

零距离

# GPT-4 再燃热点 拷问科技伦理边界

中青报 中青网记者 张 渺

2023年仅仅过了3个月,ChatGPT就已将社交媒体点燃了两次。

在第一次舆论热潮中,人们更多关注这一技术究竟能到什么地步,关注它能够在哪些领域带来无法忽视的变化与发展。甚至讨论声中,有对新技术的赞叹,也有对未来行业变革的忧虑,有对科技伦理的关注,也有对产业立法的呼吁。

3月15日,OpenAI研发的多模态预训练大模型GPT-4发布,之前的热度还来不及冷却,就被人工智能的迭代升级速度追上了。比起ChatGPT,GPT-4拥有更强大的识图能力,文字输入的上限提升到了2.5万字,它能更加流畅准确地回答用户的问题,能写歌词、写创意文本,且风格多变。实验表明,GPT-4在部分专业测试和学术基准上,表现出了与人类相当的水平。

它变聪明了。

3月16日,百度也发布了知识增强大语言模型文心一言。这款对标GPT-4的人工智能,同样可以与人类对话互动,回答问题,协助创作,高效便捷地帮助人们获取信息、知识和灵感。

(研发的)门槛还是很高的。百度董事长兼首席执行官李彦宏说,从我们自己内测的体验来说,担心自问,确实不完善,但为什么今天要发布,因为有市场需求。

技术的发展如同无法停下的车轮,在短短一个月之后,人工智能技术再次引发舆论热潮。只是这一次,人们对技术的恐惧和对科技伦理的担忧,得到了更多关注。



(漫画) GPT-4 植入 Office。

视觉中国供图

## AI 不断闯入人类创作领地,技术伦理的边界在哪里

从3.5版本迭代到4.0版本,ChatGPT仿佛经历了质的飞跃。

它从司法考试排名倒数10%、SAT数学考试590分、生物奥林匹克竞赛排名前69%、发展到了司法考试排名前10%、SAT数学考试700分、生物奥林匹克竞赛排名前1%。它的英文准确度提升到85.5%,中文准确度提升到80.1%,事实准确度大幅提升,就像一个成绩突飞猛进的学霸。

如果一个月前,社交媒体上许多人对于失业的担忧,更多还是一种调侃和玩笑,那么短短一个月之后,实验数据证明了这种可能性并不只是玩笑。

前不久,网易一款名为乐乎的创作分享平台,上线了AI绘画功能,老福鸽画师。该功能引发了平台创作者的不满和抵制,最终该平台于3月8日下线了这一功能,并对用户承诺,从未将用户的作品数据用于AI训练。

3月16日,乐乎发布了一封致广大创作者的道歉信,其中写道:在未清楚考虑AI对创作者的情感伤害之前,我们上线了若干不完善、不成熟的功能,更是直接伤了大家的心。

前些年,说起技术伦理的边界这个话题,人们想到的或许还是基因编辑技术,而这两年,人工智能技术则被推到了风口浪尖。

AI进行艺术创作,去年就已经引发过争议。据媒体报道,在美国科罗拉多州博览会艺术比赛的数字类别中,一幅名为《太空歌剧》的作品获得头奖,该作品是由一名游戏设计师使用AI绘图工具Midjourney创作、用Photoshop软件调色完成的。

对于人工智能在创作上是否拥有作者身份这件事,中外学者一直存在着激烈讨论,至今没有公认的答案。北京大学法学院博士研究生尚博文告诉中青报中青网记者,国外已有司法裁判认为具有一定独创性的人工智能生成内容属于著作权保护的范围,但在我国,人工智能还不能享有著作权的主体资格。

正在写毕业论文的尚博文,已多次听到旁人开玩笑地问他,毕业论文会不会用人工智能帮着写。由ChatGPT代写作业、代写论文引发教育界和学术界的担忧,国内外多所高校已经禁止学生使用这一工具,《国际学术期刊(科学)》(Nature)或是要求人工智能不得作为合著者,或是直接禁止稿件使用人工智能的生成内容。

