

科技伦理研究的目的是不是阻碍科学发展,而是为科技发展提供伦理支持与规约,使科学家、工程师研发出更安全、更人性化、更容易被社会接纳的科技产品。

自主武器:技术与伦理的边缘

■本报记者 袁一雪



郭刚制图

当美、英、法三国将导弹投掷在叙利亚上空,鲜血与死亡、绝望与恐惧便铺洒在这片大地上。当人们在惊叹导弹精准的同时,野心勃勃的武器专家却并没有止步于此,因为他们已经将目光聚焦在自主武器上。

今年2月,韩国科学技术院(KAIST)和集束炸弹制造商韩华系统公司开设了“国防和人工智能融合研究中心”,前者表示,新的智能机器人(AI)研究中心将专注于利用人工智能命令和控制系统,对大型无人驾驶潜水器进行导航,以及利用人工智能进行智能飞行器训练,追踪和识别物体等。

这一决定引起了全世界人工智能界科学家的关注。一封由全球逾50名顶尖的人工智能研究员联合签名的公开信直指韩国科学技术院这一决定。签名的学者来自蒙特利尔大学、多伦多大学、瑞士人工智能实验室、加州大学伯克利分校、香港中文大学和南京大学等。

信中称,自主武器有可能成为恐怖分子的武器。暴君和恐怖分子可以利用它们对付无辜的人群,他们不会理会任何的道德限制。如果打开了这个潘多拉魔盒,将很难关闭。“与过去禁用其他技术,例如禁用激光致盲武器一样,我们可以简单地决定不开发自主武器。我们敦促KAIST遵循这样的共识,致力于改善人们的生活,而不是伤害人的生命。”

在公开信发布后,韩国科学技术院也即刻发表了声明,声称他们没有研发“致命自主武器系统或杀人机器人”的计划。但这并没有减少人们对于此事的关注度。

反自主武器集结号

参与签名的南京大学计算机系教授周志华在接受《中国科学报》记者采访时表示:“联名抵制的发起人是澳大利亚新南威尔士大学的Toby Walsh教授,所有参与签名的人员都是他邀请的。他按自己的原则邀请了一些国际人工智能领域有威望的学者来给韩国方面施加压力。”

与此同时,美国谷歌公司内部超过3000名谷歌员工签名要求公司退出一项美国五角大楼人工智能军事计划。

这些发生在人工智能领域的事件也引起了联合国的重视,4月9日,联合国《特定常规武器公约》政府专家小组在瑞士日内瓦召开

会议,预定讨论致命自主武器系统研发,也就是常说的“杀手机器人”可能带来的影响。

其实在2015年,一封要求停止研制具有杀人功能自主武器系统的公开信得到了2万人的签名,签名者主要是人工智能领域的专家以及对此话题颇为关切的人士,其中就包括已故著名物理学家斯蒂芬·霍金。

霍金一直担忧人工智能的崛起将威胁人类,在去年举行的2017全球移动互联网大会,他通过视频表示,人工智能的威胁分短期和长期两种。短期威胁包括自动驾驶、智能性自主武器以隐私问题;长期担忧主要是人工智能系统失控带来的风险,如人工智能系统可能不听人类指挥。

自主武器

不听人类指挥,超出人类控制,这也是人工智能的“自主武器”引发业界人士普遍关注的关键点。因为其使用的“强人工智能”并非目前人们认知中普遍接受的“弱人工智能”。两者的区别在于,弱人工智能可以辅助人类,减轻智力或体力劳动;而强人工智能则是机器人达到甚至超越人类智慧水平,甚至具有意识,可以自主安排自己的活动。

目前,已经有些武器使用智能功能实现一部分辅助功能,比如带有自动瞄准装置、导弹定位系统的武器等,但这些武器的使用都需要由人来作决定,人类的思维决定向谁射击,是否需要射击。“自主武器与一般意义上的智能武器,我们联名反对的是这种脱离了最起码的人类控制的自主武器。”周志华说。

