可对用户提出的问题进行解构,然后以易于理解的方式提供准确答案,用户仅需通过视频化的解题过程便可清晰得知解题步骤和解题思路。使用户“知其然,更知其所以然”,最终提高学习效率。其二,协助查找资料。在查询资料过程中,人们通常易受无效信息或资料的困扰,通过与 Sora 的人机交互方式,有助于提高检索资料的准确性与全面性^[16]。除此之外,在智慧推广服务中,通过 Sora 文生视频能力,图书馆可以制作图书预告片用于介绍书籍、作者或出版社,增加读者对特定书籍的了解和兴趣。还可以视频的方式发布书评和阅读体验,根据不同主题、风格或需求向读者推荐图书,提高读者选择的准确性和适配度。

Sora 虽然能为图书馆的发展带来前所未有的机遇,但与此同时也会对图书馆人力资源结构带来深刻影响。从人力资源结构变化的角度来看,Sora 将使图书馆某些传统职业的减少或被替代,以达到可持续发展和就业机会转型。具体而言,伴随 Sora 技术的发展,将增加以下就业机会:(1)视频图书创作人员。这些人员将专门根据纸质图书或电子图书将文字内容转化为视频,确保视频内容符合文本内容。(2)视频内容审查和监管人员。随着 Sora 技术的普及使用,图书馆需要有人专门负责审核视频内容的正当性、合法性与合理性。(3)视频后期处理人员。这些

人员将主要负责视频的后期制作、处理,如对视频内容进行标识来源、分类管理等。(4)视频技术研发人员。这些人员应当具有人工智能、算法、大数据合成等专业知识,能够对视频内容、效果的生成技术和增强技术进行改进和修复。

2.4 用户个性化定制服务创新

如果说 ChatGPT 在用户个性化定制方面已表现出强大能力,那么 Sora 所能提供的个人化定制将会更加丰富。从教育服务的角度来看,Sora 文生视频的能力有助于用户知识学习变得简单而高效^[17]。将 Sora 与图书相结合,可将书籍里原本晦涩难懂的知识点变得更易理解。尤其是抽象概念、关系等“只可意会不可言谈”的知识点,通过 Sora 的展示和讲解可在现实世界中找到其具体表现;过去曾发生的历史事实和文献,通过 Sora 的视频模拟,将能情景再现。对于普通用户而言,在学习上经常面临知识碎片化的考验,如果能够利用 Sora 将各学科知识以案例再现、虚拟场景的形式进行视频内容转化,无疑将帮助用户建立起完整的知识体系。

在个性化定制服务方面,图书馆可通过 Sora 为用户定制个性化学习计划和知识素材。在日常生活中,学校的标准化教育通常难以实现每个人的个性化发展和需求,往往缺乏对每个学生实际需求的关注。而 Sora 可根据每个人的观看记录、人机交互行为数据,对用户的兴趣爱好、学习进度、知识储备、学习习惯、学习目标等进行评估,从而生成个性化的课程视频,为用户提供更合适的学习资料。在个性化体验方面,Sora 能够生成视觉上、审美上不同学习风格的视频供用户选择,以提高学习效率和学习兴趣。通过分析用户的行为数据、成长经历、学习效果,Sora 可以在不同的时间、空间环境下为其提供合适的学习内容。考虑到 Sora 的模型训练结合了 Transformer、Patches、Diffusion 等先进技术,代表着算法应用层面的最新发展,因此图书馆可有效利用 Sora 技术来提高自身的算法创新,更新原有数据和算法模型,从而促进通用模型的自主性和持续学习能力。通过算法技术优化以后,图书馆将能够为读者提供更加精准、智慧化的推送服务^[18]。如基于用户的观看习惯、用户评价、兴趣等,Sora 能够进行播放列表定制和视频

内容推荐,为其创建个性化的播放列表,方便用户查看自己感兴趣的视频内容。

3 人工智能 Sora 在图书领域中的法律风险

Sora 在图书馆中的应用可谓前景广阔,尤其是在文生视频方面呈现出良好的发展态势,势必会对未来的图书馆建设带来重大变革。由于 Sora 主要依托海量语料库和视频数据等进行训练和学习,其在图书馆中的应用可能会面临一些潜在的法律风险。这些法律风险分布在 Sora“生成视频的过程中”与“生成视频后”两个阶段,主要表现为:侵犯人身权风险、知识产权风险、数据安全风险、虚假信息风险等。

