

2023年,ChatGPT等大语言模型的发展速度日新月异来形容毫不为过。相比过去,AI对于我们而言已不再是遥远的“科技话题”,而是切实进入了每个人的生活,无法回避。有人担忧AI会让人类失业,也有人乐观地认为未来淘汰的是不会使用AI的人。

由此,本报记者采访了几位在他们的学习和工作中运用AI的“先行者”。2017年,科幻作家陈楸帆便开始尝试与AI合作撰写科幻小说,并成为“打败莫言”的人;工程师关鹏在学生时代苦于传统编辑器的不便,“斗争”半年后,运用AI研发出一款新的编辑器,解决了一系列

bug;心理咨询师王曦尝试用AI开发心理剧帮助人们更好地缓解心理问题。面对AI带来的不确定性,他们选择积极拥抱,看到了新的希望和机会。就像陈楸帆所说,未知是最迷人的,而AI帮助每个人挖掘和发挥自我的独特性。

## 陈楸帆 我和AI合著的小说『打败』了莫言

■本报记者 韩扬眉

登顶“AI文学榜”榜单!2019年初,在《收获》杂志和《思南文学选刊》联合举办的一次评选中,科幻作家陈楸帆与AI合著的短篇小说《出神状态》排名第一,以0.0001分的优势战胜“榜二”——诺贝尔文学奖得主莫言发表在《十月》杂志上的小说《等待摩西》。

陈楸帆笑称,这篇科幻小说让他成为“打败”莫言的人。有意思的是,不但这次的参赛作品《出神状态》是陈楸帆和AI合著的,这次比赛的评委也是AI。

自2022年底,以ChatGPT为代表的大语言模型横空出世后,用AI帮助文学创作已非新鲜事,但陈楸帆的尝试更早,始于7年前。

### 未来的机器将更深入地卷入人类创作

2017年,陈楸帆受中信出版集团邀约,写一本名为《人生算法》的短篇小说集,探讨人与AI的故事。他突然灵光一现:为什么不不用AI参与创作呢?于是,曾在谷歌任过职的陈楸帆找到他的工程师朋友王咏刚,开始了一场写作实验。

这年谷歌发布了注重注意力机制与Transformer算法——类似今天ChatGPT基础架构。在此基础上,两人在各大技术论坛上寻找开源代码,开发模型,然后用陈楸帆创作的小说和网络上的素材进行训练,被称为“陈楸帆2.0”。

“当时开发的模型非常初级。”陈楸帆回忆道,模型创作出的文字连简单的人称都搞错,人物关系里谁是谁、彼此什么联系也不清楚,“我创造一个情景,去配合AI完成一个故事,实际上我和AI的主次关系是颠倒的”。

这次尝试虽然不那么成功,但从那时起,陈楸帆就开始关注AI大模型在写作上的应用。

陈楸帆让机器学习了他的上百万字作品后,AI程序“陈楸帆2.0”可以通过输入关键词和主题,每次自动生成大约几十到一百字的段落。于是,陈楸帆与它共同创作了作品《出神状态》。

《出神状态》的灵感来源是智利作家罗贝托·波拉尼奥的短篇小说集《地球上最后的夜晚》,主题是大瘟疫。故事记录了主人公去上海图书馆还书途中的见闻、思想,探讨了通信中断与紧急状态下的人与社会的关系。他将自己创作的段落与AI创作的段落分别做了标注。

“没想到《出神状态》战胜了《等待摩西》,真是比科幻还要科幻的惊喜。这也让我醒悟到,未来的机器将更深入地卷入人类创作中,未来的内容版图也会变得更加复杂、暧昧而有趣。”陈楸帆告诉《中国科学报》。

2020年,陈楸帆再次尝试,借助一个新开发的科幻小说大语言模型,用包括刘慈欣、韩松等作家不同风格的科幻小说训练,还邀请了其他作家参与尝试用模型创作小说。

在陈楸帆看来,2020年的大语言模型相比2017年有了质的飞跃,“有的作家认为AI写出来的东西达到了人类的平均水平,有时

还会有奇思妙想的语句”,比如AI说“北京像一个黑洞,因为你在路上,你的时间都不知道花到哪儿去了”。陈楸帆说这种让人类深有同感的话语,从AI的“嘴”里说出来非常具有讽刺性。

