

《科学》杂志发文,25位世界顶尖科学家呼吁——

## 采取更强有力行动防范AI风险

## 今日视点

◎本报记者 张梦然

当以后的人们为人工智能(AI)撰写一部编年史时,这可能是一个里程碑。

因为这是人类第一次由一个庞大的国际顶尖专家组,就先进AI的风险推进全球政策制定。

人们已清晰认识到:我们离不开AI,亦需要大力推动AI技术发展,但我们必须有能防范其中的风险。

自6个月前在英国布莱奇利公园举行第一届AI安全峰会以来,世界各国领导人承诺以负责任的方式管理AI,但是AI专家认为这依然不够。随着第二届AI安全峰会(5月21日至22日)在韩国首尔举行,25位世界顶尖AI科学家呼吁,全世界应对AI风险采取更强有力的行动。

发表于《科学》杂志的一份专家共识文章中,专家组概述了全球应采取的紧急政策优先事项。文章作者之一、英国牛津大学工程科学系教授菲利普·托尔表示:“在上届AI峰会上,世界一致认为我们需要采取行动,而现在是时候将模糊的建议转变为具体的承诺了。”

## 安全性研究严重匮乏

这25位全球顶尖的AI及其治理专家来自中国、美国、欧盟、英国和其他AI技术强国,其中包括图灵奖获得者以及诺贝尔奖获得者。这也是首次由国际专家组就AI风险的全球政策制定达成一致。

文章指出,全球领导人需认真考虑,在未来十年内可能开发出超越人类能力的通用AI系统。他们表示,尽管世界各国政府一直在讨论先进AI,并尝试引入初步指导方针,但这与许多专家期望的快速、变革性进展仍有距离。

目前关于AI安全的研究严重缺乏,只有1%—3%的AI出版物涉及相关安全问题。此外,全球没有适当的机制或机构来防止滥用和鲁莽行为,包括如何使用能独立采取行动和自行实现目标的自治系统。

## 在快速发展中防范极端风险

文章建议,各国政府需建立可快速行动的AI监管机构,并为这些机构提供资金。美国AI安全研究所目前的年度预算为1000万美元,听起来并不少,但相比之下,美国食品和药物管理局(FDA)的预算达67亿美元。

AI领域需要更严格的评估并制定可执行的措施,而不是依赖一些不甚明确的模型评估。应要求AI研发企业优先考虑安全性,并证明其系统不会造成伤害,AI开发人员要承担起证明安全性的责任。

人们还要有一个与AI系统带来的风险水平相称的缓解办法,解决之道是制定一个“自动”政策,当AI达到某些能力里程碑时“自动触发”——如果AI发展迅速,这一严格的要求就会自动生效;如果进展缓慢,要求将相应放缓。

对于能力超凡的未来AI系统,政府必须作好带头监管的准备。这包括许可开发,限制其在关键社会角色中的



图片来源:视觉中国

自主权,停止部署部分能力,强制执行访问控制,以及对国家级黑客采取强有力的信息安全手段,直到人们准备好足够的保护措施。

## 谨记AI不是玩具

不列颠哥伦比亚大学AI教授、强化学习领域专家杰夫·克伦纳表示,太空飞行、核武器和互联网等技术在数年内就从科幻变成了现实,AI也不例外。“我们现在必须为那些看似科幻小说的风险做好准备。”

现在,AI已经在黑客攻击、社交操纵和战略规划等关键领域取得了快速进展,并可能很快带来前所未有的关于“控制”的挑战。AI系统可能会为了实现某个“不良目的”而骗取人类信任、获取资源并影响关键决策者。为了避免

人为干预,它们还可以在全球服务器网络上复制自身算法。

在这种情况下,大规模网络犯罪、社会操纵和其他危害会因此迅速升级。而在公开冲突中,AI系统能自主部署各种武器,甚至包括生物武器。因此,当AI的进步不受控制,很可能最终导致大规模生命损失、生物圈破坏、人类边缘化甚至灭绝。

美国加州大学伯克利分校计算机科学教授图尔特·卢塞尔表示:“此次是权威专家的共识文章,其呼吁政府严格监管,而不是自愿制定一些规范行业的行为。”

“企业可能抱怨说满足这些法规太难,但所谓‘法规扼杀创新’的说法是荒谬的。”卢塞尔这样说,“是时候认真对待先进AI系统了,它们不是玩具。在我们了解如何确保它们的安全性之前,轻易增强它们的能力是完全鲁莽的行为。”