3.1 侵犯人身权的风险

近年来,随着生成式人工智能技术发展的突飞猛进,其对个人姓名权、肖像权、名誉权、隐私权等人格权益的侵害风险逐步呈扩大趋势。如利用自然人声音作为数据进行 AI 模型训练。根据民法典规定,这种行为构成侵犯他人声音利益,也不属于《著作权法》中“改编、演绎作品”所界定的法定许可范围。生成式人工智能 Sora 的出现将使得个人的声音、肖像等可以被轻松提取与合成,从而引发侵权纠纷^[19]。根据《民法典》第 1023 条之规定,对自然人声音的保护,参照适用肖像权保护的有关规定。肖像权是指自然人对自己的肖像享有再现、使用或许可他人使用的权利^[20]。未经肖像权人同意而使用他人肖像构成侵犯肖像权的行为。Sora 视频大模型是根据视频、图像数据而生成模拟视频,如果使用者未经他人同意输入他人的肖像或声音从而生成相同人物图像和音频的一段视频,便可构成侵权。如在实践中,部分网络用户未经死者近亲属同意,利用逝世明星的肖像、声音制作、发布“AI 复活”系列短视频,以温情之名非法牟利,构成侵犯死者人格利益的行为。倘若不对 Sora 所生成视频的内容进行适当审查或规范,那么 Sora 无疑将成为侵权人实施侵权行为的“帮凶”。

3.2 知识产权风险

从 Sora 的视频大模型训练过程来看,其在核心技术中常会利用原始视频或图像进行压缩、转化为

一维向量,然后使 Sora 在潜在空间中进行训练与学习,最终生成视频。在对视频大模型进行训练时,可能会发生侵犯他人著作权的行为。根据《著作权法》第 26 条规定,使用他人作品应当经他人许可或同意。如果未经他人许可或同意,且不构成合理使用或法定许可使用的,则构成侵犯著作权的行为。在视频大模型训练中,设计者输入原始文本、图像或视频等作品的行为,属于对原作品的复制行为;而在后续的使用过程中,如果基于原作品进行改写或者扩展成视频,则属于对原作品的改编行为。上述两种行为如果未经过原作品的著作权人许可,且不构成合理使用的,那么就可能构成侵犯著作权的行为。如在我国首个 AIGC 平台侵权案件中,法院认为被告 AIGC 平台多次使用侵权数据训练 AI 大模型并生成相应侵权图片,侵犯奥特曼系列作品的复制权和改编权;其未建立投诉举报机制、欠缺潜在风险提示和显著标识等行为即表明被告未尽到合理注意义务,应承担侵权责任^①。

结合图书领域而言,如果要对文生视频大模型进行训练,或者想通过在先作品生成视频,图书出版商、图书馆等相关主体应事先取得原作品的著作权人许可或同意,以避免发生著作权侵权纠纷。除此之外,当 Sora 生成视频以后,该视频是否具有独创性与可版权性,还可能存在著作权归属争议。其一,Sora 生成的视频是否属于著作权法上的“作品”?其二,出版社或图书馆通过 Sora 生成视频后,谁是享有视频作品著作权的法律主体?由此可见,Sora 在图书馆中的应用,将为图书馆知识产权保护机制带来新挑战。

3.3 数据安全、隐私和算法风险

从 Sora 的运行机制可知,它通过视频图像数据、大量语料库数据来训练大模型。这个过程中会涉及视频图像数据的抓取、使用者个人信息的收集以及行为数据的采集。如用户与 Sora 在人机交互过程中,会产生交互行为数据^[21];利用传感设备、移动操作设备等高精度设备可以获取文献信息资源。在 Sora 大模型训练过程中,设计者通常利用图书资源数据和用户数据训练文生视频人工智能模型,以满足机器