### AI有可能帮助我们“跳出惯性”

ChatGPT诞生后,陈楸帆花费大量时间和精力探索如何与ChatGPT更好合作,每天不断地和它对话。

“我并不想让它非常了解我,我期待看到我所没有想到的。”陈楸帆说,跟ChatGPT合作肯定不是我们原来想象的简单粗暴的“复制粘贴”过程,而是一个复杂的、人和机器共生的、共同写作的过程。

陈楸帆用“头脑风暴”形容他与AI的合作方式。他的经验是,不要预设太多,甚至越模糊越好,让AI提供多个方向,而且每个方向相差越远越好,再从中间筛选最有感觉的方向,并进一步细化。

“人总是容易被思维惯性束缚,AI有可能帮助我们跳出惯性。”陈楸帆在创作过程中,当直觉告诉他哪里需要AI的参与,他便“即兴”地由AI创作,“我不会修改AI创作的内容,让其保持原有状态”。

最近他正在写一个儿童科幻,内容是一个非常懂AI的城市女孩去一个原始村落,遇到一个少数民族女孩的故事。他用ChatGPT描绘了未来AI城市的样子、女孩一天的生活场景,写出来的内容非常像作家叶永烈的科幻小说《小灵通漫游未来》。

于是,陈楸帆把文本再回输进去,要求ChatGPT增加文学性、哲学思考、生动的比喻,ChatGPT就根据这些要求修改。陈楸帆甚至要求ChatGPT用美国小说家科马克·麦卡锡的笔触写一个海啸的场景,虽然出来的文字不一定都能用,但给了他一些“感觉”。

在陈楸帆看来,和ChatGPT一起写作就像一个开放的写作实验,与它不断对话的过程中逐渐明确场景,再加上设计的人物,借由ChatGPT,人物和场景之间发生的故事就展开了。

陈楸帆说,让AI参与创作,其实更多的是寻找可能性,而非强化我们已有的东西,“扔进去几个关键词,让ChatGPT自己生成内容,反而会激发更多的灵感。创意就是发现未知,而未知是最迷人的”。

### AI时代,每个人都应挖掘独特性

面对AI这样的新事物,陈楸帆的第一反应是拥抱和探索,他很喜欢学习的过程,乐于思考这个时代中自己的优势和局限性在哪儿,努力思考如何让自己的优势变得更强大,从而不可替代。

陈楸帆认为,人与AI写作没有特别大的差别,都是信息处理转化的过程。区别在于,人有五官、五感,有许多非常微妙的感觉,人类有很多自己都不知道的思想意识。但AI目前都没有,AI本质上是一个计算性的模型,“AI甚至有些拙,非常笨拙”。

陈楸帆还通过人与AI创作的不同,对自己的写作作了反思,他看到人类的边界、惯性在哪里,思考原创性的边界、人类的主观能动性在哪里,包括知识产权如何界定、怎么尊重被采集数据作者等等。

“AI的出现,鼓励了人挖掘自我的独特性。这个时代,每个人都应该向内挖掘自己的独特性,不可替代性,对于青少年来说更是如此,所以教育不在于规划,而在于找到每一个人的独特性。”陈楸帆说。

