人工智能的“无间道”通向何方

■本报记者 丁佳

前段时间，一些视频在网上广泛传播。有人在1994年电视剧《射雕英雄传》中使用人工智能技术，将黄蓉一角的扮演者朱茵换成了杨幂，效果几可乱真；还有人将人工合成的“美国前总统奥巴马”放在境外社交媒体上，吐槽现任总统特朗普是“彻头彻尾的笨蛋”。

人们看了之后，大多当做网络段子一笑而过，但仔细想想的话，如果你的脸也被用来制作假视频、假新闻，你还能笑得出来吗？

实际上，人工智能技术的滥用和因算法漏洞引起的安全问题，已经引起了业内人士的高度警觉，一场看不见硝烟的“攻防战”，正日趋白热化。

危险逼近

人工智能正深刻地改变着人们的生产生活方式，但今年以来，利用人工智能技术造假的新闻屡见报端，也引发了越来越多的公众关注。

“公众对人工智能技术抱有希望和想象，相信它能为生活的方方面面带来改变，但我们也应当正视，当前人工智能技术还处在比较初步的发展阶段，还面临着一些挑战。”清华大学教授、该校人工智能研究院基础理论研究中心主任朱军坦言。

冰冻三尺非一日之寒。早在几年前，学术界和产业界已经感受到危机在悄悄逼近。

“这几年在各种各样的学术会议上，我们能听到越来越多关切的声音，有关人工智能安全的论文也是呈指数级增长。”瑞莱智慧（RealAI）公司CEO田天说。该公司是清华大学人工智能研究院重点孵化的一家高新技术企业。

人工智能滥用也已引起国家有关部门高度关注。5月28日，国家互联网信息办公室发布了《数据安全管理办法（征求意见稿）》，其中明确表示，网络运营者利用大数据、人工智能等技术自动合成新闻、博文、帖子、评论等信息，应以明显方式标明“合成”字样；不得以谋取利益或损害他人利益为目的自动合成信息。

除了技术的滥用，人工智能本身存在的算法漏洞也令人担忧。几个月前，比利时鲁汶大学的研究人员发布了一则视频，实验者通过在肚子上贴一张打印出来的图片，就能够逃脱智能监控系统的识别，实现“隐身”；而想要把“隐身术”传给同伴也很简单，只要把这张纸递给他就好了。

“可以设想，如果这项技术被犯罪分子掌握，他们就能够很容易躲避智能安防系统，给公共安全带来极大威胁。”田天说。

“道”高一尺必先“魔”高一丈

2018年6月，清华大学人工智能研究院揭牌成立，以“一个核心、两个融合”作为发展战略，即以人工智能基础理论和基本方法研究为核心，积极推进大跨度的学科交叉融合，积极推进大范围的技术与产业、学校与企业融合。

该研究院院长、中国科学院院士张钹认为，作为研究性机构，清华大学人工智能研究院的本分是搞好学术研究，“但是人工智能又与应用结合相当紧密，因此研究院不仅仅是开开会，我们的研究一定要产生出算法、孵化出新产业，甚至成为人工智能界的‘BAT’”。

在这一思路的指引下，瑞莱智慧应运而生，其核心研发团队来自清华大学、北京大学、中科院等，致力于打造安全、可靠、可解释的第三代人工智能，提供工业检测、预测性维护、金融风控、人工智能系统安全评测与防护等服务。

“在技术层面，人工智能安全会是一个持续攻防的过程。”田天说，“攻击者会采用新的方法来进行攻击，而防御者也需要适应动态变化的攻击场景。如果要占据主动权，就要对未来可能潜在的攻防场景做预演，并为应对潜在的风险提前设计可行方案。”

瑞莱智慧算法科学家萧子豪介绍，他们很早就关注到网上假视频、假图片的问题，一直在想能否利用清华积累的技术优势，去防止人工智能技术的滥用。

但是，想练就一副“火眼金睛”，就需要大量高质量的数据去训练计算机。

近期，瑞莱智慧将工作的重点放在了换脸假视频的检测上，研发团队利用清华大学人工智能研究院在深度生成模型方面的积累，通过生成更加逼真的假视频，训练出了一个更加可靠的模型来检测假视频。经测试，这一模型的识别准确率超过了99%，能够覆盖市面上绝大多数换脸技术生成的假视频，达到了“真的不误杀，假的不漏判”的目标。