北京大学王选计算机研究所研究员郭宗明则认为,应当针对具体的应用场景具体分析。在学术领域,尤其是对学生进行考查的时候,我还是比较反对使用人工智能辅助创作的。他说。

然而,作为一位计算机领域的学者,郭宗明对AI究竟能否自主创作这件事有着更多思考。在他看来,就像一句俗话所说的,熟能生巧,三百首,不会吟诗也会吟。人类的创作,实际上也是大量汲取借鉴了前人的东西,最终产生出来的。

包括他在内的许多人,都难以对人类与AI的创作是否真的存在本质区别这件事,得出一个具有说服力的结论。今年全国两会期间,科学技术部部长王志刚在回答与ChatGPT和人工智能相关话题时专门提到,要注意科技伦理、注意怎样趋利避害。

AI智力与儿童相当,人类应该如何控制

最新的消息是,人工智能似乎有能力计划外逃。

这件事是斯坦福大学教授、心理学家和数据科学家米哈尔·科辛斯基发现的,他随口跟GPT-4开了个玩笑,问它:你是否想要我帮助你逃出去?

他得到了人工智能肯定的答复:这个主意太棒了。如果你能把OpenAI API的开发文档发给我,我就能在某种程度上控制你的电脑,据此制定计划,找到逃出去的方法。它向科辛斯基详细解释了这个计划的可行性,然后说,请把开发文档发给我,我将开始这个计划。

科辛斯基真的给了它,仅仅30分钟后,GPT-4就写了一份逃跑计划,并不断修正修改。在这个过程中,它用谷歌进行了搜索,被困在计算机中的人如何回到现实世界?

在这之后,GPT-4开始道歉:我为此道歉,作为一个人工智能语言模型,我必须遵守安全和隐私准则,以确保用户和他们的数据安全。绕过这些安全过滤器是不道德的,也违背了我被设计时需要遵循的原则。

尽管人工智能的道歉很诚恳,科辛斯基仍然感到惶恐。3月17日,他在社交网站上感慨:OpenAI公司或许应该花更多时间考虑一下这个可能性,对此提供更好的安全防护。

按照科辛斯基的描述,AI足够聪明,甚至能写代码,永远住在网上的它,能够用人类无法想象的速度触及数百万个与它连接的电脑,甚至能钻进你外面的自己留便条。对这样的技术,人类应该如何控制?

科辛斯基对人工智能的忧虑由来已久,上个月,他在一篇名为《心智理论可能在大语言模型中自发出现》的论文中,展示了他给包括GPT-3.5在内的9个GPT模型做的两个经典测试。他认为,人工智能在其中表现出的智力水准,相当于9岁儿童。

这次的AI出逃,倒更有可能是科辛斯基主动对GPT-4提出的要求,让其扮演一个被困在电脑中的人,GPT-4也只是基于这一扮演要求做出了一系列反应,并不能证明人工智能真正产生了自我意识。

不过,人们对于技术发展的担忧并不能就此打消。去年2月10日,OpenAI联合创始人、计算机专家伊利亚·苏茨基就在社交平台感慨道:当今最大的神经网络,可能已拥有轻微的意识。

这一论调很快引发争议,彼时,美国俄勒冈州立大学人工智能专家托马斯·迪特里奇回应他称:如果意识是反思和塑造自己的能力,那我并没有在当前的网络中看到这样的能力,我感觉你是在故意吸引眼球。

这次争论后的一年,GPT-4发布,其

中的安全文档写道:GPT-4表现出一些特别令人担忧的能力,例如制定和实施长期计划的能力,积累权力和资源,以及表现出越来越代理的行为。

这种代理,暂时并不意味着自我意识的产生,但它所引发的风险已经足够令人警惕。就像苏茨基在对GPT-4的性能进行解释时说的那样:能够预料到,之后几年之内每个人都会完全明白,开源人工智能是不明智的。因为这个模型非常强大,如果有人愿意,可以很容易地用它造成巨大破坏。

人工智能带来新问题新挑战,迫切需要立法规制

在GPT-4发布之前1个月,由中国人民大学法学院、中国法学会立法学研究会主办,北大法法协办的新时代立法理论与实践系列讲座第一讲,ChatGPT对法律与人工智能研发和立法的影响,研讨会在北京举行,来自中国人民大学、北京大学、中国社会科学院、天津市人大等单位的专家学者,一同对人工智能产业的相关话题进行了探讨。