①广州互联网法院(2024)粤 0192 民初 113 号判决书。

自主学习的需求。Sora 在运行时所收集的用户相关数据和信息越多,其能为用户提供个性化服务的质量就会越高,也能推动图书馆在内容生成、服务创新、知识提取等领域的智能升级^[22]。然而,在这个过程中,极易出现 Sora 过度收集用户行为数据和个人信息的情形发生。倘若图书馆或设计者在对视频大模型进行训练或者在生成视频的过程中,不遵守法律法规的相关规定,将可能存在数据泄露和数据滥用的风险。如 ChatGPT 曾被曝光在未遵守“告知同意”规则下,过度收集用户个人信息且将用户信息共享给第三方平台,从而导致用户数据、信息、隐私泄露。更甚者,如果收集、使用的数据涉及公共利益和国家利益,那么数据泄露则将导致重要机密流失海外。

此外,与其他人工智能一样,以 Sora 为代表的文生视频类生成式人工智能并不是一个没有任何偏见的工具,其所生成的视频内容会受到算法规则的影响。算法是人为的产物,Sora 所生成的内容很大程度上取决于人类在视频模型训练过程中使用了哪些数据和原始视频作为素材对其进行训练和学习。这意味着,Sora 可能会受到人类的特定政治倾向、价值取向或其他利益的影响,最终产生算法偏见与歧视^[23]。如有媒体曾曝光谷歌旗下一款 AI Gemini 在其生成的图像中存在种族歧视问题,随后谷歌对此回应称 Gemini 所呈现出的问题,不过是 AI 领域长期存在的歧视问题的另一种表现。在 Sora 后期的深度学习和自主学习过程中,由于视频大模型需要不断接受数据喂养,如果使用者提供的数据失真、数据样本不全面或误导性信息,那么其所生成的视频内容也将会带有片面性和偏见性,无法确保输出内容是客观且公正的。

3.4 虚假信息风险

随着生成式人工智能被广泛应用于文本、图像、音频和视频等各大领域,信息传播正经历着深刻变革,大量虚假信息滋生其中。生成式人工智能技术在生成虚假信息方面具有速度快、体量大和仿真性高的特征,它与虚假信息的耦合为虚假信息泛滥提供了“新路径”^[24]。虚假信息是指不符合事实真相的误导性信息内容。虚假信息的生产者和传播者通常以隐瞒事实真相为目的,具有故意欺骗和误导他人的主

观动机。既可能是出于社会动机,也可能是出于个人动机或对抗需要,如操控民众、损害他人声誉等^[25]。随着生成式人工智能技术与算法演进,信息的深度合成与伪造将会在个人、社会以及国家安全等不同层面产生风险隐患^[26]。在实践中,目前深度合成伪造的现象频发,如人像换脸、深度美颜、合成特效等情景,对视频拍摄的内容进行操纵,增加了虚假信息产生的几率。从社会秩序角度来看,生成式人工智能技术使得虚假新闻、虚假视频制作成本更低、传播速度更快、欺骗性更强,给社会秩序和社会信任机制带来挑战^[27]。在国家安全方面,虚假信息将对政治稳定、国防安全、经济安全等带来隐患。因此,在图书出版领域中,出版商、图书馆在生产视频内容、提供图书视频服务过程中,应尽量确保视频内容的妥当性与合法性,必要时对视频制作与内容进行全面审查和监管。

4 人工智能 Sora 在图书领域应用风险的有效治理

对生成式人工智能 Sora 在图书领域的适用存在的风险进行有效治理,应以促进行业可持续发展为目的,在追求知识生产和风险防控之间寻求平衡,最终实现技术革新与知识传播。有鉴于此,结合人工智能 Sora 的技术特征与图书行业的创新发展,应分别从个人权益保护、著作权保护与权利归属、算法监管、提升合规能力等维度进行回应,将法律风险置于可控之范围内。

4.1 加强对人格权、知识产权的法律保护

在数字图书领域中,出版商、图书馆等相关主体利用文生视频类人工智能将文字图书转化为视频图书,在制作视频时应避免侵犯他人姓名、肖像、名誉、声音等人身利益的行为。对视频大模型进行训练时,如果训练数据样本中涉及使用他人肖像、声音等人格利益,应当事先取得他人授权或许可,且授权的内容和范围亦应明确清晰。如果需要收集用户、读者的个人信息进行模型训练,根据《民法典》《个人信息安全法》的相关规定,个人信息处理者在收集个人信息时应当取得个人的同意,且应遵循比例原则,将可能给个人造成的损害限制在最小范围。使用个