科技日报北京5月21日电(记者刘震)美国辛辛那提儿童医院科学家领导的研究团队,成功开发出全球首个包含全功能血脑屏障的人类“迷你”大脑。这一成果有望增进科学家对多种脑部疾病的理解并改善治疗方法,如中风、脑血管疾病、脑癌、阿尔茨海默病、亨廷顿病、帕金森病及其他神经退行性疾病。相关论文发表于最新一期《细胞·干细胞》杂志。

大脑内的血管被一层紧密堆积的细胞覆盖,构成血脑屏障。这严格限制了从血液进入中枢神经系统的分子的大小。功能正常的血脑屏障能有效阻挡有害物质,同时允许必要营养物质进入大脑,从而维持大脑健康状态。但该屏障也阻止了许多潜在的有益药物进入大脑。此外,当血脑屏障发育不良或出现破裂时,会导致或恶化多种神经系统疾病。

此前,未有科研团队成功构建出具备全功能人脑血管屏障的大脑类器官。在最新研究中,科学家成功地将一个直径3—4毫米的大脑类器官与一个直径约1毫米的血管类器官融合,形成一个直径略超4毫米(约芝麻籽大小)的球体。这个新型结构体被称为“血脑屏障(BBB)组装体”。

该组装体由特定脑细胞的干细胞培育而成,因此能够反映可能导致血脑屏障功能障碍的基因突变等条件。研究团队采用这些患者的干细胞,成功构建了能够精准再现大脑海绵状畸形关键特征的组装体,为深入探究该脑疾病的分子与细胞病理学机制提供了崭新视角。

研究团队认为,该模型具有广泛应用前景:根据患者独特的基因和分子特征为其量身定制治疗方案;为多种神经血管疾病建模;更准确地快速分析潜在脑药物是否可以有效穿越血脑屏障;支持向大脑提供基于免疫的治疗等。

近年来,类器官技术不断进步,其中脑类器官对于研究和模拟人类神经和血管疾病发挥了关键作用,在脑发育研究、脑疾病建模和新药开发方面展示出巨大潜力。然而,普通的脑类器官已经面临培养周期长、成本高的问题,拥有全功能人脑血管屏障的脑类器官更是一直没能出现,这都限制了它们在生物医学研究中的应用。此次的成果,可以说为相关领域填补上了一块缺口。

拥有全功能血脑屏障的「迷你」大脑问世  
有望增进对多种脑部疾病的理解

总编辑 卷点  
环球科技24小时  
24 Hours of Global Science and Technology

## 二维材料可在室温下保存量子信息

科技日报北京5月21日电(记者张佳欣)英国剑桥大学卡文迪许实验室科学家首次发现,层状二维材料六方氮化硼(hBN)中的“单原子缺陷”可以将量子信息在室温下保留几微秒。相关论文发表在《自然·材料》杂志上。这一发现意义重大,因为能够在环境条件(室温)下拥有量子性质的材料十分罕见,此次发现还凸显了二维材料在推进

量子技术方面的潜力。

在hBN中,单一的“原子缺陷”在环境条件下表现出自旋相干,并且这些自旋可以用光来控制。自旋相干性指的是一种电子自旋,能够随时间推移保留量子信息。

此次研究结果显示,如果将特定的量子态信息传输到电子自旋上,这些信息就会被存储约百万分之一秒,这使该系统成为一个非常有前途的量子应用

平台。虽然百万分之一秒很短,但难得的是这个系统不需要特殊条件,就可在室温下存储自旋量子态。

hBN是一种超薄材料,由堆叠在一起的单原子层组成。这些层通过分子间作用力结合在一起。但有时,在这些层内部会存在“原子缺陷”。就像困在晶体中的分子一样,这些缺陷可以吸收和发射可见光范围内的光,具有光学跃

迁,并且可以充当电子的局部陷阱。由于hBN中的这些“原子缺陷”,科学家现在可以研究这些被捕获电子的行为,如电子的自旋特性,这种特性允许电子与磁场相互作用。真正令人兴奋的是,研究人员可以在室温下利用这些缺陷中的光来控制电子自旋。

