

·特别策划:ChatGPT 与图书情报事业发展·

生成式人工智能给图书馆带来的机遇、挑战及应对策略

李颖婷

(1.呼伦贝尔市图书馆 内蒙古呼伦贝尔 021008)

摘要 近年来,在算法、数据和算力的综合推动下,生成式人工智能(AIGC)的发展进入爆发期,AIGC 的应用在各行各业迅速展开。对图书馆行业来说,AIGC 的应用可大大提升用户体验,为图书馆创新服务模式和内容带来无限可能,但也潜藏着诸多不利因素。如何规避风险,最大化利用 AIGC 助力图书馆发展,需要图书馆提前谋划,制定出有效的应对策略。

关键词 生成式人工智能;AIGC ChatGPT;图书馆

中图分类号:G251 文献标识码:A DOI:10.11968/tsyqb.1003-6938.2023021

Opportunities, Challenges, and Strategies for Libraries in the Era of Artificial Intelligence Generated Content

Abstract The advent of Artificial Intelligence Generated Content (AIGC), exemplified by OpenAI's large-scale model ChatGPT, has garnered global attention and sparked a new wave of AI revolution. While AIGC offers opportunities for intelligent information retrieval, transformative library experiences, and the development of metaverse libraries, it also has the potential to alter readers' search and reading habits, amplifying the existing technology and talent shortages in libraries. There are also risks involved, including copyright infringement and privacy breaches. To address these challenges, libraries can begin by enhancing and expanding their services through the implementation of chatbots and exploring the use of open-source models to train on existing library literature data. It is crucial for libraries to proactively address and mitigate the risks associated with the application of AIGC.^①

Key words Artificial Intelligence Generated Content (AIGC); ChatGPT; intelligent libraries; intelligent retrieval; opportunities; strategies

2022年11月30日,由美国人工智能研究公司 OpenAI 研发的生成式人工智能聊天机器人程序 ChatGPT 正式上线,短短 5 天内注册用户数就突破百万,到 2023 年 1 月末,其活跃用户数已达到 1 亿,成为史上活跃用户增长最快的应用程序。ChatGPT 的推出引发了相关领域企业的强烈反映。2月2日,OpenAI 的合作伙伴微软公司宣布旗下所有产品将全线整合 ChatGPT,包括搜索引擎必应、Office 以及云计算平台 Azure 等,使微软的所有产品具备 AI 能力。在国内,2月8日,阿里巴巴确认正在研发“阿里版聊天机器人 ChatGPT”;2月10日,京东宣布了其产业版 ChatGPT——智能人机对话平台 ChatJD 的落地应用路线图;3月16日,百度正式发布其生成式 AI 产品“文心一言”。与此同时,ChatGPT 更是引发了

各行各业对生成式人工智能应用场景及其对各行业影响的大讨论,甚至已经在为 AIGC 可能带来的社会变革做好准备。可以预见,作为通用型多任务人工智能对话系统的一个里程碑,ChatGPT 的推出将引发更多的机构加入到 AIGC 的应用及开发中来,对包括图书馆在内的各个行业也必将产生深远影响。本文通过分析生成式人工智能(AIGC)的特点、发展趋势及图书馆应用 AIGC 面临的机遇与挑战,提出图书馆的应对策略,以期为业界的相关研究和实践提供参考。

1 AIGC 的特点及发展趋势

2014 年,由前谷歌大脑研究科学家 Ian Goodfellow 等首次提出的生成对抗网络模型(GAN)被认为

收稿日期 2023-04-17 责任编辑 刘婷

是生成式人工智能发展的开端^[1]。GAN 使用两个相互对立的神经网络即生成器和识别器经过交替周期训练赋予了人工智能深度学习模型以想象力,使其更接近人类思维^[2]。而近几年来,随着新的模型结构如 Transformer 等的出现以及硬件能力的不断提升,具有庞大数据参数的生成式人工智能大模型逐渐代替了之前的小模型,如 OpenAI 的 GPT-3 已经具有 1750 亿参数、96 层 Transformer 解码器^[3]。由于参数越多的模型其表现越出色,加上软硬件技术的发展提供的巨大算力,2022 年 ChatGPT 的出现,使生成式人工智能的发展迎来了爆发期^[4]。综合来看,目前 AIGC 有如下特点及发展趋势:

(1)基础生成算法模型的不断突破创新使 AIGC 能够创造和生成内容的类型日益多元,质量也在不断提升。除初期的 GAN 之外,近几年涌现出一大批性能稳定、各具特色的基础生成算法模型,如 2015 年 Stability AI 推出的能够根据文本描述生成高质量图像的 diffusion 模型,2017 年谷歌推出的可并行、通用性强的 Transformer 模型,2021 年初 OpenAI 推出的拥有强大特征抽取能力、能够将图像和任意文本联系起来 CLIP 模型等。尤其是 2022 年,AI 生成模型更是层出不穷,如文本-图像生成模型 Stable Diffusion、Imagen、DALL-E2,文本-3D 生成模型 Magic3D、Dreamfusion、Point-E,图像-文本生成模型 VisualGPT、Flamingo,文本-视频生成模型 Soundify、Phenaki,文本-音频生成模型 Whisper、AudioLM、Jukebox 以及文本-代码生成模型、文本-科学知识模型等^[5]。这些强大的算法生成模型使得目前 AIGC 可生成文字、图像、语音、视频、程序代码甚至 3D 物体和交互内容等多种类型的数据。

(2)大型预训练模型和迁移学习技术只需标记少量数据就可实现微调预训练,很多情况下无需开发人员从头开始训练,因而有效降低了 AI 训练成本^[5]。同时,预训练模型可以通过增加数据集和模型优化来不断扩展其应用场景和反应的准确性。理论上讲,人类可以将世界上所有可获得的文本、图片等不同类型的加入预训练数据集即可自动学习其中蕴含的知识点而无需人工介入。而预训练模型在特定任务或领域上的微调模式也使其具有更大的伸

缩和适应性,因而能够适用于多任务、多场景、多功能需求。成本的降低和适应性提升了 AIGC 的通用化能力和内容生产能力,使得同一个 AIGC 模型可以高质量地完成多种内容输出任务,成为数字内容生产的“流水线”。

(3)多模态人工智能成为 AIGC 重要的发展和应用方向。多模态技术能够将不同模态数据,如图像、文本、视频进行关联从而实现跨模态检索或不同模态数据的相互转化和生成,大大提升了 AIGC 生成内容的多样性和 AIGC 的通用化能力^[6]。如 OpenAI 推出的 DALL-E2 模型、CLIP&DALL-E 模型,谷歌的 Imagen 模型等均可将用户的自然语言转化为图片,根据用户的描述生成逼真的图像,CLIP&DALL-E 模型还可实现跨模态检索;DeepMind 的 Gato 模型能够同时执行多达 600 多项任务,可以一边视频游戏,一边聊天,同时通过机械臂堆叠积木等^[7]。现实世界中,人们很容易根据图像、文本、视频等多模态数据全面而准确地学习和理解知识,因而多模态技术的应用也使得 AIGC 向接近人类思维迈了一大步。

(4)在算法、数据和算力的综合推动下,AIGC 应用已迅速展开。目前的应用模式主要分为三大部分:一是基础生成算法模型和通用预训练大模型应用,往往由实力较强的企业推出,通过受控 API、开源等方式提供模型应用,如 Stability.ai、OpenAI;二是按需对大模型进行“专业训练”和调试形成定制化的小模型,使其成为能够为教育、金融、传媒、游戏等各行业提供服务的技术平台;三是在上述基础模型和定制化小模型的基础上进一步开发面向个人和企业用户的应用,形成最终产品,如聊天机器人、图像生成软件等。

2 AIGC 应用给图书馆带来的机遇

2.1 多模态技术助力全新的智慧检索体验

AIGC 中多模态技术的应用不仅解决了信息检索中的“异构鸿沟”问题,还可将多模态数据进行更深层的精细关联,从而实现精准的跨模态信息检索而为用户带来全新体验。自媒体时代,图像、文本、视频等多模态数据呈爆炸式增长,图书馆文献资源