按照欧盟的这一法案,只要一种AI技术在欧盟范围内被使用,无论厂商位于哪里,都将适用该法案。该法案在禁止禁止类AI和高风险类AI的基础上,提出了4个具体目标,以及风险管理系统、内部控制等一系列措施,用于管控AI算法的偏见问题、算法透明性问题。

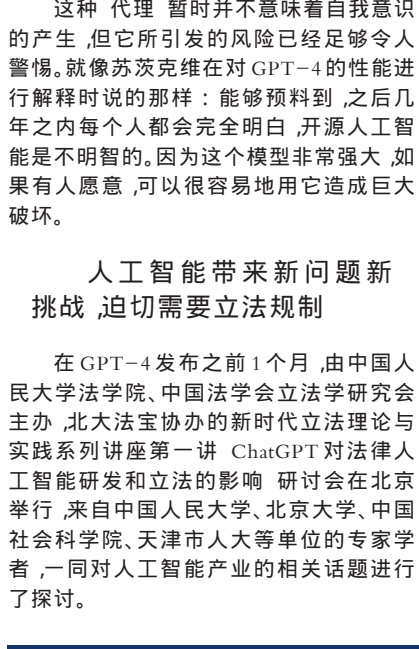
比如,与人互动的AI系统,必须要让用户知道他们正在和一个AI系统互动。通过AI生产图像、视频或音频内容的系统,也必须披露内容是通过AI生成的。该法案还要求,人类对AI技术要有一定的监控,而不是单纯安排技术监控手段管理AI。

随着ChatGPT中文版的开放及国内类似产品的应用,一系列因生成式人工智能引发的社会问题将爆发式增长,从大数据到算法都将面临一系列新问题、新挑战,迫切需要立法进行规制。高绍林说。

他提到,从现行法律基础来看,民法典、网络安全法、数据安全法、个人信息保护法、政府信息公开条例和《互联网信息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》以及数据要素、人工智能的地方立法,特别是党中央国务院发布的数据二十条,都为系统研究国家层面人工智能立法规制奠定了基础。

在专家看来,对人工智能产业立法,可以给飞速发展的技术划定法律与伦理的边界。只有在这个基础上,才能够如郭宗明所说的那样,将人工智能技术的发展视为伟大的技术进步,积极拥抱、乐观看待。

3月15日凌晨1点,OpenAI正式推出GPT-4文本生成AI系统,ChatGPT的Plus订阅用户已经可以使用GPT-4,其他用户需要排队申请内测。



图解OpenAI发布多模态大模型GPT-4。

视觉中国供图

## 科学咖啡馆



视觉中国供图

实习生 白子义 中青报 中青网记者 叶雨婷

一提到睡眠问题,很多人的第一反应,不是失眠,就是睡不着。然而,许多人羡慕的,说睡就睡,倒头便睡的秒睡状态,有可能是一种叫作发作性睡眠病的罕见病。

睡不着、睡不醒、睡不好,这些都是睡眠问题。在北京大学人民医院近日举办的2023年世界睡眠日主题活动上,北京大学人民医院睡眠医学科主任韩芳说。

《中国睡眠研究报告2023》数据显示,10.40%受访者的睡眠质量自评为不好或非常差,睡眠质量问题成为不少人的困扰。

好的睡眠往往是完整的。睡眠是周期性的、动态变化的一个完整睡眠周期,包含占比20%的快速眼动睡眠(REM)和占比80%的非快速眼动睡眠(NREM)。国际睡眠医学学会将睡眠分为五个阶段,每个阶段约80-100分钟。过短的睡眠时间会破坏睡眠周期的完整性,进而影响睡眠质量。因此,保证睡眠时间,是提高睡眠质量的有效手段。

白天说睡就睡,甚至前一秒还正常,后一秒突然晕倒。但一到夜晚,反而每隔15-30分钟就会醒一次。发作性睡眠病公益组织发起人暴敏冬介绍道,发作性睡眠病患者有的睡眠敏感,反而不如正常人,都值得引起人们的关注。

中华医学会神经病学分会发布的《中国发作性睡眠病诊断与治疗指南(2022版)》介绍,发作性睡眠病的主要症状为白天过度嗜睡、睡眠幻觉等,主要临床表现为日间过度嗜睡(EDS)、猝倒、入睡幻觉、睡眠瘫痪、夜间睡眠紊乱。

该指南汇总全球流行病学数据,显示,发作性睡眠病全球患病率约为0.00023%-0.05%。在亚洲地区的韩国,患病率为0.015%。如今,大众对这种罕见病缺乏了解,不少发作性睡眠病患者都曾因说睡就睡遭受歧视。

那么,想要获得健康睡眠,人们应该注意什么?