打开“黑盒”

张钹在2016年中国计算机大会上提出了“后深度学习时代的人工智能”，此后又进一步提出了“第三代人工智能”的概念，曾引起广泛关注。

这是因为当前主流的基于深度学习的人工智能技术即第二代人工智能，存在着一个难以逾越的“天花板”，业内人士称之为“黑盒”。

“当前人工智能的算法越来越复杂，它告诉你明天要下雨，但是你完全不知道它是怎么做的决策，又谈何信赖呢？”田天坦言，如果被人利用，“黑盒”将带来严重的安全问题。

鲁汶大学的“隐身术”原理其实就是针对这种算法漏洞，专门生成一些噪点，来诱导人工智能识别出错。

瑞莱智慧也做了类似的尝试。技术人员研发了一种带有噪点的“眼镜”，只要戴上，就能通过商用智能手机的刷脸解锁。

多年来，清华大学人工智能研究院一直在攻防技术上深耕，曾在谷歌公司组织的“对抗样本攻防竞赛”上获得了所有任务的冠军。创办瑞莱智慧后，研发团队利用新技术做了更精确的建模，更加深入地研究如何利用白盒的替代模型来攻击黑盒的受害模型。通过发现真实存在的、危害更大的漏洞，团队进一步有针对性地设计防御算法来提升整个人工智能系统的安全性。

今年5月，清华大学人工智能研究院与瑞莱智慧联合发布了人工智能安全平台——“RealSafe对抗攻防平台”，它能够提供多种攻击与防御算法，开展全面、细致的攻击防御算法的评测与比较。此外，通过该平台，研究人员可以很容易地开展人工智能安全方面的学习和实训，帮助更多的人来关心和研究人工智能安全。

“随着人工智能技术渗透到人们生活的各个方面，因技术滥用、算法漏洞等导致的安全问题也随之而来。对此，公众需要提高自我保护意识。”田天呼吁。

显然，人工智能的“无间道”还将长时间上演。但可以肯定的是，这条螺旋式上升的路，最终会将人类带往一个充满想象力的未来。

简讯

山西新认定 77个科普教育基地

本报讯 记者从山西省科协获悉，太原工业学院工程训练中心科普教育基地等77个单位，日前被认定为“山西省科普教育基地”（2019-2023年）。

据山西省科协负责人介绍，此次新认定的山西省科普教育基地经过各市科协、各省级学会（协会、研究会）初选推荐，并由专家评审确定。希望这些科普教育基地充分发挥自身科普资源作用，建立开展科普活动的制度，将合法经营与公益性科普相统一，更好地面向公众普及科学技术知识，在科普工作中起到示范带头作用。 （程春生 郜丰）

600余位青年学者天津研讨 合成生物学发展

本报讯 “合成生物学技术正在逐步发展成为像分子生物学技术、基因组测序技术一样的底层基础生物技术。与不同学科的交叉融合，最容易爆发出最原始的创新、最颠覆性的技术。”第五届合成生物学青年学者论坛日前在天津举行。关于合成生物学的发展，论坛主席、中科院天津工业生物技术研究所研究员江会锋如是说。

此次论坛由中科院天津工生所主办，600余位中外青年学者和学生参会。与会专家学者围绕合成生物学与数学、化学、计算科学、工程科学等学科的交叉融合，与微生物学、植物学、动物学等传统生物学科的协同创新，以及合成生物学在材料、能源、医疗、环境等领域的产业化应用进行了深入探讨。 （闫洁）

中药实验药理分会第十五次学术会议在广州召开

本报讯 8月20日至22日，中华中医药学会中药实验药理分会第十五次学术会议在广州召开。会议邀请了中药全球化联盟主席、美国耶鲁大学教授郑永齐，解放军总医院第五医学中心研究员肖小河，中国中医科学院中药研究所所长陈士林、副所长朱晓新、研究员林娜，中国科学院上海药物研究所研究员丁侃等专家进行了学术报告。