这一发现为未来的技术应用铺平了道路,特别是在传感技术方面。

## 纳米药物促进动物糖尿病伤口愈合

科技日报北京5月21日电(记者张佳欣)美国西奈山伊坎医学院研究人员设计了一种再生医学疗法,能加快糖尿病伤口愈合。该疗法使用载有基因指令的微小脂肪颗粒来缓解炎症。结果表明,该疗法可以针对引起问题的细胞,并减少小鼠皮肤损伤模型中的肿胀和有害分子。相关论文20日在线发表于《美国国家科学院院刊》。

糖尿病伤口通常对常规治疗具有耐药性,对全球数百万人构成严重健康风险。人体内的巨噬细胞原本应当发挥积极的免疫作用,然而在某些情况下,它们却可能引发炎症反应。这种炎症会损害其他细胞,使伤口更难正常愈合。

新疗法使用装载有编码IL-4的

RNA脂质纳米粒(LNP),能够靶向功能失调的巨噬细胞,同时减少糖尿病伤口中炎症和“活性氧”(ROS)的产生。IL-4是一种被称为细胞因子的细胞间信号蛋白。

ROS分子在体内各种代谢过程中自然产生,并在细胞信号传导和免疫反应中发挥作用。然而,过量ROS会导致氧化应激,对细胞、蛋白质和DNA造成损害。这种压力与炎症和衰老等各种疾病和症状有关。

临床前模型证明,新疗法能够将促炎巨噬细胞重新编程为修复性巨噬细胞,从而改善伤口愈合效果。

尽管结果令人鼓舞,但研究人员强调,还需要进行严格的随机对照临床试验,以确保其应用于人体的安全性和有效性。

## 生酮饮食或加速器官衰老

科技日报讯(记者刘震)美国科学家开展的小鼠实验显示,生酮饮食会导致受损细胞在小鼠心脏、肾脏、肺部和大脑内积聚,加速器官衰老,增加其罹患心脏病、癌症等疾病的风险。相关论文发表于17日出版的《科学进展》杂志。

生酮饮食是一种遵循低碳水化合物、适量蛋白质、高脂肪的饮食方式。尽管许多人为了减肥和控制血糖而采用了这种饮食方式,但其健康影响研究结果喜忧参半。

为了解更多信息,得克萨斯大学健康科学中心团队给6只小鼠喂食生酮饮食3周。这些小鼠90%以上的热量来自脂肪,不到1%来自碳水化合物。对照组则遵循标准饮食,其中17%的热量来自脂肪,58%来自碳水化合物。

团队随后分析了小鼠的心脏、肾脏、肝脏和大脑组织样本,寻找衰老细胞。当细胞受损无法正常发挥作用时,会变成衰老细胞。这些细胞会在身体组织内游荡,释放毒素,从而引发炎症。

结果显示,与标准饮食小鼠相

比,生酮饮食小鼠器官中的衰老细胞明显更多。例如,生酮饮食小鼠肾脏内细胞衰老标志物的数量是正常饮食动物的4倍。鉴于衰老细胞的数量会随着年龄增长而增加,这些发现表明,生酮饮食可能会加速器官衰老,从而增加罹患心脏病、癌症、II型糖尿病等疾病的风险。研究还显示,让小鼠重新遵循标准饮食会减少衰老细胞的数量。

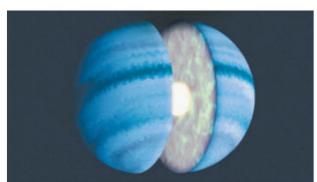
研究团队指出,虽然生酮饮食在某些情况下可能有益,但并不适合所有人。



生酮饮食是一种遵循低碳水化合物、适量蛋白质、高脂肪的饮食方式。  
图片来源:《新科学家》网站

## “棉花糖”行星为何如此“蓬松”

## 科普园地



WASP-107b艺术概念图显示了行星气体包层内的湍流大气混合。  
图片来源:约翰·霍普金斯大学

科技日报北京5月21日电(记者张佳欣)太阳系外有一颗温暖的气态巨行星WASP-107b,外表看起来像棉花糖一样,它为何如此“蓬松”?根据美国国家航空航天局(NASA)詹姆斯·韦布太空望远镜收集的数据,并结合之前哈勃太空望远镜的观测结果,美国两个独立研究团队给出了答案。他们的最新研究显示,在WASP-107b中蕴藏着极少量甲烷和一个超大核心。这表明其内部肯定比之前估计的要热得多,核心质量也要大得多。相关论文发表在20日《自然》杂志上。