载体也日益多元,通过多模态技术实现以图搜文、以文搜图,使用户更方便、快捷地找到自己想要的文献。如用户可通过输入文字找到自己想要的视频或图片,也可通过输入图像找到文字或进行混合搜索,改变了传统的用文字搜文字、用图片搜图片的单一检索模式。这不仅能有效提高文献利用率,还可为图书馆的资源建设、资源组织和知识管理提供新的思路。首先,检索方式的改变,使得图书馆馆内资源、网络资源、馆外机构资源和其他资源的整合与共享更加容易,囿于用户无法统一检索而带来的资源整合障碍有望尽快消除;其次,AIGC在文本、图像、视频等多模态数据间建立精细关联的功能也有助于建立资源结构性和非结构性数据之间的关联,从而实现文献资源的重新聚类 and 分类揭示,为用户提供更精准和多样化的信息。目前,跨模态检索技术已经在公安、传媒及医学领域展开应用,但模态数据特征表示还不够精细、不同模态特征之间细粒度对应关系的联想捕捉等问题,导致其在真实场景下的泛化能力受限^[8]。2022年以来AIGC的爆发性发展使得业界对多模态技术向着更深的领域探索,其应用领域也在逐步扩展,在图书馆行业的应用也指日可待。

2.2 AIGC应用助力智能图书馆向智慧图书馆转变

事实上,人工智能技术早已被广泛应用于图书馆,如智能检索、智能咨询、智能导航等,相较于传统图书馆服务,智能技术的应用使图书馆服务更加专业化、智能化、人性化。建立在基础算法模型、预训练模型等通用大模型基础上,并经过专业训练和调式的AIGC应用则会给图书馆服务带来更大的变革,有望推动图书馆服务由智能化向智慧化转变。以图书馆聊天机器人为例,目前的图书馆聊天机器人采用的是基于规则模型,通过规则来匹配最正确的答案,因而仅能根据预设规则回答用户有限的问题,对于用户提出的新问题或复杂查询难以处理,且回答问题生硬机械,无法满足用户的实时交互和个性化需求,达不到参考咨询服务的智慧化。而基于AIGC应用开发的聊天机器人如ChatGPT在自然语言生成任务上已达到了非常高的水平,不仅拥有语言理解和文本生成能力,还具有编写和调试计算机程序、翻译、撰写论文,进行电视剧、故事、音乐、诗歌、绘画等

的创作能力。在处理参考咨询任务方面具有更大的查询范围和精确度、较高的处理复杂问题的能力,能根据聊天上下文进行互动,与真实人类聊天场景几无差异。理论上讲,图书馆可以将世界上所有可获得的文本、图片等不同类型的数据加入预训练数据集,让聊天机器人更智慧、更接近人类思维,从而满足用户的多样化需求。除了聊天机器人,AIGC依据自然语言自动输出内容的特性,在图书馆阅读推广、特殊群体服务、用户培训等方面也会发挥重要作用。如文献的多语言翻译、盲文翻译、馆藏文献的多元化展示等;在用户培训方面,AIGC应用将创造高质量的沉浸式内容,如按需生成视频、创建三维空间导览等,可为用户带来多感官、互动性的时空体验,获得最佳培训效果。

2.3 AIGC助推图书馆元宇宙的发展

元宇宙是对虚拟世界的高度模拟,是趋近真实、虚实融合的互联网应用和社会形态。元宇宙时代的图书馆是实体空间和虚拟空间、实体馆藏和虚拟馆藏相互融合,为用户同时提供实体空间服务和虚拟空间服务的无所不在的图书馆,馆员是拥有活跃在虚拟图书馆任意空间的数字替身的超级馆员^[9]。因此,元宇宙和元宇宙时代的图书馆需要大量的数字内容来支撑,未来AIGC应用作为数字内容生产的“流水线”,可以高质量地完成多种内容和数据输出任务,因而AIGC技术将成为元宇宙重要的基础设施之一。而AIGC技术在图书馆的应用,还将为图书馆自行进行元宇宙建设提供更多的技术上的主动权,让图书馆有更多的机会按照自身的特点和用户需求进行图书馆元宇宙的规划、设计和建设,而不是一味地被外包商牵着鼻子走。同时,AIGC模型的通用化能力和工业化水平,能够使更多的图书馆用户融入进来,参与内容创建;而对于图书馆来说,则可以将复杂、效率低、成本高的事交给AIGC去做,如虚拟馆舍和家具建设、虚拟馆藏建设、用户服务场景仿真、分布于不同场景的虚拟分馆建设等,馆员则腾出更多时间进行元宇宙规划和设计,届时AIGC将为图书馆节省大量的时间成本和金钱成本。未来,AIGC在图书馆的应用意味着图书馆将从繁冗的基础性工作中解脱出来,把更多的精力放到服务设计