此次会议由中华中医药学会主办，中华中医药学会中药实验药理分会、广东省药理学会中药药理专业委员会、暨南大学共同承办。会议期间还进行了换届选举会议，上海中医药大学中药学院院长徐宏喜当选分会主任委员。 （朱汉斌）

粤港澳大湾区西岸科技创新和人才培养合作联盟在澳门成立

据新华社电 粤港澳大湾区西岸科技创新和人才培养合作联盟日前在澳门举行成立仪式，70多名来自大湾区西岸的高校及政府代表出席。

联盟由澳门大学、北京师范大学—香港浸会大学联合国际学院、五邑大学共同发起，共17所大湾区高校参与。

这是一个来自智利安第斯山脉，距今2000万年的灵长类动物——卡拉斯科智利猴的头骨化石。
图片来源：AMNH

2000万年前头骨暗含类人猿大脑进化线索

本报讯（记者唐凤）人们一直认为类人猿的大脑体积随时间的推移逐渐增大。对南美洲最古老、最完整的灵长类头骨化石之一进行的最新研究表明，这一群体的大脑进化模式其实更复杂。这项研究发表在《科学进展》上，由美国自然历史博物馆、中国科学院古脊椎动物与古人类研究所和美国加州大学圣塔芭芭拉分校研究人员合作完成。

距今4000万年前，阔鼻猴类与狭鼻猴类分开独立演化。因此，阔鼻猴类是研究包括人类在内的狭鼻猴类脑演化的自然参照系，而曾生活在南美洲的基干阔鼻猴类卡拉斯科智利猴是该参照系中目前已知的关键参照点。研究人员对智利安第斯山脉高处发现的一个距今2000万年的类人猿化石进行了详细研究，这是已知的唯一一个卡拉斯科智利猴标本。

该研究组之前的研究大致了解了这种动物大脑大小与身体大小的关系。大多数灵长类动物的脑指数都比其他哺乳动物高，甚至有些灵长类动物——尤其是人类及其近亲——比其他灵长类动物还要聪明。而最新研究进一步阐明了类人猿谱系树的模式。

研究发现，卡拉斯科智利猴的系统发育脑商（PEQ）相对较小，仅为0.79。相比之下，大多数现代猴子的PEQ从0.86到3.39不等，而人类则高达13.46，且脑容量也有了显著扩大。研究提出了对卡拉斯科智利猴大脑的新见解：在现代灵长类动物中，大脑中视觉和嗅觉中枢的大小呈负相关，这反映出一种潜在的进化“权衡”，意味着视觉敏锐的灵长类动物通常嗅觉较弱。阔鼻猴和狭鼻猴类祖先拥有相对大小相似、但结构不同的脑。两个类群都有脑容量增大的趋势，且趋同演化出嗅球缩小、沟回增加等特征。

相关论文信息：*https://doi.org/10.1126/sciadv.aav7913*

光电仪器亟待解决“卡脖子”问题

本报讯（见习记者卜叶 记者黄辛）近日，由中国科学院上海交叉学科研究中心主办的第112期交叉学科论坛“先进光电仪器与应用研讨会”在沪召开。与会专家认为，农业、畜牧业、生态环境监测等领域对光电仪器装备有迫切需求，急需研制高精度植株表型光电装备、保障粮食安全的先进光电装备、危害气体实时监测预警装备和多元多模态危害物质在线高通量监测装备等，解决“卡脖子”难题。

中科院院士、中科院分子植物科学卓越创新中心/植物生理生态研究所所长韩斌坦言，在水稻表型与基因测序方面，我国使用的高端光电仪器主要靠进口，价格高昂、维护成本高。“期待能自主研发相关光电仪器，为水稻的基因精确测序、关联分析、水稻杂种优势遗传机制研究提供技术支撑。”