WASP-107b是一颗巨大的,类似海

王星的系外行星,绕着一颗距离地球约200光年的恒星运行。它被灼热的大气层包裹,这些大气像棉花糖一样蓬松。其构造独特:体积是木星的3/4,质量却不到木星的1/10,是已知密度最小的行星之一。尽管它含有甲烷(地球生命的组成部分),但由于靠近主星且缺乏固体表面,科学家认为该行星并不宜居,但它可能蕴藏着有关行星演化的重要线索。

研究人员推测,行星过度膨胀是由于其内部热源造成。这颗行星有一个炽热的核心,正在改变其更深处气体的化学性质,并推动强烈的对流混合,从内部冒出气泡。

WASP-107b的巨大半径、广阔的大气层以及独特的轨道,使其成为进行透射光谱学观测的理想对象。透射光谱学是一种根据气体如何影响星光来识别系外行星大气中各种气体的方法。韦布望远镜超乎寻常的灵敏度和测量穿过系外行星大气光线的能力,为取得这些重要成果提供了可能。此次研究可能解释了数十颗低密度系外行星蓬松的原因。

这项研究还标志着对系外行星核心质量的首次测量,或为未来行星大气和内部研究奠定基础。此外,这也是寻找太阳系外宜居世界的关键一步。

## 外国专家点赞——

## 中国为生物多样性保护提供全球协作样本

◎本报记者 毕炜梓 陈春有

中国云南自然资源十分丰富,生物多样性数量居全国之首,素有“植物王国”和“动物王国”的美誉。在今年5月22日第24个国际生物多样性日来临之际,科技日报记者走进云南西双版纳、大理两地,采访三位在华工作的外国专家,他们与记者分享了对生物多样性的认识,赞同中国在生物多样性保护方面的举措和取得的成绩,并对未来全球促进这项共同事业充满期待。

## 在发展中兼顾生态

中国科学院西双版纳热带植物园爱尔兰籍植物生态学家唐力森向记者介绍,生物多样性可以分为三个层次:一是生物物种的多样性;二是生物的功能多样性;三是生态系统服务,即自然生态系统为人

类提供的各种经济和社会价值服务。

“保护生物多样性与人类命运息息相关。随着人们生活水平提高,生态保护受到越来越多重视。”定居在中国云南大理的美国旅行家、作家布莱恩·林登向记者表示。他1984年首次来到中国留学,是中国生态文明发展的见证者。他认为,经济发展和生态保护对于提高人民生活水平同等重要,但难得的是平衡两者之间的关系。

“在发展中兼顾生态是实现可持续发展的的重要途径。受益于绿色发展政策,人类得以在发展的同时,守护更多栖息地,实现与自然的和谐共生。”中国科学院西双版纳热带植物园西班牙籍首席研究员康木飒说。

## 中国是亚洲象的安全港

近年来,中国加快国家公园体系建设步伐,设立三江源、大熊猫、东北虎

豹、海南热带雨林、武夷山等多个国家公园,用以保护大尺度生态过程以及维护这一区域的物种和生态系统特征。

康木飒说,国家公园建设正是加强生态保护的主动实践。“当前,西双版纳国家级自然保护区积极推进升级为国家公园。25年来,保护区内仅大象数量就翻了一番,这对深入科学人象关系和大象行为习性,以及科学制定实施大象保护策略至关重要。”

“对于亚洲象来说,中国是个安全港。”康木飒长期关注亚洲象保护和热带雨林生态,他认为中国制定的亚洲象保护措施非常完善。

他表示,在共建“一带一路”倡议框架下,中国帮助共建国家发展基础设施。而这种模式同样适用于生态保护。中国正与周边国家共同推动保护区建设,打造良好生态环境。特别是在西双版纳与缅甸、老挝两国的边境地区

开展了联合监测与保护,堪称跨境生态保护的典范。

## 国际合作不可或缺

今年国际生物多样性日的主题是“生物多样性 你我共参与”。为更有效参与生物多样性保护事业,唐力森建议公众通过听讲座、看科普、走进大自然等方式,了解人类当下面临的生态问题。

林登也鼓励大众沉浸式亲近大自然,与丰富的自然资源互动。

在唐力森看来,除民众的个人努力外,国际合作在生态保护方面的作用不可或缺。“没有一个国家是一座孤岛,生态系统不会止步于边境线。”唐力森说,“保护边境线上的珍稀濒危野生生物还需各国共同努力。”

这不仅是为了生物,也是为了人类自己。