和创新中,这也是 AIGC 应用背景下图书馆人工作方式的整体趋势。

3 AIGC 应用给图书馆带来的挑战

对图书馆行业来说,AIGC 的应用可大大提升用户体验,为图书馆创新服务模式和内容带来无限可能,但也需要警惕 AIGC 应用给图书馆发展带来的不利影响。

3.1 用户检索与阅读习惯改变带来的挑战

信息环境的变化会直接影响用户行为和习惯,往往会对图书馆的服务模式和内容产生根本性的冲击。事实上,近 20 年来,随着互联网、搜索引擎、短视频等社交媒体的发展,相较于传统的阅读方式,人们的检索和阅读习惯已经发生了很大变化,图书馆也在不断适应这种变化的过程中做出了很大改变,搜索引擎先进的索引技术和理念、人工智能技术等也在图书馆的应用,使数字图书馆的建设日趋成熟,智慧图书馆建设也已提上日程。然而依然有一部分用户在检索信息时首选搜索引擎或互联网而非图书馆,图书馆的用户流失已成既定事实。如果说互联网、搜索引擎的出现深刻地影响了图书馆服务,那么以 ChatGPT 为代表的 AIGC 应用的出现则会对搜索引擎产生颠覆性的挑战,甚至有人预言谷歌只需要一两年时间就会被彻底颠覆^[10]。为应对危机,各大搜索引擎相继推出了自己的人工智能聊天机器人,如谷歌的 Bard、百度的文心一言,而微软直接将 ChatGPT 嵌入到其搜索引擎中,如 Bing 搜索上的“问答搜索功能”,就采用 GPT3.5 提供支持;甚至在其 Edge 浏览器上也添加了聊天和撰写功能,帮助用户解答问题和撰写邮件、简历、文档等。搜索引擎与智能聊天机器人的竞争不仅能够快速提升各自的性能,也使得人工智能技术进入人们现实生活的门槛更低,两者通过聊天场景与用户建立联系也就更容易,人们的检索和阅读习惯也就更容易发生改变,对图书馆来说,用户的流失则会进一步加重,除非图书馆能提供与两者截然不同的、对用户更有帮助的服务。

3.2 图书馆相关技术与人才弱势带来的挑战

众所周知,计算机和信息技术是推动图书馆发展的重要动力。但长期以来,图书馆对于信息技术的

应用多依赖于企业相关产品的研发或技术外包,更多地是在购买和使用服务,导致大多数图书馆的技术部门只是充当图书馆硬件设备或常用数据库和软件的购买、维护、使用和培训工作,不具备软件开发能力,只有少量顶级高校图书馆或省级公共图书馆的技术部门具备一定的技术研发能力。以图书馆智能聊天机器人为例,图书馆往往或采取直接购买,或与企业合作开发的方式,能够自主研发的图书馆少之又少。相关技术与人才弱势使得图书馆在应对计算机和信息技术高速发展,尤其是在具有里程碑意义的新技术出现时显得力不存心,这也给图书馆在瞬息万变的信息时代的生存带来了危机。ChatGPT 上线以来,很多行业以最快的速度做出了反应,图书馆界虽然也很快加入了相关讨论,但却很难在短时间内做出应对。2023 年 3 月 1 日,OpenAI 宣布以收费方式开放 ChatGPT 的应用程序接口(API),允许第三方通过 API 将 ChatGPT 集成至自己的应用程序和服务中,从而构建自己的人工智能聊天机器人,为用户提供更深层的服务。根据 OpenAI 官网展示的应用案例,已经有社交平台、学习平台、生鲜配送平台、购物程序等接入 ChatGPT API 创建了自己的聊天界面^[11]。ChatGPT API 接入图书馆最直接的应用就是将馆藏资源输入 ChatGPT 中,使其通过学习为用户提供知识服务,将图书馆智能聊天机器人提供的基本参考咨询服务扩展到知识咨询与服务。然而,如果图书馆行动不够快速,各大数据库商可能会抢占先机,图书馆将再次沦为购买和作为信息中介的被动局面,用户的流失将不可避免。