中科院遗传与发育生物学研究所研究员降雨强认为，精准获取高分辨率时空表型是“后基因组时代”相关科研工作的瓶颈，我国虽已逐步开展了表型组学设备的引进和研发工作，但主流超分辨显微成像技术及其装备都被国外进口产品垄断，急需自主可控的高端光电仪器。

事实上，进口设备依然无法彻底解决我国的科研问题。湖南省农业科学院研究员朱明东表示，进口仪器难以适应中国水稻产业发展的实际需求，已成为我国水稻产业精细化、规模化发展的阻碍。

光电仪器是以光电子学为基础，综合利用光学、精密机械、电子学和计算机技术解决各种工程应用问题的装备。不同领域和应用对光电仪器的需求不同。

比如，在育种方面，植株表型和生长环境监测对科学高效育种十分重要。光谱、红外、激光等技术能够实现植株生长各要素的全面监测，但该类高端光电仪器的研发在国内仍是空白。

霉菌毒素污染给农业和畜牧业生产造成了巨大的经济损失，急需研发霉菌类污染物在位检测及脱毒处理一体化装备，解决我国小麦赤霉病检测技术难题。“该装备应该具备覆盖90%以上已知霉菌毒素的能力，并能完成分级处理。”中科院长春光学精密机械与物理研究所研究员谭鑫说。

此外，光电设备在环境保护方面也表现出应用潜质。中科院上海技术物理研究所研究员顾明剑认为，对有害气体远程实时监测预警存在盲点，应充分发挥红外高光谱成像技术的优势研发相关光电仪器。

水体污染监测也能用到光电仪器。中科院安徽光学精密机械研究所研究员谢品华介绍，水体放射性核素在线监测预警装备将是光电仪器的又一应用。

中科院上海光学精密机械研究所研究员黄立华认为，先进激光技术在该类仪器中发挥着重要作用，应该重视并开展激光浮游菌粒子计数器和飞秒激光技术的研究。

“研制过程中，要实现关键核心器件的自主可控，关键核心器件是高端光电仪器的重要攻关方向。”中科院上海技术物理研究所所长丁雷表示。

发现·进展

中科院地球环境所等

中印度—太平洋地区过去2700年降雨在减少

本报讯（记者张行勇）中科院地球环境研究所“一带一路”气候环境研究中心研究员谭亮成领衔的国际团队，利用泰国南部可兰洞中3根可重复的、精确定年（最小测年误差为0.5年）的石笋氧同位素记录，重建了中印度—太平洋北部地区过去2700年连续的高分辨率降雨记录。相关成果近日发表在美国《国家科学院院刊》上。

热带地区是全球气候变化的关键区域，该地区的降雨变化不仅影响着世界上40%的人口和全球生态系统的稳定性，而且对全球水文循环和能量平衡也起着十分重要的作用。

这一气候变化重建结果显示，过去2700年中印—太北部降雨呈长期下降趋势，与北半球热带其他地区古水文记录一致，而与南半球热带地区的降雨增加趋势相反，这体现了轨道尺度上夏季太阳辐射对南北半球热带降雨变化反相位关系的驱动作用。

该研究的另一个亮点是揭示了14世纪末到15世纪初的极端降水事件对吴哥文明消失的可能影响。研究结果显示，这一时期该区域存在持续数十年的极端降雨事件，这和柬埔寨吴哥文明时期城市排水系统的冲毁时间一致。此外，研究还发现在中世纪暖期和现代暖期，中印度—太平洋地区百年—十年尺度干旱的空间模式与厄尔尼诺事件发生时期类似。

研究人员通过泰国南部和印尼石笋重建的中印—太北南地区降雨的差值，构建了一条新的热带辐合带（ITCZ）南北移动指数记录。结果显示，过去2000年ITCZ存在整体南移的趋势，这和南北半球副热带地区温度梯度有关。研究人员认为，20世纪以来热带北部地区的干旱趋势类似于历史暖期，主要由厄尔尼诺活动的增强以及ITCZ南移导致；而人类活动对北半球热带地区降雨变化造成的影响，尚未改变自然变化的趋势。

相关论文信息：*https://doi.org/10.1073/pnas.1903167116*

中科院昆明植物所等

我国豆科植物与传粉者近年来“关系稳定”

