安全可靠：新一代人工智能面临核心大考

□ 科普时报记者 陈杰

人工智能技术一直是当前科技领域的热点，不论是从图像分析到自然语言理解，再到科学领域。得益于神经网络的“深度”变革，以大量数据和算法加持的算力去处理复杂的数据已成为日常，而这在过去是无法想象的，背后则是人工智能技术突飞猛进发展的必然结果。

在刷脸、自动驾驶、智慧家庭等各种人工智能应用场景纷纷落地日常生活的当下，从消费者的视角来看，人工智能已经让生活变得更美好了，但这一技术是否已经足够强大了呢？学术界的答案却恰恰相反——当前的人工智能技术在安全可靠方面其实完全不够！

视觉中国供图

人工智能并没有那么可靠

在日前召开的首届“青年科学家50²论坛”上，清华大学教授朱军对正在高速发展的人工智能产界泼出一盆冷水：今天的人工智能远远算不上完美。首先是深度学习仍然面临着鲁棒性差的问题，比如很多手机都会配备人脸识别方案，但通过一副打印了特殊纹理图案的眼镜，就能轻松解锁很多手机；再说到安全需求更高的自动驾驶系统，黑客们只需要在标识牌上加上特定图案，就能让系统将限速标识识别为停止标识，导致致命事故的发生……

这些看似非常符合人们直觉或者是常见的问题，放到很智能的系统面前，却会出现错误的结果，背后的原因是什么？

“这里面一个关键的问题，是现在的人工智能技术‘不知道自己不知道’。”朱军表示，人类都能够遵循“知之为知之，不知为不知，是知也”的道理，但当前的人工智能主流技术，绝大部分都欠缺这种能力。比如在一个很多猫和狗照片的比对系统中放一张人工智能还不知道的苹果照片，让人工智能来分析这是什么，结果答案要么是猫，要么是狗，而这取决于这张苹果照片和猫、狗哪个更相似一些。

不用怀疑人工智能的学习能力，但确实缺乏“知道自己不知道”的能力。学术界认为，一个人工智能系统如果知道自己“不知道”，其实代表着其有更高的智能水平。

“具体来说，在现实世界里存在两种不确定性会影响到人工智能对此的认知，一种是比较直观的环境和数据的不确定性导致的信息采集不完全；另一种则是更抽象的、更难以被感知的信息，这种模型的不确定性，让大数据的模型存在歧义，让人工智能在训练时也许表现得非常好，但是测试上就会发现性能距离很大。”朱军表示，目前的解决思路是更深入地研究贝叶斯智能理论，解决如何定量、客观地计算这种不确定性，得到最优的结果。推动贝叶斯智能理论研究的深化，则需要产学研界的共同努力。

安全是必须跨越的鸿沟

人工智能的概念产生最早可追溯到1956年的达特茅斯会议，但直到近十年来这一领域才真正地深刻改变了大众生活，甚至攻克了多种人类世界的顶级挑战，比如战胜围棋顶级选手、预测蛋白质结构、击败人类飞行员等等。但目前来看，当前的人工智能技术不仅在可靠性方面存在很大问题，安全和可信方面也开始进入瓶颈期。

8月3日，在首届全球数字经济大会的人工智能产业治理主题论坛上，中国科学院院士张钹抛出了一个观点：当前人工智能的安全治理是迫在眉睫的，随着技术潜入到“深水区”，研究发现，人工智能的算法存在一些根本性的问题，本身就带有不安全性，容易受到攻击。“人工智能技术所面临的安全问题是非常特殊的，其问题不在于设计，而是源自算法本身的不安全性。这种算法的不安全性，是由第二代人工智能所引起的。”

“学术界已经开始思考技术下一步的发展方向，新一代人工智能的发展路径是融合第一代的知识驱动和第二代的数据驱动的人工智能，同时利用知识、数据、算法和算力等4个要素，建立新的可解释和鲁棒的人工智能理论与方法，发展安全、可信、可靠和可扩展的人工智能技术。”朱军表示，“安全可靠“作为第三代人工智能的核心发展目标逐渐成为共识，数据与算法安全也成为学界和业界人士重点关注的研究主题之一。

张钹认为，发展第三代人工智能是一项长期任务，技术的路径也非常艰难。“当前，应对人工智能进行安全可信方面的治理，防止它被误用、制止它被滥用，这需要从技术创新层面发力去‘治本’，同时也要从法律法规、伦理规范、行业共识等不同层面去‘治标’。”