人信息时,不得超出授权范围而使用,不得另作其他商业性用途。

从作品授权层面而言,生成式人工智能大模型的训练通常需要海量素材和原始作品作为数据样本,才能生成文本内容、图像或视频。因此,在大模型训练时,为减少基于使用在先作品带来的授权许可压力和成本,可以通过建立一个集体管理组织来专门负责在先作品著作权的授权和许可工作。通过集体管理组织的授权或许可,有利于降低缔约成本和时间成本,生成式人工智能的设计者、服务提供者便可在大模型训练阶段付费使用在先作品进行训练。如果发生侵权行为,集体管理组织还可代表原著作权人进行维权。

4.2 明确 Sora 生成视频的著作权归属

出版商、图书馆使用文生视频类生成式人工智能 Sora 生成视频后,尚且存在两个问题有待解决:第一,Sora 生成的视频是否属于著作权法上的“作品”?第二,该作品的著作权归属主体是谁?有观点认为摄影作品的著作权保护路径对生成式人工智能创作物的可版权性具有重要的启示意义^[28]。那么摄影作品的独创性体现在哪些方面?通常而言,摄影作品是由摄影师事先设定相机参数、构图创意、选取场景和拍摄角度、光影技术等再按下快门键。由此所拍摄的照片便能体现摄影者的审美取向,通常会被认为具有独创性^[29]。对比生成式人工智能 Sora 而言,人们在使用 Sora 时,通过设定提示词或限定词的方式按照算法规则输入最终生成符合使用者预期的文本内容或图像。这个过程同样能体现使用者的创造性劳动,而非简单的机械性过程。如在我国首个“图片类生成式人工智能创作”判决书中,法院认为使用者在设计人物的呈现方式、选择提示词、安排提示词顺序、设置相关参数、选择符合预期的图片等方面均体现使用者的智力投入和个性化选择,因此人工智能创作物具备“智力成果”和“独创性”要件,应当认定为作品^①。虽然这起案件本身涉及的是“图片类”生成式人工智能作品,但其裁判结论对文生视频类生成式人工智能 Sora 所生成的视频被视为作品具有重要

的参考意义。从 Sora 生成视频的运行原理来看,出版商、图书馆作为使用者输入提示词、设定参数后通过 Sora 生成相应视频,其所输入的提示词、设定参数等行为同样体现了出版商、图书馆作为使用者的个性化选择和智力投入,因此生成式人工智能所生成的视频构成著作权法意义上的作品,应受著作权法保护。

生成式人工智能 Sora 生成视频以后,该视频的著作权归谁?对这一问题,目前存在两种不同观点。一种观点认为,文生视频大模型生成的视频作品融入人类的智力活动,属于人类的智力成果^[28]。另一种观点认为,机器学习技术下的算法自由实现了算法创作物的表达自由。其表达过程已不再受人类控制,表达结果已超出人类预期,不能将输入提示词得到 AI 创作物的用户确立为作者^[30]。对此,笔者认为,文生视频类生成式人工智能 Sora 生成视频的著作权归属应先遵循意思自治原则,没有约定或者约定无效的,则按贡献程度的动态认定模式进行确定。意思自治乃民法之根基,知识产权法虽作为民事特别法,亦应遵循民法“私人自治”的核心价值。如 OpenAI 公司作为生成式人工智能的设计者和提供者,其在“使用协议”中已明确约定:“在用户和 OpenAI 之间,用户保留输入内容的所有权;且同时 OpenAI 将输出中的所有权利、所有权和利益转让给用户。”由此可知,当存在著作权归属的有效约定时,该约定只要不违反法律法规属有效,应优先适用。如果关于著作权归属的协议约定无效或者未明确约定的,则按照使用者与设计者、服务提供者之间的贡献程度进行确定。在人机协作过程中,AI 使用者、AI 设计者、AI 服务提供者之间实际上存在着共同创作的关系。在这种共创模式下,很难一概而论地就断定哪一方就是著作权主体。因此,可以采用“按照贡献程度”的动态认定模式去判断 AI 作品的著作权归属。这种认定模式与著作权法中作品的“智力成果”要件以及“独创性”要件之标准相契合,是一种根据个案中各方主体在 AI 作品内容中的实质性贡献程度大小来认定作品的著作权归属,有利于平衡各方之间的利益关系,实现公平正义价值。