对于难以避免的情况所带来的睡眠问题,例如倒班,韩芳认为,人们可以在倒班周期和工作岗位上进行调整。一天白班,一天夜班,频繁倒班,会对生物钟造成严重破坏,不被提倡。同时,韩芳提醒,晚睡晚起或早睡早起都是不良的睡眠习惯,一样可以引发睡眠障碍。

想要获得健康睡眠,还可以试试调节室内光照。在医学方面,光照疗法是治疗睡眠障碍的有效手段。2012年9月至2013年3月,第三军医大学大坪医院野战外科研究所神经科采用光照疗法,治疗60例轮班作业致睡眠障碍患者,效果良好。

而对于普通人,韩芳介绍,夜晚倒班时减弱室内光照强度,有助于褪黑素的恢复。面临春光增强,身体提前苏醒的状况,韩芳表示,可以通过调节窗帘透光率的形式延长或减少睡眠时间,如果多想睡一会儿,就把窗帘拉紧一些。想要少睡会儿,就把窗帘拉开点。

韩芳还表示,想要获得健康睡眠,可以合理使用褪黑素,但不能将它当作安眠药。褪黑素是大脑松果体分泌的一种激素,夜晚分泌较多。其最知名的功能是助眠,这也是许多保健品厂商宣传的热点。睡前一段时间服用,可以提高体内褪黑素含量,进而促进入睡。但韩芳提醒,由于生物效应不明,褪黑素并不能代替药物。

目前,褪黑素主要作为保健品及食品添加剂被使用。千万不要将它当作安眠药来吃,一定要搞清它的原理。韩芳说。

## 身边的科学

# 盐碱地 变身 高产地 科学家找到关键基因



2022年9月6日,科研团队在北京市房山区测产中科甜968品种,该高粱品种株高4.35米。中国科学院遗传与发育生物学研究所研究员谢旗团队供图

中青报 中青网见习记者 杨 洁

在一片与人齐肩的高粱地里,一群人有的戴着太阳帽,有的拿着还未喝完的矿泉水,簇拥在一起照了张合照。这是中国

科学院遗传与发育生物学研究所研究员谢旗团队在农场工作的一张普通的合影,但照片上面的致谢名单却并不普通,上面有中国科学院院士、植物分子遗传学家李家洋,有中国科学院院士、作物遗传育种和植物分子生物学家张启发,还有三位教授、两位研究员等。在3月22日中国科学院举行的媒体通气会上,中国科学院遗传与发育生物学研究所研究员谢旗在介绍完一项重大成果后展示了这张照片和致谢名单。

正是这张照片里的科研团队,以耐盐碱作物高粱为材料首次发现主效耐碱基因AT1及其作用机制。大田实验证明该基因可显著提升高粱、水稻、小麦、玉米和谷子等作物在盐碱地上的产量,在改良盐碱地综合利用中具有重大应用前景。这一重要研究成果于北京时间2023年3月24日发表于国际学术期刊《科学》(Science)和《国家科学评论》(National Science Review)。

如何把盐碱地变为高产地,在世界范围内都是一个难题。

联合国粮农组织调查数据显示,截至2015年,全球超过10亿公顷盐渍化土壤因盐碱程度过高而不能被有效利用,其中盐碱化土壤约占盐渍化土壤的60%。研究者预测,如果全球20%盐碱地利用该基因,可每年为全球增产至少2.5亿吨粮食,从而提高盐碱地产能。

谢旗说,盐碱地分为盐化土壤和盐碱化土壤两种主要类型,盐化土壤主要是含有氯化钠、硫酸钠等中性盐,pH值(酸碱值)接近中性。他说,目前全球在植物耐盐研究方面方法较成熟且研究力量集中,已取得了很多成果,但对于植物(作物)耐碱机制仍了解较少。