3.3 AIGC 自身的局限性给图书馆带来的风险

从安全的角度来看,首先存在版权风险。以 ChatGPT 为例,其用于训练的大型文本数据语料库的数据主要来自互联网,如果这些数据中存在侵犯版权的情况,则其生成内容也可能面临知识产权纠纷,正如 OpenAI 公司自己所承认的那样,即使 ChatGPT 在预训练时已经在尽量规避版权风险,但不能保证所有数据都经过了原始作者的许可^[12]。由于 AIGC 的预训练模型数据量庞大,图书馆在应用时无论是否输入自有的数据,也会面临版权风险;其次是存在用户隐私、敏感和机密信息泄露的风险。AIGC 依靠

收集和存储大量用户数据来提高性能,当用户输入个人数据或其他敏感信息时,AIGC可能会将其纳入语料库从而产生泄露风险;第三是存在产生偏见和歧视性结果的风险。一是用于训练人工智能的语料库中如果含有偏见或歧视性的数据,模型可能会不加区别地进行处理,进而把这种偏见和歧视反映到输出内容上。二是如果算法潜藏偏见和歧视,也会导致偏见或歧视性结果。虽然目前AIGC应用工具都在设法规避“毒性”语料,但其算法被认为是“黑箱”操作,透明度缺失,仍然存在输出偏见或歧视性结果的风险^[13];最后是助长学术造假的风险。以ChatGPT为例,由于其具有翻译、撰写论文、编写代码等创作功能,已有国外大学生利用其发表学术论文、撰写课程作业,引发了学术界和教育界的担忧^[14]。在线课程提供商Study.com对1000多名学生进行的调查显示,89%以上的学生使用ChatGPT来帮助其完成家庭作业^[15]。已有包括《Nature》在内的多家知名学术期刊明确表示不接受由ChatGPT生成的论文投稿^[16]。但如何区分人类写作文本和人工智能合成文本却困难重重,针对这种情况,2023年2月,OpenAI公司推出AI Text Classifier的文件检测器,但其准确性也不容乐观,仅为26%^[17]。

4 图书馆应对AIGC的策略与途径

4.1 从聊天机器人入手优化和扩展其服务范围

聊天机器人是目前很多图书馆提供在线参考咨询服务的方式。但现有的图书馆聊天机器人主要被用来回答最基本的有关图书馆服务的问题,复杂的问题还需要馆员来解答。参考ChatGPT理念与模型训练方法,优化和丰富训练语料库,创建图书馆特色的人工智能聊天机器人,使其既能回答最基本的和较为复杂的参考咨询问题,同时也能真正扮演聊天对象,为用户提供范围更广的聊天服务,改善用户体验,这应该说是目前较为简单的图书馆与AIGC技术对接的方式。一项对60多个品牌的客服聊天机器人的研究发现,约40%的用户的请求是感性的,即用户可能只是需要向聊天机器人宣泄情感,而不是寻求具体的产品或服务信息^[18]。加州大学欧文分校图书馆对其智能聊天机器人ANTswers的用户咨询问

题进行统计后发现,与图书馆服务无关的问题数量之多出乎了馆员的意料。由于用户是匿名咨询,他们更喜欢把ANTswers当作人来对待,询问它怎么样以及生活情况等。也就是说,用户可能只是想找个人聊聊天,而聊天机器人充当了这个角色。统计结果使欧文分校图书馆放弃了将这些与图书馆服务无关的聊天数据从语料库中删除的决定,因为这样可以为用户带来更好的体验^[19]。因此,将图书馆聊天机器人语料库向通用型靠拢,增加更多的通用型内容对其进行训练,使其能够回答用户除图书馆服务外更为广泛的问题,逐步深化大规模语言模型在图书馆参考咨询服务中的应用。