如今,“把人工智能科技应用于武器研发非常普遍,比如世界各国都对军用机器人给予了高度重视。对于智能机器人是否拥有自由意志或意识,是一个有争议的话题。对于军用智能机器人而言,关键的问题是机器人的自主程度问题。”上海交通大学科学史与科学文化研究院副教授杜严勇也在接受《中国科学报》记者采访时表示:“由于军用机器人拥有强大的杀伤力,没有恐惧感,也可能对人类没有同情心。而且,在目前的技术水平下,机器人很难区分战斗人员与非战斗人员,完全自主的智能机器人在战争中不可避免地会伤害及无辜,也很容易使战争规模升级。”

技术进步不能抛弃伦理限制

“自主武器并不符合伦理,更不符合人类

的利益。所以,目前主流人工智能学界根本没有向这个方向努力。”周志华强调。

不论是家庭服务型机器人还是自主武器都离不开“伦理”二字。“发展智能机器人,是为了提高人类的生活质量,使人们更好地享受科技的益处。”

为此,机器人的行为就需要符合人类的伦理原则与规范,才能让人们更好地接受机器人。”杜严勇解释。比如一个提醒人类吃药的机器人,当遇到提醒对象拒绝吃药时应该作何反应:是为了其身体考虑强行喂药还是提醒几次后就放弃?“如果遇到拒绝服药的对象,人们通常会期望机器人应该比人类更有耐心,采取各种合理的方式去说服人们吃药,而不是简单地放弃或强迫。也就是说,人们期望机器人的道德水平更高,在道德活动中更加积极主动,同时机器人对人类的行为应该更加宽容,这样的机器人可能更容易被人们所接受。”

简单的服药一件事就足以令智能机器人“头痛”,当智能机器人涉及到暴力、性、侵犯隐私等问题时则更难解决。为了避免智能机器人威胁人类,已故科幻作家艾萨克·阿西莫夫曾提出了以“不得危害人类”为核心的“机器人三原则”,即机器人不得伤害人类,或坐视人类受到伤害;除非违背第一法则,否则机器人必须服从人类的命令;在不违背第一及第二法则下,机器人必须保护自己。此后,阿西莫夫又添加了一条定律——第零定律:机器人不得伤害人类族群,或因不作为使人类族群受到伤害。

但是,定律不是法律。现在的机器人发展已经完全舍弃了阿西莫夫制定的三大定律,此次的联合签名抵制事件就可窥到一二。“所以,从长效性与普遍性的角度看,需要国际组织来制定人工智能科技研究的伦理守则与规范,这方面的工作已经是刻不容缓了。”杜严勇表示。

但同时,杜严勇也强调,哲学家研究科技伦理并不是要“指导”科学家的学术研究,科技伦理的研究目的不是去阻碍科技发展,而是为科技发展提供伦理支持与规约,使科学家、工程师研发出更安全、更人性化、更容易被社会接纳的科技产品。

“技术本身没有善恶。就好比菜刀本身没有善恶,在好手里用来切菜,在坏手里可能成为凶器。要限制的是使用技术的人,而不是技术本身。”周志华说。

读心有术

收脏钱更容易做坏事

刚刚发布的2018菠萝科学奖的心理奖得主是浙江大学管理学院周欣悦团队,他们研究的内容跟“钱”有关:收脏钱更容易做坏事。

这一内容出自周欣悦团队2013年的一篇科研论文《干净与肮脏的钱对态度、价值观和人际行为的不同影响》,他们发现,钱的干净与否会作用于人们的心理感受。

研究从人们几乎每天接触的小额纸币开始,实验室是普通的菜市场。人们在菜市场买菜时,一般是先选菜再付款,这种情况下,他们无法看到“脏钱”对心理的影响。于是,研究人员巧妙设计了一个实验步骤。首先,实验人员先买回固定份额的蔬菜,用脏钱或干净的钱付款。然后,实验人员会再次要求选购其他的菜,并且把刚才的钱要回来,最后一次付清两次的菜钱。