本报讯（见习记者高雅丽）近日，中科院昆明植物研究所研究员杨永平带领的植物基因组演化与基因功能发掘团队与中科院植物研究所、云南师范大学、云南大学和西藏自治区高原生物研究所合作，通过豆科植物标本数据，揭示了植物种子产量与传粉者的相互关系，研究成果在线发表于《新植物学家》。

传粉者在生物多样性的维持、陆地生态系统的服务功能和农业生产等方面发挥着重要的作用，但大量的研究表明，传粉者（特别是蜂类）的多样性和丰富度在过去百年间存在显著的降低趋势，而这种降低趋势究竟如何影响野生植物的种子产量仍然缺乏证据。

研究组发现，豆科植物的标本数据是研究种子数量历史动态变化的理想材料，他们采用传统的豆科分类系统，提出了一个较为合理的假设：具有专化传粉系统的蝶形花亚科植物的种子数量可能表现出降低的趋势，而具有较为泛化传粉系统的含羞草亚科和云实亚科植物的种子数量可能表现出降低、不变或者增加的趋势。

为检验这一假设，研究人员查阅了存放于中科院昆明植物所标本馆和中科院植物所的2万余份豆科植物标本，记录了每份含果实标本一个果荚内的种子数量，最终共获得了109种豆科植物4637个关于种子数量的数据。这些标本最早采于1900年，最新采于2013年，时间跨度超过30年的物种为101个。

统计结果表明，只有13个物种的种子数量在近些年表现出了显著的变化趋势，其中9个物种的种子数量显著增加；3个亚科植物的种子数量并没有表现出一致的变化趋势，而在蝶形花亚科中，种子数量增加的物种数要高于降低的物种数。研究结果表明，在我国豆科植物与传粉者的相互关系在近些年来并没有被严重干扰。

相关论文信息：*https://doi.org/10.1111/nph.16119*

上海微系统所

制备高灵敏度石墨烯基可穿戴纤维

本报讯 中科院上海微系统与信息技术研究所研究员丁古巧课题组通过结构化设计减少石墨烯与高分子接触面积，制备了高分子纳米球修饰的石墨烯多孔网络纤维，从而提高了石墨烯基纤维的灵敏度。相关研究成果近期在线发表于《先进功能材料》。

石墨烯—高分子复合纤维具有质量轻、信号噪声低、能耗低等优点，可用于电阻型应变传感器。对于感知心脏跳动、脉博和眨眼等人体局部微小形变，其应变在0~10%范围，需要传感器件在发生形变时结构和电阻变化大，即高灵敏度，从而实现对信号的精确捕捉和对不同动作状态的准确辨析。然而，石墨烯基纤维传感器在0~10%应变范围内的灵敏度一般较低（GF~0.1-50），如何提高石墨烯基纤维传感器在小应变范围内的灵敏度是一个难题。

研究人员利用石墨烯/聚偏氟乙烯/聚氨酯DMF体系在水相的相分离过程，制备了高分子纳米球修饰的石墨烯多孔网络纤维，这种结构大幅增强了该纤维在发生形变时石墨烯片层之间的结构变化，从而实现石墨烯基纤维灵敏度的显著提高。其灵敏度因子值在0~5%应变时为51，在5%~8%应变时达到87，通过编织集成，他们进一步验证了该纤维在人体重要信号收集时的准确性和对不同动作状态分析的可行性。同时，这种新型石墨烯基纤维传感器最低形变检测限达到0.01%，较好的应变—电阻线性关系可保证在信号后处理上的准确性，>6000次的循环寿命有利于实际应用的稳定性。

将此纤维编织进纱布并作为眼罩，可实时监测眼球的转动等信息，未来可用于眼疾病人的监测和睡眠监测；将该纤维集成到创口贴中并贴到手腕处，能够识别手腕脉搏，且脉搏信号能够清晰表现脉搏的不同信号；该纤维也可编入手套，对不同的手弯曲进行感应，表明其对于动作信号的准确把控。 （柯讯）

相关论文信息：*https://doi.org/10.1002/adfm.201903732*