①北京互联网法院(2023)京 0491 民初 11279 号判决书。

就图书出版领域而言,图书馆在使用生成式人工智能 Sora 生成视频后,该视频的著作权归属认定逻辑是:约定优先,没有约定或者约定无效的,则按照贡献程度的动态认定模式进行确定。其中“贡献程度”应聚焦于最终呈现的版权法意义上的作品“内容”进行判断,根据图书馆、AI 设计者、AI 服务提供者等主体与 AI 作品之间的关联性、内容策划和设计、研究和分析、数据收集和统计、资金投入、智力投入程度、有偿或无偿等要素动态认定。如图书馆将图书文本内容作为提示词输入生成式人工智能 Sora 生成相应视频,输入的提示词越详细,其与最终视频作品内容的关联性越高,该视频的著作权应由图书馆享有。

4.3 强化人工智能算法监管

我国在算法领域目前制定了以场景和架构为基础的算法标准、法律法规。如《信息服务算法推荐管理规定》《算法综合治理的指导意见》《深度合成管理规定》以及《机器学习算法安全评估规范》(征求意见稿)等。但面对人工智能技术带来的种种挑战,仅凭法律一己之力实难完全应对。基于人工智能的技术风险外溢所呈现的安全隐患,有必要重新审视人工智能的治理理念。就此而言,对生成式人工智能算法治理需要秉持“协同共治”的治理理念,强调多元力量共同参与。

在协同共治的治理理念下,对生成式人工智能 Sora 的治理主要包括治理主体多元性、法律功能互补性、治理模式协同性、价值取向公共性等内容。(1)治理主体多元性。人工智能技术汇聚互联网、大数据、云计算等现代信息技术,在创新发展与场景运用过程中,各方主体都对人工智能的治理存在不同关切。政府、人工智能开发者和设计者、图书馆、公众等,在涉及人工智能的关系网络中都不同程度地扮演“实质性参与”角色。鉴于此,应当建立一个由政府主导、图书馆和研究技术人员等社会力量参与的多元主体共治平台。(2)法律功能互补性。承前所述,人工智能 Sora 在生成视频过程中与视频生成后可能存在多方面法律风险,甚至可能涉及刑事风险。因此,不仅涉及民事法律,而且关乎行政、刑事法律法规,故需要私法与公法的协同和互补。(3)治理模式协同性。现阶段我国正尝试建立以政府监管部门为主、行业

合规自治为中心的动态调整治理模式。但在技术发展过程中,对人工智能产品引发的新问题和新风险的防控机制尚存在不足。面对同一问题,德国建立了由算法伦理委员会和自动化决策系统专家等第三方组成的独立组织,适当介入参与治理。加强第三方监管,可在一定程度上保证自动决策和算法的公正性和透明性^[31]。且第三方参与具有一定独立性,有利于降低监管成本。(4)价值取向公共性。人工智能算法治理应弘扬人文价值,这不仅是人工智能技术发展的道德要求,也是伦理规范。为了应对人工智能 Sora 在图书领域运用的风险与挑战,应强化对人机交互关系的伦理规约和行业自律规范,防止“算法黑箱”对人的主体地位、认知团结和共识凝聚的潜在威胁。当前,算法黑箱正在“加剧社群内部分化,部分弱势群体或将面临信息不透明的挑战,又或是成为算法霸凌的受害者,动摇普遍意义上的人本主义立场”^[32]。人类社会不能以人机关系替代人际关系,不能让人工智能判断替代人类价值判断,不能让机器数据和算法共识替代人类的普遍共识。因此,在算法监管方面还应当坚持法律与道德、伦理以及行业自律规范协同共治、并驾齐驱的治理模式,共同维护国家主权、公共利益和图书行业的发展利益,使治理成为一种善治。