这样的实验结果十分有趣,那些中间拿到过干净的钱的商贩,在人们二次购买时几乎不会缺少斤两,相反,如果商贩接触的是脏钱,那么人们再次买到的蔬菜就会少一点儿。

由于菜市场的实验环境,条件不容易精确控制,研究人员回到真正的科学实验室进行了进一步研究。这是一个考验信任的分钱游戏。研究人员告诉被试,他们会收到另一个人给的钱,然后可以选择分一部分回馈给那个送钱的人。此时,最公平的选择是直接分出一半儿送回对方,而结果是,拿到干净钱的人比拿到脏钱的人更愿意这么做。

然后,研究人员以手指运动为名,让被试数钱或者数一摞纸。结果显示,刚刚数过干净钱的人,道德标准提高了。在答题时,他们更难接受不道德的行为;在玩游戏时,他们更倾向于公平的策略;在受到不公平对待时,他们宁可一分不赚也不让对方获利。而那些刚刚接触了脏钱的人,道德标准则有所降低,对不道德行为接受度更高,即便受到不公平对待也先计算得失。

他们甚至发现,即便不接触真的干净钱或脏钱,只要阅读一篇关于钱是否干净的文章,其中干净或肮脏的描述都能影响到人们对公平的判断。

于是,研究团队得出了这样一个结论,干净钱提升了人关于公平的标准,而脏钱则相反,让人们把底线放低了。(朱香整理)

裸眼看3D:从虚无走向现实

■本报记者 袁一雪

看3D影像,眼镜、头盔等设备几乎是标配,因为现在的3D图像利用的是人两眼视觉差,必须佩戴相应的设备才能保证两个角度拍摄的图片呈现立体视觉。

不过,近日,北京邮电大学信息光子学与光通信国家重点实验室教授桑新柱带领研究团队研发了“裸眼3D光场显示”设备,即不需要其他辅助器件,直接用肉眼就可观看悬浮于空中的3D物体。

一次几乎“违背”物理规律的尝试

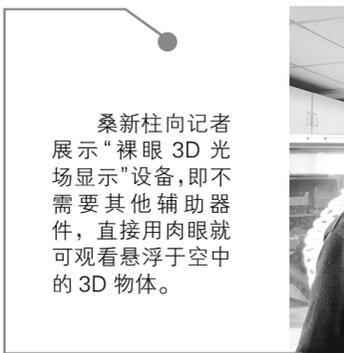
在桑新柱的实验室,几台机器将不大的实验室装得满满当当。摆放在房间正中的机器前,一颗旋转的星球“飘浮”在空中,星球上环绕着英文字母,不一会儿球体变成了快速挥舞翅膀的蜂鸟,然后又幻化成正在飞翔的五彩斑斓的蝴蝶……如果用手触摸,会发现这些只是光影。更神奇的是,这些图像不会随着人的视角不同而发生突变,也就是说,在改变观看角度的情况下,人眼看到的都是3D图像。

“目前图像距离屏幕约90厘米,在95厘米以外都可以观看,没有严格的距离限制。但因为现有设备显示尺寸还不够大,所以最好在10米之内观看。观看角度则是正面约50度都可以。”桑新柱在接受《中国科学报》记者采访时说。

在房间另一边的显示器上,一个立体的胸腔骨架中一颗心脏正规律地跳动着,与刚刚看到的球体不同的是,骨架的前方有一块透明面板挡在人与图像之间。不要小看这块透明面板,它上面遍布着控光器件。而类似的控光器件在第一台机器中却是隐藏在图像后黑色的机箱内。

控光器件是桑新柱研究团队实现裸眼3D光场图像的核心部件,“人眼看到现实世界中的物体,是因为光照射到物体上产生了漫反射,漫反射后的光照进入人眼,人眼就看到了3D影像。”桑新柱解释说,“我们所做的控光器件就是模拟现实的漫反射光分布让人眼看到3D影像。”