智能网联汽车数据安全无小事

□ 科文

随着智能汽车产业的爆发式增长，智能网联汽车已经逐渐走入我们的生活。然而，智能网联汽车越来越像是一台装有四个轮子的“手机”，集成了摄像头、雷达、测速仪、导航仪等。由此，远程控制、数据窃取、信息欺骗等手机安全问题同样会出现在汽车上，也会产生诸如驾驶自动化系统随机故障、功能不足等引发的道路交通安全问题，以及在线升级（又称OTA升级）改变车辆功能、性能可能引入的安全风险，甚至直接危及人身安全。

所谓智能网联汽车，包括新能源智能汽车，带有一定辅助驾驶、自主驾驶及信息搜集相关智能化功能的汽车都可以纳入其中。数据是智能网联汽车很重要的组成部分，它包括三方面：个人信息相关的数据、行驶数据和图像采集数据，这些都会产生大量的安全风险问题。

为推动智能网联汽车产业高质量发展，工信部8月12日发布《关于加强智能网联汽车生产企业及产品准入管理的意见》（以下简称《意见》），要求加强汽车数据安全、网络安全、软件升级、功能安全和预期功能安全管理，保证产品质量和生产一致性。这一系列新要求的提出，将有望减少用户的相关担忧。

《意见》从加强数据和网络安全管理、规范软件在线升级、加强产品管理、保障措施等方面提出11项具体意见。

在加强数据和网络安全管理方面，《意见》要求企业落实主体责任，加强汽车数据安全、网络安全、软件升级、功能安全和预期功能安全管理，保证产品质量和生产一致性。《意见》明确在境内运营中收集和产生的个人信息和重要数据应按照有关法律法规规定在境内存储。需要向境外提供数据的，应通过数据出境安全评估。

在规范软件在线升级方面，《意见》要求企业生产具有在线升级（又称OTA升级）功能的汽车产品的，应建立与汽车产品及升级活动相适应的管理能力，具有在线升级安全影响评估、测试验证、实施过程保障、信息记录等能力，确保车辆进行在线升级时处于安全状态，并向车辆用户告知在线升级的目的、内容、所需时长、注意事项、升级结果等信息。

在加强产品管理等方面，《意见》要求企业生产具有驾驶辅助和自动驾驶功能的汽车产品的，应明确告知车辆功能及性能限制、驾驶员职责、人机交互设备指示信息、功能激活及退出方法和条件等信息。企业生产具有组合驾驶辅助功能的汽车产品的，应采取脱手检测等技术措施，保障驾驶员始终在执行相应的动态驾驶任务。

《意见》还规定要加强自动驾驶功能产品安全管理。企业生产具有自动驾驶功能的汽车产品的，应当确保汽车产品至少满足四项要求。

为了有效复盘事故，汽车产品应具有事件数据记录系统和自动驾驶数据记录系统，用于事故重建、责任判定及原因分析等。

此外，企业应当确保汽车产品具有安全、可靠的卫星定位及授时功能，可有效提供位置、速度、时间等信息，《意见》鼓励支持智能网联汽车产品接受北斗卫星导航系统信号。

据了解，工信部后续将重点针对具有自动驾驶功能的智能网联汽车产品编制出台相关技术规范，推动《意见》落实。

行业会讯

“工业互联网+AI”专业赛启动

8月13日，由中国工业互联网研究院、大连金普新区管委会联合主办的“工业互联网+人工智能”专业赛启动报名，线下决赛将于9月底在大连金普新区举行。

中国工业互联网大赛“工业互联网+人工智能”专业赛是致力于我国工业互联网与人工智能领域融合创新的全新赛道。大赛将进一步聚焦人工智能技术在工业互联网上的创新应用，促进工业互联网与人工智能全面创新、链接、融合。

针对此次赛事，中国工业互联网研究院将发挥主办方优势，重点遴选和培育一批具有发展潜力和推广价值的工业互联网与人工智能解决方案。大连金普新区将发挥自身“中国工业百强区”的产业政策优势，助力优质项目落地应用，促进工业互联网与人工智能协同发展。

从大赛组委会了解到，本次大赛聚焦国内重点工业场景，将在装备制造、石油石化、生物医药、汽车整合和零部件等行业广泛征集遴选一批优质解决方案和产品，以促进人工智能技术从多维度为工业互联网各类场景赋能，全面助力工业互联网应用价值提升。

