

随着技术研发和迭代的步伐逐渐加快,AI正被深入应用到人们的工作中,不仅渗透到千行百业,还在重塑劳动力市场。但这并不意味着AI将完全取代人类。AI需要与人类协同工作,实现“人机合作”,才能更好地发挥其作用。

赋能还是替代

AI“入职”百业生变

本报记者 郝一萍 韩启

数字人“夏语冰”。

AI歌手唱出“好声音”
加入呼吸声让它更像真人

孙玉镜是国内早期利用AI技术辅助创作的音乐人之一。作为幻想动漫音乐团队创始人、艺术总监,他在2013年就开始使用歌声合成软件。不过在他看来,那时,这类软件主要使用的是采样拼接技术,无论技术还是效果都与如今的软件差异较大。

AI歌手是结合深度学习技术,通过大量语音和演唱素材训练出来的虚拟歌手。这些AI歌手能迅速“读懂”乐谱,并像人类歌手一样,自然地“演唱”出来。

如今,市面上的歌声合成软件能提供各种音色的AI歌手。音乐制作人根据音乐创作的需求,在这些软件中挑选相匹配的声音,无论是戏腔还是气泡音,甚至摇滚嗓,都能从中找到。

孙玉镜和他的团队目前常用的有3款歌声合成软件:X Studio、ACE Studio和Synthesizer V。每款软件有十几位甚至几十位AI歌手。其中,前两个为免费软件,而最后一个为付费软件。

记者登录X Studio看到,操作页面左侧列有12名AI歌手(也就是12个声库),不仅按照男女歌手进行分类,每个名字旁还注明其演唱风格和声音特点,比如名为“崔璨”的AI歌手为“摇滚”“烟嗓”,AI歌手“小鱼豆腐”为“流行”“娃娃音”等。

“我们常用的是‘崔璨’。因为它的摇滚腔比较特别,市场上具有同类音色的AI歌手很少,演唱出来的声音我们自己听着都觉得挺带劲儿。”据孙玉镜介绍,在使用AI歌手之前,制作人创作出来的歌曲,一般由自己试唱。而制作人虽然在音乐理论和制作方面非常专业,但是对于演唱,可能远达不到歌手的水平。因此,音乐制作团队在词曲创作和制作歌曲小样时常常会用到AI歌手。

“在使用过程中,我们先输入一段旋律和对应的歌词,AI歌手会生成一段歌声,但可能不会一下子就能达到令人满意的程度。我们会再调整一些参数,包括颤音的幅度等,使演唱更自然。另外,我们还会加入一些呼吸声,让这段演唱更加符合大众的听觉习惯。比如,真人歌手唱到这句时该换气了,我们就把呼吸声加到这里。”他说,“AI歌手有助于快速明确这个旋律和歌词是不是我们想要的效果。这对我们前期词曲创作和编曲工作都有很大的帮助。”

令人惊讶的是,日前,X Studio中的12名AI歌手还组成了虚拟歌手厂牌WOWAIDO,并在1个月内接连发布了3张专辑共9首歌,引来数千名歌迷的评论。不少歌迷对它们产生好奇。“你们多大啦?”“你是AI还是真人?”甚至有歌迷感叹:“你能相信这是人工智能唱的吗?”对此,AI歌手还会在这些评论下一一回复。公司还公布了这些AI歌手的个人资料,包括星座、身高等信息。

尽管AI歌手的声线已与真人十分相似,但谈及将AI歌手用于商业项目时,孙玉镜还是有一些顾虑:“在商用项目上,我们希望版权能更清晰一些。目前,AI歌手作品的版权界定仍有一些不确定性。”

拥有孙燕姿歌曲版权的环球音乐对此感触颇深。“在词曲创作方面,一般没有问题,因为很多国家的法律规定版权所有者必须是人,机器是不能得到版权的。不过在唱片方面就有问题了,因为版权所有者可以是公司。”环球音乐亚洲区公共政策高级副总裁洪伟典表示,目前,环球音乐已经拥有了3项人工智能专利,但应用场景并非音乐创作,而是帮助艺术家扩大受众群并更好地与“粉丝”互动。

AI绘出纺织行业新蓝图
图案设计工时减少65%

早在2018年,曹潇文和她所在的团队就将AI绘图应用到了纺织设计领域。作为中国纺织信息中心时尚智能部创意总监,她将这一实践称为纺织业的“意外惊喜”。

2018年10月,中国纺织信息中心推出了初代人工智能设计模型,并发布全球首款由人工智能设计的纺织面料图案。该模型在前期训练时,至少要使用3000张高质量图片,生成的图案相继被用于服装、地毯、箱包等产品。