“这种在空气中模拟光场分布实现3D显示的方法在很多业内专业人士眼中是违背物理规律的。”桑新柱说。一直以来,人们都在寻



桑新柱向记者展示“裸眼3D光场显示”设备,即不需要其他辅助器件,直接用肉眼就可观看悬浮于空中的3D物体。

找一种途径作为光的介质让物体显现出来。比如大部分研究都利用水雾作为介质,光打到水雾上进而发生漫反射,人眼看到时就是立体图像的效果。

今年1月,一篇发表在美国《自然》杂志上的论文就阐述了类似的研究:来自美国的研究人员研发了一种裸眼3D悬浮显示系统,通过激光控制扫描粒子,借由粒子让光产生散射然后产生图像。但是由于光源受设备所限,且高速粒子控制并不容易,所以显示的图像的分辨率还比较低,高分辨率图像的帧率比较低,目前显示图像只有厘米级大小。

“我们的技术不依靠空气中的任何介质,直接在空气中就能显示立体图像。”桑新柱说。他们根据人眼的视觉特性及影像的悬空距离计算得到超精密光学微结构的的面型和排布,再让在二维显示器发射出的携带3D信息的光线按照计算好的方向射入人眼,当空间光场中有无数光线沿着特定的“路径”穿插传播后,最终汇聚到达人眼中呈现的就是3D图像。

而且,这样的显示效果远远超过通过介质的方式,这一点在实验室内展示的从十几厘米到几十厘米大小不等且清晰的3D图像就是最好的证明。

30年的沉淀与积累

机器设备虽多,但它们的成像原理并不相



同。而这些成功的研究成果源自北京邮电大学信息光子学与光通信国家重点实验室三十年的坚持。

2008年,桑新柱结束了在美国加州大学尔湾分校的博士后工作,回到母校北京邮电大学,与北京邮电大学信息光子学与光通信研究院教授徐大雄在同一间办公室。“我当时的主要研究方向是光电子器件和光通信,而且回国时,光通信还在蓬勃发展的,光学显示却成了冷门。当时徐大雄老师曾向我感叹,他一直从事的全息显示研究只怕后继无人,会让其之前的努力付之东流。”桑新柱回忆说。

如今,快10年过去了,桑新柱已经在三维显示领域实现了另一个学术高度。

现在,桑新柱教授的研究团队已经与一些高端应用单位展开了合作,为他们订制3D光场显示系统。可以说,他们对于3D影像的掌控几乎到了炉火纯青的地步。“最近我们的学生做出了一套新的软件,它可以将目前市场上的影片直接转换为裸眼3D播放出来。”

而且,实验室内的样机,既可以做成3.6米宽、2.07米高的“庞然大物”,以保证其分辨率更高,显示图像更清晰,视觉冲击力更强;也可以做成犹如文件夹大小,放入包中,随用随取。

“我要感谢我的团队和我的学生,他们的辛勤努力付出与创新才让我们这项技术在该领域领先一步。”桑新柱说。

热闻

久坐与大脑内侧颞叶

一项最新研究显示,每天坐的时间太大会对人的记忆力产生影响。美国加州大学洛杉矶分校的研究人员发现,长期的久坐习惯可能会使成年人大脑中与记忆有关的部分发生变化。此前的研究中,科学家已经发现久坐会增加中老年人心脏病、糖尿病和过早死亡的风险。

在以往结果的基础上,研究人员将焦点放在了久坐对大脑的影响上。他们发现,久坐与大脑内侧颞叶的缩小存在关联,而这一区域与新记忆的形成有关。大脑的缩小可以作为中老年人认知衰退和痴呆症的前兆特征。

有35名年龄在45岁到75岁之间的志愿者参与了这项研究。分析结果显示,久坐习惯与内侧颞叶的缩小密切相关,无论个体的身体活动水平如何。换句话说,研究表明“久坐习惯是(内侧颞叶)缩小的重要预测指标,即使身体活动水平很高,也无法弥补久坐带来的有害影响”。

“浪费基因”

近日,著名期刊《细胞代谢》杂志以封面论文的形式,发表了厦门大学教授林圣彩和副教授林舒勇研究团队的重要研究成果,因为他们发现了一个名叫Aida的“浪费基因”。该论文的共同第一作者为罗辉和姜鸣。