新IT进入高速增长期

数字经济高速发展的趋势下，对IT需求的增长也进一步加速焕新传统IT行业，“硬件战”红海走向“转型服务战”蓝海。据IDC预测，2021财年中国智能产品市场规模将达到1050亿美元；智能基础设施市场规模将达到430亿美元；智慧服务市场规模将达到1090亿美元，“3S”新业务带来更多空间，迈入可持续增长周期。

8月11日，联想中国区公布的2021/2022财年一季度报显示，联想中国区季度营业额年比年提升超过50%，PC业务实现超预期增长，销量市场份额达到40%，以PC整体市场2倍速增长。“3S”转型新业务占总营收比例超过1/4，季度总营收年比年增长71%。

“3S业务市场空间比PC更具想象空间”，联想集团执行副总裁兼中国区总裁刘军在内部信中表示，联想中国智能产品、智能基础设施和智慧服务业务的收入分别实现了173%、58%和43%的同比增长，3S转型业务成为打开中国区未来发展空间的核心赛道。

专家表示，未来深入智能化转型将迎来发展新势能，为千行百业智能化转型企业提供源源不断的驱动力，也为行业和社会变革打开新的蓝海航道。

TikTok新增青少年心智推送

近日有外媒报道称，TikTok近期列出了一系列措施，以改善其平台上青少年的安全和隐私，也许所有变化中最有趣的是其对推送通知的态度改变。TikTok称，希望帮助青少年对社交媒体形成一种更健康的态度，以便不影响他们的睡眠。

根据儿科专家和青少年福祉倡导者的建议，如果用户年龄在13-15岁，TikTok将在晚上9点后停止发送推送通知，而16-17岁的用户将在晚上10点停止收到通知。希望这一变化能让青少年减少影响，而不是必须对每一个弹出通知做出反应。

据悉，TikTok正在引入的另一个变化是，允许16-17岁的青少年了解下载的工作方式。默认情况下，下载青少年用户视频的选项是禁用的，但可以手动启用。在允许他人下载视频之前，用户必须确认他们的选择。对于16岁以下的用户，允许他人下载其视频的选项被完全禁用，不能开启。此外，16岁以下的青少年在去发布视频时，会看到弹出窗口，询问应该允许谁观看该视频。

最后，TikTok已默认禁用16-17岁用户的私信，尽管可以打开。这个新的默认值建立在之前的变化之上，即完全禁止16岁以下账户的私信。

百度大规模校招揽AI人才

8月12日，百度启动2022届校园招聘，面向海内外2022届毕业生，开放技术、产品、设计、政企解决方案、专业/职能五大类型岗位，是百度史上规模最大的校园招聘。

与去年相比，今年百度招聘岗位更加多元，新增芯片开发、人力资源、整合营销、公众沟通、业务研究与分析等大量岗位，还在北京、上海、深圳之外开放了广州、成都等城市的全新就业机会。同时，面向管培生、技术和产品精英的高端校园人才项目也同步启动，计划打造AI领域的未来之星。

除五大类型岗位之外，百度今年也加强了对高端AI校园人才的招聘和培养力度。其中，去年正式发起的“管理培训生”计划将再次启动，2022届优秀MBA及顶尖应届毕业生均可报名，百度董事长兼首席执行官李彦宏及集团各大副总裁将亲自选、亲自带，旨在培养AI时代科技创新企业新一代年轻管理者；面向技术和产品岗位的AIDU技术精英与AIDU产品精英人才招募同步进行，将重点网罗AI技术研究和项目上有优异表现、或拥有出色产品意识和用户思维的应届毕业生，提供高精尖的团队指导、职业培训与颇具竞争力的薪资待遇。

孪生技术助力行业数字化转型

8月13日，在“Tech 2021数字中国技术年会”上，神州信息携手华为、中青智造、智慧足迹、优锘科技发布园区智能体解决方案，这一解决方案意在利用数字技术赋能智慧场景，实现业务在数字世界的孪生形态，帮助企业高效决策，便捷管理。

园区智能体是智慧园区、智能制造、智慧交通、智慧物流等现实业务场景在数字世界实现镜像的统称。神州信息信创BG技术总监袁晔栋表示，园区智能体通过构建业务全数字化、系统全联接、数字全融合的一个完整的生态系统，形成数字资产、数据仓库，并通过孪生技术，在数字世界展现整体的业务环境，使业务变得可控、可管、可视，这也是园区智能体建设的核心意义，是行业数字化转型的最佳实践。