4.2 尝试利用开源模型训练图书馆现有文献数据

目前,大多数人工智能开源模型基于Transformer模型,并具有上手快、易操作等特点,适合入门级开发人员使用,对技术力量薄弱的图书馆来说,可以快速自主开展适应图书馆自身业务需要的人工智能数据训练,为AIGC技术在图书馆的成熟应用打下基础。国外已经有图书馆开展了类似的工作。如瑞典国家图书馆就采用开源自然语言处理平台Hugging Face的Transformer模型对几个世纪以来的馆藏资源进行人工智能训练。在过去的500年里,瑞典国家图书馆收集了几乎所有用瑞典语出版的各种资料,这些资料包括书籍、报纸、广播和电视、互联网内容、博士学位论文、明信片、菜单和电子游戏等,资源的多样化非常适应进行人工智能训练。2019年,瑞典国家图书馆成立KBLab,开始在1TB的瑞典语文本上训练其模型,为提高模型性能,还尝试向数据集添加荷兰语、德语和挪威语的内容。随着馆藏资源的不断数字化,每个月都会增加约50TB训练数据,小型工作站已不能满足要求,瑞典国家图书馆随即采用了英伟达的超级计算机DGX,并增加了将声音转录为文本的人工智能工具,使图书馆庞大的广播电视馆藏能够转录为文本,利用多模态数据进行人文研究。随着训练工作的不断推进,瑞典国家图书馆现在已开始开发生成文本模型,并正在研究可以处理视频和能够自动描述视频内容的模型^[20]。图书馆开展AIGC模型应用是一个循序渐进、逐步推进的过程,利用现有馆藏积极进行初期的应用不失为一种跟进

人工智能技术、改善用户知识服务的良好途径。

4.3 提前进行 AIGC 应用的风险规避和防范

在安全问题方面,版权和用户隐私问题事实上一直是数字图书馆建设中不可避免的话题,在 AIGC 技术应用背景下,这一问题更加复杂,图书馆要做的工作相对更多。在用户隐私保护方面,要通过培训和教育提高馆员和用户的安全意识。在很多情况下,往往是用户在不经意间将自己的隐私信息泄露出去的,如在使用聊天机器人时,输入了自己的隐私数据而毫不知情。因此,需要对用户进行必要的信息安全教育,并不断完善用户隐私政策。馆员可以说是用户隐私的“守护者”,应对他们进行数字安全技能培训以及职业道德教育,增强对用户隐私的保护意识。在版权方面,首先要保证图书馆用于人工智能训练的数据的合法性,除图书馆自有数据外,其他数据要确保获得许可和授权或符合著作权法的其他规定;其次,有关 AI 生成内容的版权问题目前还存在争议。OpenAI 公司在其用户协议中规定,用户拥有基于输入内容而生成的输出内容的所有权,并对内容负法律责任,ChatGPT 则拥有这些内容的使用权^[21]。而根据我国《著作权法》的规定,“著作权属于作者。创作作品的自然人是作者。也就是说目前我国的法律仅保护自然人创作的作品,对于由人工智能和自然人共同创作的作品还没有明确的规定。图书馆作为信息文献中心,有责任也有义务关注并对相关法律法规的制定、修订提出自己的建议,这需要图书馆在服务实

践中不断地积累相关的用户使用情况和意愿,并将之纳入业务研究范畴,关注涉及人工智能生成内容的版权纠纷案件。另外,自 2023 年 1 月 10 日起,我国正式实施《互联网信息服务深度合成管理规定》,对深度合成服务的提供者、使用者、支持者所要承担的法律义务进行了规定^[22]。2023 年 4 月 11 日,国家互联网信息办公室发布《生成式人工智能服务管理办法(征求意见稿)》,对生成式人工智能产品或服务应当遵循的法律法规提出了要求^[23]。AIGC 若在图书馆使用,则作为服务提供者的图书馆需要承担相应的法律责任,这也是图书馆行业应该提前规划和关注的方面。

5 结语

2023 年 3 月 22 日,非营利组织生命未来研究所(Future of Life)发布了一封《暂停大型人工智能研究》的公开信,呼吁所有人工智能实验室立即暂停比 GPT-4 更强大的人工智能系统训练 6 个月,共同制定和实施人工智能设计和开发共享安全协议,并由独立的外部专家进行严格的审查和监督,通过建设 AI 治理系统防止人工智能对人类造成的危险^[24]。每一项新技术的应用都会带来潜在的风险,虽然 AIGC 的潜力还存在很多不确定性,但毫无疑问它对图书馆行业必将产生深刻的影响。未雨绸缪,提前谋划;持续跟踪,积极实践,是当前图书馆应对这一危机和影响的必由之路。

参考文献:

- [1] Poulomi Chatterjee.What Excites Yoshua Bengio about the Future of Generative AI[EB/OL].[2023-03-26].<https://analytic-sindicamag.com/what-excites-yoshua-bengio-about-the-future-of-generative-ai/>.
- [2] 魏富强,古兰拜尔·吐尔洪,买日旦·吾守尔.生成对抗网络及其应用研究综述[J].计算机工程与应用,2021,57(19):18-31.
- [3] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners[J]. Advances in neural information processing systems, 2020, 33:1877-1901.
- [4] Gozalo-Brizuela R, Garrido-Merchan E C. ChatGPT is not all you need. A State of the Art Review of large Generative AI models[J]. arXiv preprint arXiv:2301.04655, 2023.
- [5] 唐亚汇.大规模预训练模型对 AI 应用与普及的影响[J].上海信息化,2021(9):52-55.
- [6] 赵朝阳,朱贵波,王金桥.ChatGPT 给语言大模型带来的启示和多模态大模型新的发展思路[J].数据分析与知识发现,2023,7(3):26-35.
- [7] Reed S, Zolna K, Parisotto E, et al. A generalist agent[J]. arXiv preprint arXiv:2205.06175, 2022.

- [8] 刘颖,郭莹莹,房杰,等.深度学习跨模态图文检索研究综述[J].计算机科学与探索,2022,16(3):489-511.
- [9] 董甜甜,卜险峰.图书馆迈向元宇宙的先行实践、未来场景及发展策略[J].图书与情报,2022(5):92-97.
- [10] 齐健,陈伊凡.ChatGPT 引爆新一轮科技军备赛[EB/OL].[2023-02-26].<https://new.qq.com/rain/a/20230212A01Z8L00>.
- [11] OpenAI.Customer stories[EB/OL].[2023-02-26].<https://openai.com/customer-stories>.
- [12] 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[J].中国出版,2023(5):16-21.
- [13] Goldstein J A,Sastry G,Musser M,et al.Generative Language Models and Automated Influence Operations:Emerging Threats and Potential Mitigations[J].arXiv preprint arXiv:2301.04246,2023.
- [14] Thorp H H.ChatGPT is fun,but not an author[J].Science,2023,379(6630):313-313.
- [15] Productive Teaching Tool or Innovative Cheating? [EB/OL].[2023-03-30].<https://study.com/resources/perceptions-of-chatgpt-in-schools>.
- [16] Editorials N.Tools such as ChatGPT threaten transparent science;here are our ground rules for their use[J].Nature,2023(613):612.
- [17] Kirchner J H,Ahmad L,Aaronson S,et al.New AI classifier for indicating AI-written text[J].OpenAI,2023.
- [18] Xu A,Liu Z,Guo Y,et al.A new chatbot for customer service on social media[C].In Proceedings of the 2017 CHI conference on human factors in computing systems,2017:3506-3510.
- [19] Kane D.Analyzing an interactive chatbot and its impact on academic reference services[J/OL].[2023-03-22].[https://alair.ala.org/bitstream/handle/11213/17624/Analyzing anInteractiveChatbotanditsIm pac tonAcademicReferenceS ervices.pdf?sequence=1](https://alair.ala.org/bitstream/handle/11213/17624/Analyzing%20anInteractiveChatbotanditsImpactonAcademicReferenceServices.pdf?sequence=1).
- [20] ISHA SALIAN[EB/OL].[2023-03-30].https://blogs.nvidia.com/blog/2023/01/23/sweden-library-ai-open-source/?utm_campaign=Learning%20Posts&utm_content=235868723&utm_medium=social&utm_source=facebook&hss_channel=fbp-1556374471294462.
- [21] Lemley M A.Terms of use[J].Minn L Rev,2006,91:459.
- [22] 国家互联网信息办公室,工业和信息化部,公安部.互联网信息服务深度合成管理规定[EB/OL].[2023-03-26].http://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm.
- [23] 国家互联网信息办公室.国家互联网信息办公室关于《生成式人工智能服务管理办法(征求意见稿)》公开征求意见的通知[EB/OL].[2023-04-11].http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm.
- [24] Future of Life Institute.Pause Giant AI Experiments:An Open Letter[EB/OL].[2023-03-30].<https://futureoflife.org/open-letter/pause-giant-ai-experiments/?continueFlag=d19c796896f3776241c518f8747afa68>.

作者简介 李颖婷,女,内蒙古呼伦贝尔市图书馆主任编辑,研究方向:图书馆管理、资源建设、图书馆服务、智慧图书馆。