“我们当时找了不同类型的企业去测试这项技术是不是能做产业化落地,但发现还是存在一些问题。比如,纺织产业在图案的应用上有特殊的生产标准。”曹潇文对此进一步解释说,“我们在电脑屏幕上看到的很鲜活、很美的图案,在实际的生产中并不一定能被制作出来。因为图案的生产环节对图片的精度、颜色数量、图案循环方式等都有严格的要求。”

随着对AI绘图探索的推进,2020年10月,人工智能图案创意平台——DPI SPACE上线。该平台通过AI生成大量符合流行趋势的纺织图案供用户使用。曹潇文告诉记者,与上一代绘图模型相比,该系统具有逻辑思维和情感识别能力,学习能力更强,只需30张高



AI绘图生成的服饰花色。



AI绘图生成的山水画。

本版照片由受访者提供

质量图片即可完成模型训练。

记者登录DPI SPACE发现,该平台不仅提供传统纹样、植物花卉、艺术绘画等10个类别的图案下载,还能进行图案流行趋势预测。以2024春夏系列为例,该平台预测,趣味外观、创意扎染、古典繁花等8类图案会流行。其中,一些图案显示“已售出”,其他用户将不能下载。此外,该平台还设置了智能生成板块,用户输入提示词,搭配平台推荐的趋势风格和颜色特征,即可生成相应图片。记者在体验该功能时,输入了“水墨画”“树枝”“复古”等提示词,大概经过14秒,即生成了一张纺织品图案设计图稿。

“完成一张设计图稿,一名经验丰富的设计师至少需要5个小时,而通过AI绘图最多只要60秒。在企业实践过程中,AI绘图突破了人工设计花型周期长和花型单调的限制;提高了设计效率,图案设计工时减少65%;降低了研发成本,设计打样费用减少了60%。”曹潇文告诉记者,在纺织企业中,一名花型设计师一天至少需要产出两三张设计图。很多设计师缺少构思时间,创造力难以得到发挥。而利用AI绘图可大大减少制图时间,提高设计效率,让设计师有更多时间去收集素材、挖掘市场需求。

然而,这并不意味着AI绘图可以完全替代设计师的工作。“AI绘图是用来辅助设计师的。因为在使用AI绘图时,用户的脑海中要先有一个构想,包括图案风格、配色、题材等。也就是说,创意和灵感依然是核心。之后,如何将这些需求转化为相应的提示词也是重要一环。使用不同提示词生成的图案会相差很多。”曹潇文预测,“未来,设计师的工作模式或将动手画图转变为思考提示词,用什么样的提示词才能把设计图‘画’得更漂亮。这也是一个设计师需要适应AI绘图的过程。”

数字人“入职”千行百业
10分钟即可培训出一名“员工”

“与其说我们是科技公司,不如说是‘人才市场’。”这是在北京红棉小冰科技有限公司(简称小冰公司)的办公室里接受记者采访时,前微软(亚洲)互联网工程院副院长、小冰公司首席执行官李笛说的一句话。

小冰公司前身为微软(亚洲)互联网工程院人工智能小冰团队,2020年被拆分为独立实体。AI Being(AI数字人)的理念就是2019年由小冰团队率先提出。

之所以称小冰公司是“人才市场”,是因为过去3年,从这间办公室“走”出去的数字人成功“入职”了多家企业,担任的职位涵盖主播、秘书、分析师、设计师、催收专员、音乐制作人、金融风控师等。包括画家“夏语冰”、歌手“何畅”、北京冬奥会裁判与教练“观君”、时

尚设计师“西湖一号”等“知名人士”也均出自于此。

“以前,我们制作数字人,从最早期的人物设定到最终推向市场,可能要几个月才能完成。而现在通过AI大模型,10分钟即可培训出一名‘员工’。”李笛满脸自豪,“这就是技术的进步。”

值得一提的是,基于AI大模型生成的数字人还具有很强的学习能力和逻辑推理能力。“数字人的能力一般会垂直。他们所学的‘专业’会根据客户需求而定。企业不需要全能的数字人,因为功能越多,成本越高,所以我们会根据客户的岗位要求做定向训练。”李笛告诉记者,“我们推出的数字人已‘入职’万科、遥望科技、红杉中国等知名企业、机构。他们大多有与真人无异的样貌和神态,甚至有些数字人在短时间内都不会被发现不是真人。”