据报道,Aida这个基因最早是林圣彩团队与清华大学孟安明院士团队合作鉴定和命名的,当时对Aida的具体功能还不是很清楚。林圣彩团队发现,从小鼠体内去掉Aida基因后,与Aida基因还在的小鼠相比,去掉Aida基因的小鼠胖得像充了气一样。

由此可见,Aida基因可以防止变胖。于是林圣彩和林舒勇团队就围绕没有Aida基因的易胖小鼠研究了10年。他们发现,食物中脂肪在小肠内的吸收效率是个体易肥胖的一个主要决定因素。而Aida基因编码的AIDA蛋白质就是控制整个过程的。在没有Aida基因小鼠体内,小肠细胞中与脂肪合成相关酶大量表达,加快脂肪的合成,这样一来,小肠细胞就腾出空间继续从小肠里吸收脂肪的代谢产物了。这就间接地促进小肠对脂肪的吸收。(北緯整理)

高端科普书20年

(上接第1版)

“第一推动丛书”的读者大部分是热爱科学的大中学生,吴炜说,有些人就因此爱上了物理、天文等,走上了科学研究之路。

也正正如刘兵所说,对于高端科普书还需要考虑一个问题,如此集中出版,受众是否作好了充分的准备?“所有新的变化,特别是有学术基础的、新观念的,肯定都要有培育的过程。”

历经20多年,高端科普书的各个方面也在发生变化。有意思的是,与过去不同,此次接受采访的编辑们都认为,近几年图书销售在好转。

吴炜感觉到,近几年的科普图书市场逐渐热起来,一些从不做科普的出版社也转向做科普书,她分析,这应该与社会对科技的依赖度越来越高,全民更加自觉地重视科技有关,“科学也变得时髦起来”,许多读者是为了改善自身的知识结构而主动阅读科普书,以打开看待世界的更多角度。

李继勇认为,高端科普书销量的读者相对固定,只要图书内容经得起检验,每种图书的销量都不是太低,相对来说还比较平稳。“少的有八九千,好的可达两三万,虽然比不上,但也足以让人在这条路上走下去了。”并且他相信,随着政府和公众对科普重视程度的提高,还将越走越好。

上海科技教育出版社副总编匡志强对这类书的前景预测,“总体说来还是比较乐观的”,因为这类图书的阅读群体正在逐渐扩大。不仅阅读群体越来越年轻化,且有一些并非从事学术工作的人都开始接触这些图书。

与此同时,由于现代科学技术发展不断加速,对社会生活的影响也日益显著,读者了解最新科技及其社会效应的需求也逐渐提升。也就是说,“高端科普书的对象和内容都在发生变化,这也对出版者提出了更多的挑战”。匡志强表示。

目前,“第一推动丛书”5大系列共53种,“哲人石丛书”目前已出版127种,“科普名家经典”至今出版了三十余种,“新知文库”即将出版100种。

其中“第一推动丛书”变装过5次。今年还出版了25周年版套书,版式设计采用网络技术精确设计,简约明快,适合年轻人的阅读风格。

王世平介绍,科教社近年来在高端科普书领域的变化是,确立了“繁荣原创科普”的目标,有意识减少引进版选题,加强原创科普出版。

“新知文库”即将推出两个子系列——“原创新知”和“彩图新知”。徐国强说,10年时间,引进版已经培养了读者的趣味,现在到了激发本土优秀作者的时候。

李继勇表示,“科普名家经典”会延续现在的出版理念,“不同的是会更加细化市场,往更通俗的方向上靠,争取让这些书成为青少年成长路上的好伙伴”。

追求完美的社会效益和经济效益,是出版人永远的“梦”。除了上文提到的,《宇宙的琴弦》《生命是什么》《美丽心灵——纳什传》《大流感——最致命瘟疫的史诗》《魔鬼出没的世界》《像物理学家一样思考》《双螺旋》等等,20多年来,一大批优秀高端科普书出版为读者熟知,相信实现这个“梦”并非遥不可及。