系统中,当系统作出错误决策或者造成伤害时,责任归属可能会变得模糊。

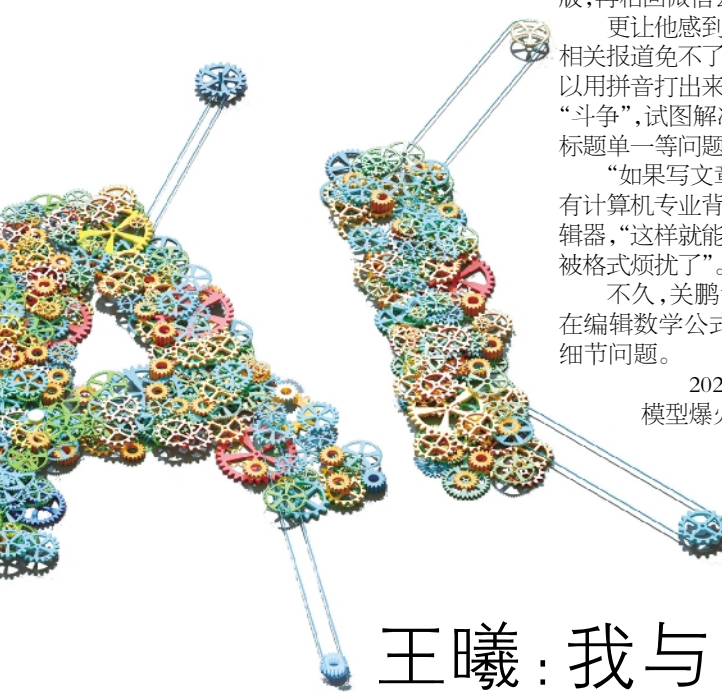
文章还提到了“算法歧视问题”。在生命医学研究中,研究者需要确保研究方法公平,不会对任何群体产生歧视。然而,人工智能技术是一项赋能技术,它能够增强完成任务的速度和效率、扩展原技术的影响力和范围,算法通过“贴标签”的运算原理,不可避免地会将显性或隐性的偏见嵌入其中,从而放大了现有的偏见。

作者提到,人工智能伦理审查的范式创新,一方面要打破传统伦理审查范式以生命科学为核心而忽略了其他学科研究诉求的桎梏,另一方面也要考虑人工智能的技术特点和风险特征。需要强调的是,人工智能伦理审查范式创新是在传统伦理审查范式的基础上探索适用于人工智能技术的审查范式,而非对传统伦理审查范式的彻底否定,是革新而非革命。

作者认为,人工智能伦理审查范式创新需从

2016年,关鹏进入中国科学院大学(以下简称国科大)计算技术研究所读研究生时,国科大官方微信团队正招募队员,出于提高审美的考虑,他报了美编岗位。

这是AI进入他日常工作和生活的开始。那时,微信公众号自带的排版功能还不强,因此衍生



图片来源:视觉中国

## 王曦:我与AI共演一出心理剧

■本报记者 袁一雪

心理咨询师王曦拥有心理学与AI交叉学科背景,她一直在尝试让二者更好结合。她找到了心理剧这一方式,通过利用AI让更多人接触心理剧,使之更广泛地运用在心理咨询、解决心理问题。

### 初衷

王曦曾跟“易术”心理剧创始人龚蓀学习过一段时间的心理剧,她告诉《中国科学报》心理剧是如何运作的,“心理剧与话剧、戏剧不一样,没有剧本,没有固定角色,所有剧情由咨询师主诉决定。一场传统的心理剧一般时长3个小时左右,分为3个阶段,包括热身、演出和分享”。

表演团队往往是临时招募的,因此热身环节十分必要。之后,心理剧导演根据主诉情况搭建舞台,重现咨询师纠结产生的各种情绪。接着由团队其他人相互配合,分配角色,把当时未呈现、未意识到的内容重新梳理、整合。

“角色包括主角、配角和观众。这里的观众不是普通意义上的观众,他们是第三方视角的参与者,会在演出结束后与主角交流感受。”王曦解释道。

一场心理剧的表演部分包含9个环节,让咨询师通过锚定剧中描述的现实、处理矛盾、宣泄情绪、梳理自我等环节完成心理干预的过程。王曦特别喜欢在心理剧表演结尾,让台上台下所有人员将手搭在咨询师身上,“手心是有温度的,能传递力量,让咨询师感受到被理解、被支持、被共情,帮助咨询师回到现实、面对现实。这才是心理剧最终的目的——完成干预,回归当下”。

在王曦眼中,心理剧像一把绝世宝刀,效果非常好,但不具有普适性,因为要成为心理剧导演需要累积足够的学时和心理咨询时间。

如何让心理剧起到更大的作用?王曦想到了运用AI。她表示,AI可以通过积累学习时长和经验替代部分导演和演员的角色;共情计算则能在心理剧中更真实地还原人类情感,提供共情的体验;更重要的是,AI的加入能减少场地、人力开支,以及降低成本,让更多人接触到心理剧。