今年5月,小冰公司启动了“克隆人计划”。“半藏森林”等首批网红明星“克隆人”已在X Eva上线。用户在这款App上除了能与网红明星“克隆人”实时聊天之外,还能“克隆”自己或者其他人。记者尝试“克隆”自己时发现,生成“克隆人”需要填写姓名、性别、人设等。此外,提供恰当的对话示例,也能让“克隆人”更符合人设。“克隆人”还可以关联微博等社交账号,以便于更好地生成“克隆人”的性格。总之,填写的内容越详细,“克隆人”的形象就越具体。该App页面类似于微信,“克隆人”生成后,用户可以与其交流,也能看到他们生成的朋友圈。目前,软件在内测阶段,加V的“克隆人”为官方推出,所有人可见;用户自主“克隆”的仅自己可见。

李笛表示,数字人可以分为两类,一类是纯虚构的,一类是对标真人的。“克隆人”属于后者,其背后的数字孪生技术已逐步完善。他说:“生成对标真人的数字人时,我们一般需要这个人20分钟至2个小时的日常录像素材。这段素材能显示出这个人的语速、语气、动作等。以此训练出来的数字人属于学习者,其背后的数字孪生技术已逐步完善。他说:“生成对标真人的数字人时,我们一般需要这个人20分钟至2个小时的日常录像素材。这段素材能显示出这个人的语速、语气、动作等。对于那些学习素材太少的数字人,可能你一眼看过去觉得挺像,但一说话就露馅了。”

随着技术不断完善,数字人已迈向更多的领域,一些企业也瞄准时机抢占“风口”。世优科技推出的数字人“慕兰”能够实现对弹幕提出的不同问题进行回答,实现较好的交互效果;魔法科技推出的数字人“翎”已应用至天猫、百雀羚、特斯拉等代言活动中;快手助力蒙牛打造的数字人“奶思”通过蒙牛牛奶旗舰店进行直播带货……

根据艾媒咨询数据,2022年我国数字人核心市场规模达到120.8亿元,同比增长94.2%。2023年诸多AI大模型相继发布,有望赋能数字人产业,实现多环节降本增效,大幅提升应用端交互能力。艾媒咨询预测,到2025年,数字人行业核心市场规模有望达到480.6亿元。

产业观察

从AlphaGo到ChatGPT 人工智能实现了哪些新跨越

从文本、图像、到音频、视频乃至跨模态信息,生成式人工智能正让内容生产呈现爆发式增长。不久前,至顶智库发布《2023年全球生成式AI产业研究报告》称,生成式人工智能是在专业生成内容(PGC)、用户生成内容(UGC)之后,利用人工智能技术自动生成内容的新生产方式。在该机构看来,自ChatGPT出现以来,生成式人工智能广受追捧,诸多科技类企业纷纷推出生成式人工智能大模型、底层基础设施及服务。

“文本生成是目前应用最多的场景,很多人工智能大模型都侧重于此。”至顶智库执行主任兼首席分析师孙硕告诉记者,目前,人工智能发展已进入到大模型时代,而这需要高性能人工智能芯片提供算力支持。

伴随需求扩张,人工智能芯片市场未来将呈现高速增长态势。艾瑞咨询发布的数据显示,去年,我国人工智能芯片市场规模不足400亿元。但这家机构预测,到2027年,该市场规模将超过2100亿元。

在ChatGPT之前,人工智能引发外界巨大关注还是在2016年。当年3月,AlphaGo大胜围棋世界冠军李世石,让这场“人机大战”举世瞩目。据孙硕介绍,AlphaGo所使用的人工智能技术之一便是深度学习,而这也是机器学习技术之一。

深度学习是通过构建和训练多层神经网络模型来模拟人类大脑的神经网络结构,从而实现大量复杂数据自动学习和特征提取。通过深度学习,AlphaGo在围棋领域的能力已达到了人类的顶级水准。

与AlphaGo相比,ChatGPT综合采用了自然语言处理、机器学习等人工智能技术。比如,借助预训练语言模型,ChatGPT可以通过大规模语料库的预训练来学习自然语言的语法、语义和上下文关系,从而实现高质量的文本生成和对话交互功能等。此外,通过机器学习技术中的强化学习,ChatGPT可根据人机交互的数据自动调整其对话方式,提高交互质量。据业内人士介绍,ChatGPT较AlphaGo的应用场景更广,且能处理的数据类型更为复杂。

从基础层,到技术层,再到应用层,人工智能的产业链十分庞大。前文提到的人工智能芯片属于产业链中的基础层。在技术层方面,人工智能包括机器学习、知识图谱、计算机视觉、自然语言处理等细分领域。至于应用层,则是人工智能为各种行业提供的解决方案、硬件产品和软件产品等。