4.4 提升图书馆的自我合规能力

由于人工智能 Sora 主要以自动生成视频为载体,因此图书馆应当建立起与视频内容相关的风险防范机制。这主要涉及视频内容审查与虚假信息防范两个方面。就视频内容审查而言,应分别建立实质性审查与形式审查相结合的模式。在内容实质性审查上,为避免人工智能生成的视频内容存在有害、色情、歧视性或不道德内容,应以公序良俗、社会主义核心价值观作为视频内容的一般审查标准。加强对中华优秀传统文化的保护,提升文化自信,杜绝包含宣扬暴力、分裂社会、制造对立、侮辱性等色彩的视频内容输出。在形式审查上,应根据《深度合成管理规定》《生成式人工智能服务内容标识方法》的规定履行内容标识义务^[26]。标识的方式主要有两种:即显式水印标识和隐式水印标识。显式水印标识是指在人工智能生成内容的显示区域中添加显示带有“由人工智能生成”或“由 AI 生成”的提示文字;隐式水

印标识是指在由人工智能生成的图片、音频、视频时,应通过时空域水印或变换域水印的方式实现标识信息。通过这种方式要求相关主体在发布、传播 Sora 生成的视频内容时应以显著的方式对视频来源、内容质量进行标识和审查,能够强化主体责任。

为了保障信息内容安全,营造清朗的网络环境与维护社会稳定,图书馆在运用人工智能 Sora 时,应构建起虚假信息的防范机制。《生成式人工智能服务管理暂行办法》第 15 条规定了“投诉、举报”规则以及辟谣机制,要求对虚假信息的生成和传播进行预测、监控及治理。在图书馆利用人工智能搭建起的数字化信息平台中,用户与 Sora 进行人机交互时可随时上传、发布、共享自己创作的视频作品。对于这些视频作品内容的妥当性、真实性与合法性,图书馆应当及时进行审核,建立动态监管机制,防止有害信息、虚假信息流动到公众视野。图书馆对虚假信息防范的范围应适当明确,即以有害信息控制为目的,以维护国家和社会安全为基本要求,推动中国式现代化进程中的数字文化强国建设。然而,对虚假信

息的治理并非一家之责,还需立足于现行法律法规结合技术发展水平,探索多手段合作治理的路径,才能真正建立起健康有序的人机共存的生态。

5 结语

以 Sora 为代表的文生视频类生成式人工智能的诞生,标志着人工智能技术发展有了质的飞跃。当下阶段,人工智能技术正成为一种“新质生产力”,衍生出多样化的应用场景与业态。文生视频类生成式人工智能在图书领域中有着广泛的应用前景,包括建立可视化数字资源知识库、建设虚拟空间、创新智慧服务模式、实现个性化定制服务体验等。但与此同时,其也带来了侵犯人身权、知识产权、数据安全、算法以及增加虚假信息等风险。为消除这些法律风险和科技隐患,让科技能更好地服务于图书行业的可持续发展,应当加强对人格权、知识产权的法律保护,厘清 Sora 生成视频的著作权归属。在此基础上进一步强化算法监管和视频内容审查,以营造清朗的网络环境,建立一个健康有序的人机共存发展前景。

参考文献:

- [1] 何自力.新质生产力理论的科学内涵和时代意义[J].中国高校社会科学,2024(3):4-14,157.
- [2] 令小雄,王鼎民,唐铭悦.ChatGPT 到 Sora;Sora 文生视频大模型对影视创作的机遇、风险及矫治[J].新疆师范大学学报(哲学社会科学版),2024(6):91-99.
- [3] 朱光辉,王喜文.人工智能文生视频大模型 Sora 的核心技术、运行机理及未来场景[J].新疆师范大学学报(哲学社会科学版),2024,45(4):149-156.
- [4] Perez-Liebana D,Liu J L,Khalifa A,et al.General Video Game AI:A Multitrack Framework for Evaluating Agents,Games, and Content Generation Algorithms[J].IEEE Transactions on Games,2019,11(3):195-214.
- [5] 周诚辰,于千城,张丽丝.Graph Transformers 研究进展综述[J].计算机工程与应用,2024(1):1-14.
- [6] Beraldo G,De Benedictis R,Cesta A,et al.Toward AI Enabled Commercial Telepresence Robots to Combine Home Care Needs and Affordability[J].IEEE Robotics And Automation Letters,2023,8(10):6691-6698.
- [7] Luo G,Zhou Y,Sun X,et al.Towards Lightweight Transformer Via Group-Wise Transformation for Vision-and-Language Tasks[J].IEEE Transactions On Image Processing,2022,31:3386-3398.
- [8] Afrasiabi M,Khotanlou H,Mansoorizadeh M.DTW-CNN:time series-based human interaction prediction in videos using CNN-extracted features[J].The Visual Computer,2020,36(6):1127-1139.
- [9] Croitoru F A,Hondru V,Ionescu R T,et al.Diffusion Models in Vision:A Survey[J].IEEE Transactions On Pattern Analysis And Machine Intelligence,2023,45(9):10850-10869.
- [10] Wu C H,Liu C H.Sentence Correction Incorporating Relative Position and Parse Template Language Models[J].IEEE Transactions on Audio Speech and Language Processing,2010,18(6):1170-1181.
- [11] Xiao J,Bi X J.Multi-Scale Attention Generative Adversarial Networks for Video Frame Interpolation[J].IEEE Access,2020,8:94842-94851.