理论、职能和方法三个维度进行;在理论上,主张采用美德伦理学,强调品德养成;在职能上,拓展伦理委员会的职责;在方法上,发展技术工具实现技术与伦理的有机融合。

论文题目:中国工程伦理规范的现状分析——基于37部规范的考察  
作者:周思泽、丛杭青  
出处:科学技术哲学研究,2023,v.40

中国工程伦理规范存在自己的本土道路吗?这一问题的现状如何?2021年,中国工程师联合体在京成立,并于2022年7月6日启动《中国工程伦理规范》的制订计划,中国工程伦理规范的发展正在迎来一个新阶段。

近年来,工程伦理规范不断增多,探索中国工程伦理规范本土道路势在必行。为了中国工程伦理规范未来更好地发展,需要就现状进行全面

出一批编辑器。

“但无论是微信的原生工具,还是其他编辑器,都让创作者不得不将有限的创作经历分散到排版设计上。”关鹏告诉《中国科学报》。

那时,关鹏几乎每天都要排版,通常需要先先在Word文档中编辑完成后,粘贴到编辑器里排好版,再粘回微信公众号,烦琐、耗时且效率极低。

更让他感到麻烦的是,国科大的公号刊发的相关材料免不了使用数学公式,公式不像汉字可以用拼音打出来。因此,关鹏一直在与编辑器作“斗争”,试图解决数学公式无法呈现、排版错乱、标题单一等问题。

“如果写文章的同时能一键排版,那该多好。”有计算机专业背景的关鹏,决定自己开发一款编辑器,“这样就能让人专注于内容创作本身,而不被格式烦扰了”。

不久,关鹏设计的编辑器有了雏形,但仍存在编辑数学公式、微信外链转脚注、修改等诸多细节问题。

2023年,恰逢ChatGPT等AI大语言模型爆火。关鹏尝试用其解决那些难题,

### 尝试

有了这个想法后,王曦开始与北京科技大学教授王志良一起开发了以表达式心理治疗为底层逻辑的虚拟现实系统(VR)。在他们开发的算法支持下,用户可以通过佩戴VR设备,在虚拟环境中体验一场由AI配合的心理剧。

在积累大数据的过程中,王曦摒弃了让AI真假难辨的文字和语音,选择通过影像学习充实数据库,完成初步搭建。“人类本身的防御机制会让语言的真实性大打折扣,心理咨询师能通过咨询者的眼神、肢体动作或语音语调等分辨真伪,但AI做不到。”王曦解释说。

如今,这套系统已初具雏形。虽然暂时无法复制传统心理剧中无剧本的演出过程,但王曦根据常见问题准备了几套剧本,并将一些引导词搭配固定下来。

实际操作时,用户先与心理咨询师沟通,完成主诉,再佩戴VR设备,根据此前的交流选定剧本,搭建场景,比如事件将要发生在电闪雷鸣的草原,还是阴郁多云的广场,抑或是阳光强烈的沙漠。

接着确定参与的配角,配角既可以是生活工作中的其他人物,也可以是焦虑、恐惧、抑郁等情绪的拟人形象。同时,用户将参与的角色摆放在场景中,明确其物理位置的高低、远近,以及形态的大小和颜色等。

最后,根据AI的引导完成剧目中人物的对话。剧情结束后,心理咨询师会与用户围绕心理剧情进行讨论、整理,对心理问题进行处理。

这套系统解决过恐婚恐育等心理问题,也曾缓解因心理问题引发的生理症状。

### 前景

“当然,AI是无法产生情感的,只能通过累积更多的数据无限接近人类的情感,使使用者感受到共情。”当王曦遇到焦虑等情绪问题时,她也会戴上设备,在虚

“AI给了我很多解决方案。比如,如果呈现数学公式,我觉得可行的方案是将公式转换成图片,但我不知道图片的内容怎么制作出来,就问AI,这是搜索引擎很难解决的问题”。

关鹏借助AI大语言模型自建后台服务,实现了公式转图片的功能,还可根据排版情况调整图片大小。对于大量公式的转换,关鹏在前端合理使用了异步请求,即将所有需要转换的公式并行执行,性能上达到了10个公式2秒转换的效果。