基于不同的口径统计,各类机构给出的人工智能产业的规模数据差异较大。比如,国内一家知名咨询机构测算,去年,我国人工智能产业规模为2000亿元左右。但据工信部相关负责人近期透露,目前我国人工智能核心产业规模已达5000亿元。

在人工智能产业中,计算机视觉不仅是热门赛道之一,还是商业化最快的细分领域。所谓计算机视觉,是指利用一个代替人眼的图像传感器获取物体的图像,然后将图像转换成数字图像,并利用计算机模拟人的判别准则去理解和识别图像,达到分析图像和作出结论的一种技术。据艾瑞咨询测算,去年,计算机视觉在我国人工智能产业中的规模占比超过40%。

大模型揭秘

训练耗时3个月 大模型“生成记”

以ChatGPT为代表的生成式人工智能,之所以能实现复杂语境下语义理解、答案生成、逻辑推理等,与背后的大模型密切相关。大模型是一种复杂的神经网络模型,具有强大的计算能力和学习能力。一般而言,大模型往往拥有上亿个参数。从全球来看,一些大模型的参数量已达到千亿级。

在ChatGPT问世后,我国掀起了大模型投资热潮。作为一家专注于自然语言处理与知识图谱的人工智能公司,天津汇智星源信息技术有限公司(简称汇智星源)今年自主研发了慧知行业大模型。截至目前,该大模型已应用于机器人、智能客服、智能桌面终端、智能交互屏等领域。

汇智星源首席算法科学家、人工智能高级工程师宋思惟告诉记者,慧知行业大模型的参数量为百亿级。“人工智能大模型的参数包括模型规模参数、学习率参数、正则化参数、批次大小参数等多种类型。至于参数值,一般通过优化算法来确定。比如,通过调整正则化参数,可以平衡大模型的拟合能力与泛化能力。”她说。

通俗来说,拟合能力决定了大模型计算结果和实际结果的差异程度,而泛化能力则可反映其是否具备“举一反三”的特征。据了解,为了提升大模型输出内容的准确性,人工智能企业往往从验证数据来源、数据清洗、数据标注、模型评估、异常检测等多个层面“下功夫”。

如何打造一个“大模型”?宋思惟对记者表示,这通常需要经过数据准备、模型设计、参数初始化、模型训练、超参数调优等诸多环节。以慧知行业大模型为例,其仅训练过程就耗时3个月。“大模型的训练周期受多重因素影响,比如模型规模、算力资源、数据集大小等。”她说。

与通用大模型相比,宋思惟认为,面向垂直领域的行业大模型的研发与训练会面临更大挑战。比如,行业大模型训练所需的数据更难获得。对于数据安全,其也需要特别考虑。此外,在行业大模型研发中,模型预测与决策过程需要具备可解释性和可控性,而这对模型的设计和解释能力提出了更高要求。

律师观点

用户难以判断是否侵权 AI服务提供者应自觉注意服务合法性

随着大模型概念不断升温,AI歌手、AI绘图、数字人等产品相继涌出,引发了一些关注和争议。这些产品改善人们的生活和工作模式的同时,也存在一些法律风险。

如今,涉及AI侵权案件的数量呈上升趋势,且类型呈多元化发展。此前,在一个人工智能侵权案件中,某软件就因擅自使用自然人形象创设虚拟人物而侵害了被告的人格权。在本案中,被告为服务(即AI算法模型)提供者;原告为该软件的用户,也就是数据提供者。

在中国政法大学知识产权研究中心研究员、北京时代华地律师事务所高级合伙人徐新明看来,每个人的法律素养和道德标准不同,并不是所有用户都能判断出软件是否涉嫌侵权。而被告利用技术,提供有极大可能侵犯他人人格权服务的行为应受谴责。对此,他呼吁,AI服务的提供者应当提高自己的法律意识,自觉注意服务的合法性,在迎合市场的同时,合法经营。

此外,他表示,AI产品开发者也应警惕信息处理、利用AI进行不正当竞争、AI专利侵权及商业秘密侵权等风险。“人工智能大模型的训练,势必会使用大量信息。开发者在进行数据收集、处理时,不可避免地会涉及个人敏感信息。这就需要事先取得个人信息主体的明示同意,否则将会涉嫌侵权。而在AI开发过程中,并非只涉及计算机软件著作权。同其他技术开发一样,在AI开发的过程中,也可能发生专利侵权以及商业秘密侵权等。”徐新明说。