- [12] 郎林芳,黄世晴,王珏,等.元宇宙图书馆阅读推广服务创新发展研究[J].图书馆杂志,2023,42(10):55-63.
- [13] 喻国明,苏芳.作为真实世界模拟器的媒介与后真相时代的“拨乱反正”[J].新疆师范大学学报(哲学社会科学版),2024,45(4):143-148.
- [14] 柯和平,龚自振,谢海先.智慧图书馆建设中混合式智慧学习空间的构建研究[J].现代教育技术,2024,34(4):112-121.
- [15] 陆道坤,陈吉钰.Sora:学校教育的“终结者”还是“拯救者”[J].新疆师范大学学报(哲学社会科学版),2024(6):110-125.
- [16] 杜牧真.智慧图书馆建设背景下图书馆交互数据的权利定位与全能[J].图书馆建设,2024(1):96-107.
- [17] 冯雨旻.AIGC 时代教育的视频化转向及其风险应对——由文生视频模型 Sora 引发的思考[J].思想理论教育,2024(5):27-33.
- [18] 尹克寒,陈纪文.AGI 浪潮下 Sora 何以赋能智慧图书馆建设[J].图书馆理论与实践,2024(3):12-19.
- [19] 李涛.论声音权在人格权编中的确立[J].三峡大学学报(人文社会科学版),2019,41(3):92-96.
- [20] 马俊驹,余延满.民法原论(第四版)[M].北京:法律出版社,2016:107.
- [21] 王艺,曹越.AIGC 技术赋能下图书选题智能策划系统的技术逻辑与体系建构[J].出版发行研究,2024(2):31-37.
- [22] 赵杨,张雪,范圣悦.AIGC 驱动的智慧图书馆转型:框架、路径与挑战[J].情报理论与实践,2023,46(7):9-16.
- [23] 钟晓东.论生成式人工智能的数据安全风险及回应型治理[J].东方法学,2023(5):106-116.
- [24] 胡泳.人工智能驱动的虚假信息:现代与未来[J].南京社会科学,2024(1):96-109.
- [25] 李卫东.数字组织论[M].武汉:华中科技大学出版社,2022:113.
- [26] 漆晨航.生成式人工智能的虚假信息风险特征及其治理路径[J].情报理论与实践,2024,47(3):112-120.
- [27] 胡泳.AI 视频的兴起:Sora 类生成式平台的可能性与风险[J].传媒观察,2024(4):5-19.
- [28] 丛立先,李泳霖.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[J].新疆师范大学(哲学社会科学版),2024(6):153-163.
- [29] 徐小奔.论算法创作物的可版权性与著作权归属[J].东方法学,2021(3):41-55.
- [30] 徐小奔.论人工智能生成内容的著作权法平等保护[J].中国法学,2024(1):166-185.
- [31] 张淑玲.破解黑箱:智媒时代的算法权力规制与透明实现机制[J].中国出版,2018(7):49-53.
- [32] 曹克亮.Sora 的意识形态效应及其治理[J].统一战线学研究,2024(3):166-178.

作者简介:李涛(1993-),男,中南财经政法大学法学院博士研究生,研究方向:民商法学、知识产权法。