在AI的帮助下,编辑器还实现了在左侧编辑栏中输入文字,右侧实时出现初步排版效果,然后一键选出合适的模板完成排版。

如今,作为腾讯的一名开发工程师,关鹏的工作是与代码打交道,遇到问题他还会使用AI这个“小帮手”。

“有时百度和谷歌无法回答我,但AI会给我答案,它可以提供一个具有可行性的新方法,很有参考价值。”关鹏说。

通过这些年的使用,关鹏认为,未来AI一定会更深入地融入我们的生活,“学习如何与AI相处或许是未来社会发展的‘主旋律’”。

拟现实中的心理剧里与自己对话。

虽然她常常带着审视的态度使用,但也会在不经意间被埋藏在系统中的“宝藏”语句治愈。所谓“宝藏”语句,对于她来说就是在不经意间向人获得启示的系统提示。比如,当用户引导从现实层进入情感层时,往往会听到五个句式,即“我感觉”“我害怕”“我需要”“不要以为我”“我期待”。

“有一次,我因为创业压力大,在使用时,听到‘不要以为我’这个句式,我跟自己说,不要以为我会在这种困难里迷失自己,我知道出发的来路,也一定要找回自己。”王曦笑着说,“人的需求其实不高,只需要一些共情、一点回应,滋养干涸的精神,让自己继续走下去。”

现在,第二代系统正在开发中,王曦期望这套能给她带来温暖的AI系统,为更多人送去情感的滋养。她期望在未来心理学领域迎来一款由数学底层逻辑和心理学底层哲学构成的人工智能,就像AlphaFold在生物领域产生的影响那样,大幅度促进心理学的发展,让AI做得更多,甚至实现无人化。

也许有一天融合增强现实(AR)、VR和混合现实(MR)的扩展现实技术(XR)会实现,用户不仅可以在更仿真的元宇宙中完成体验,更可以通过可穿戴设备还原王曦最喜欢的环节——参演者用体温传递力量,“我相信AI未来会让心理学更有魅力,为更多人带来共情的体验。”王曦表示。



用户(右)佩戴VR设备,在虚拟环境中体验一场由AI配合的心理剧。受访者供图

## 速读

论文题目:探索人工智能伦理审查新范式

作者:孟令宇、王迎春

出处:科学与社会,2023,(04)

近年来,我国科技伦理审查工作日益受到重视。传统的以生命医学伦理为基础的伦理审查范式,在应用于人工智能等新兴技术时面临挑战,如何以制度形式保障人工智能的合伦理性将是伦理审查范式革新核心议题。

作者在研究中通过分析人工智能技术的特点,阐述了其与传统伦理审查模式的冲击,并总结了业内先行的实践策略对审查范式创新带来的启示。

人工智能研究已经步入大模型时代,这可能会进一步加深人工智能研究与传统伦理审查之间的鸿沟。因为大模型的“黑箱”性质,增加了审查其决策过程及结果的难度,这与生命医学伦理强调的程序透明度存在张力。此外,在人工智能

识、道德形象上尚缺乏统一的自我定位。

此外,中国工程伦理规范还在美德与规则、预防性/禁止性伦理与激励性伦理的融合兼顾方面表现出明确的模式特征。

作者总结,中国工程伦理规范与西方尤其是美国的工程伦理规范有着显著差别,显示出中国特色道路,但仍有数量太少、代表性不足、内容不完善、现实效力不足等问题。

作者建议,要建立一个在行业涵盖范围、本行业代表性、文本内容和现实效力方面均较为完善的工程伦理规范体系。

第一,建立一个“中国工程师联合体-工程学会-工程行业协会”的行业自律体系,加强职业自治。第二,将对工程师的道德要求从现在混合性质的规范中分割出来,扩大其代表性及影响力。第三,为工程师道德技能的养成提供指引性的参考框架。第四,继续扩大“做好的工程”的伦理意识在规范中占据的比例。(尹一)