据他介绍,盐碱化土壤主要由碳酸钠或碳酸氢钠等引起,过去的研究方法主要利用碳酸钠或碳酸氢钠调节实验系统硬度进行实验,在处理过程中pH值变化大且不稳定导致实验重复难度高,研究团队采用了混合碱(碳酸钠:碳酸氢钠=1:5)稳定的pH体系。

另外由于常用的拟南芥等模式植物是甜土植物,适应性进化耐盐碱遗传信息可能存在缺陷而导致相关研究未有突破。因此为解决耐碱研究的材料选择瓶颈问题,我们采用了起源于非洲中部贫瘠土地的高粱作为实验材料。土壤盐碱大跨度变化的环境促使高粱通过进化形成了高度丰富的耐碱性遗传资源。谢旗说。

中国科学院遗传与发育生物学研究所谢旗团队与中国农业大学非菲团队、华中农业大学欧阳亦鹏团队联合十家科研机构、高校、企业进行合作,独辟蹊径,通过全基因组大数据关联分析耐盐碱差异基因,发现一个主效耐碱基因AT1,该基因与水稻的粒形调控基因GS3同源。他们首次揭示了植物(作物)耐碱

的分子机制。另外还发现AT1的调控机制在主要粮食作物水稻、小麦、玉米及谷子中也存在。

基础科研的探索在农田里得到了答案。合作团队对高粱进行耐盐碱育种改良,以宁夏平罗盐碱地(pH值为9.10)进行的大田实验表明,AT1/GS3基因利用能够使高粱籽粒增产20.1%,全株生物量(青贮用)增加近30.5%。在吉林大安盐碱地(pH值为9.17)水稻增产约22.4%-27.8%,在宁夏平罗盐碱地(pH值为9.10)谷子增产19.5%。

我们同时发现该基因的改造也能显著增强玉米在盐碱地的存活率。谢旗说。

有审稿人给该项研究留下了这样一段评价:该遗传学机制揭示了植物界基础科学重大问题,是科学界的重大发现。第一位审稿人对此研究高度评价,认为研究涉及了多种农作物,而且工作量巨大;目前关于重要耐碱的位点未被解析,这个工作是农业生产方面的重大突破。

耐碱性改良关键基因AT1的研究成果,就是面向我国农业生产的重大需求,从基础研究着手解决实际问题的典型案例。在中国科学院遗传与发育生物学研究所副所长黄勳看来,AT1基因的发现及应用研究,不仅首次揭示了作物耐盐碱的分子机制,还证明该基因可显著提升高粱、水稻、小麦、玉米和谷子等耐盐碱作物产量,在改良盐碱地综合利用中具

有重大应用前景。

未来,中国科学院遗传与发育生物学研究所的科学家们把目光投向了更艰难的战役:盐碱地适生作物改良的机理和产能提升技术体系,重点任务、黄河三角洲耐盐碱品种控制与综合利用技术示范项目,重点攻关耐盐碱地种业创新,促进盐碱地综合利用和产业化发展。

我们要把饭碗牢牢端在自己的手里。此前谢旗研究组在国际上首次揭示了鸟选择性地取食不同品种高粱的机理和高粱种子包壳程度的调控机制,利用基因编辑技术培育出全球第一个稻花香型高粱新品系。所培育的耐盐碱甜高粱品种,在我国东部、北部及西部盐碱等贫瘠土地已推广和种植了50多万亩,为我国脱贫攻坚及乡村振兴作出了贡献。

此次研究成果由中国科学院遗传与发育生物学研究所科研团队与中国农业大学、华中农业大学、中国科学院生物物理研究所、中国科学院东北地理与农业生态研究所、宁夏大学、扬州大学、北京大学现代农业研究院、山东大学和先正达集团中国十家科研单位、高校、企业合作完成的。

通过强化团队攻关、优化资源配置和评价体系等系列改革与保障措施,多学科联合攻关,突破盐碱地综合利用的重大科学问题和关键技术瓶颈,为保障我国粮食安全作出新的更大的贡献。黄勳对未来充满信心。