据了解，神州信息、华为、中青智造、智慧足迹、优锘科技优势汇聚，形成五大基础能力，保障行业客户园区智能体能够快速实现接入和运营。首先是设备接入能力，神州信息物联网可实现千万级各类传感器、终端的有线接入，设备接入套件降低了设备接入成本，提升了连接稳定性；其次是高并发能力，由华为云提供的极致算力支持千万级设备同时并发，并保证设备的稳定性；第三是数据分析与可视化能力，智慧足迹、优锘提供了数据实时计算引擎、数据算法模型以及专业的可视化套件；第四是安全能力，神州信息作为信创先锋军，从云、管、边、端全方位支持信创，并提供专业服务；最后是场景定制能力，中青智造构建定制化场景，优锘提供数字孪生。

现实的行业场景完成数字孪生后，将在园区智能体运营中心得以全面展现，包括每一个场景的实时运营，安全、性能状况等等，真正做到软件系统资产化以及可视、可控和可管。目前神州信息与生态伙伴已经构建了智慧枢纽、孪生工厂、智慧园区、智慧矿山、智慧社区、智慧城市、智慧能源、智慧物流和仓储管理等多样业务场景的数字孪生，实现了城市体检评估中心、交通出行大数据监测中心、广东制造园区、银行数据中心等多个园区智能体的项目落地。

在园区智能体的帮助下，管理者可以从最顶层的高度去看整个业务系统，产生新的业务逻辑，从而不断创新甚至是颠覆。此外，园区智能体的数字技术还可以减轻资源利用，减少资源浪费，实现绿色低碳。而从智慧港口、智慧医院、智慧银行到智慧交通，未来所有园区智能体都将一一汇集，逐步构建出城市的数字坐标图，形成真正的智慧城市。 （捷闻）

“.中国”域名得到广泛认可

□ 科普时报记者 李禾

随着中国元素、中国形象在世界上的影响力逐渐加大，以“.中国”为后缀的中文域名以其明确的“中国”形象得到众多用户认同。腾讯今年以来陆续开通解析了“QQ.中国”和“腾讯.中国”两个域名，其中“QQ.中国”解析到腾讯网，“腾讯.中国”解析到腾讯官网。这意味着，网民在浏览器地址栏中输入“QQ.中国”和“腾讯.中国”，就可快速登录腾讯网和腾讯官网。

互联网名称与数字地址分配机构（ICANN）ccNSO理事、国际互联网工程任务组（IETF）EXTRA工作组联合主席姚健康博士认为，中文域名是用中文开启互联网的钥匙，非常适合国人使用。从文化角度来讲，中文域名是在互联网上重要的文化标识，有利于传承和弘扬中国文化。从安全角度来讲，“.中国”顶级域名是我国在互联网上的中文国家顶级域名，由我国全权管理，安全有保障。从技术角度来讲，中文域名技术符合IETF国际技术标准，全球畅通。

中文域名拉近了企业与用户的距离，同时降低了中文用户接触、记忆与使用互联网功能的门槛。比如腾讯启用“.中国”域名，一方面，便于中文用户联想记忆，更易拉近中文用户群，方便访问；另一方面，中文域名能有效改善互联网入口匮乏问题。腾讯网启用中文域名举措，表明国内互联网行业对中文域名越来越重视。2020年，百度开通解析了“小度.中国”，为旗下智能装备平台小度提供中文互联网入口，以此树立中国品牌形象，增强中文用户认知。“当当网电子书.中国”“华大基因.中国”“中国移动.中国”“新華社.中国”“beijing2022.中国”等也都先后开通解析。随着国内知名企业和机构纷纷注册启用，“.中国”域名赢得了越来越多行业认同，其注册量在中文域名中排名第一。

自互联网诞生以来，英文域名一直占据着统治地位，2000年起，在中国互联网络信息中心（CNNIC）、中文域名协调联合会（CDNC）等机构的共同努力下，中文标识逐渐出现在域名体系中。2010年，“.中国/.中國”域名成为全球首个被正式写入全球互联网根域名系统的纯中文顶级域名。至此，网民可方便使用中文域名访问互联网。与此同时，中文域名的应用环境也在逐年完善，在以CNNIC为代表的中文域名社群不懈推动下，国内各大主流浏览器已全面支持中文域名访问；OFFICE2016与FOXMAIL等应用均支持多语种电子邮件技术（EAI）标准和中文域名内容显示超链接。腾讯、网易、电信等国内邮件服务厂商也在为个人及企业用户体验中文域名电子邮件进行了尝试和努力。

2021年4月30日，国家标准化管理委员会批准了中文域名总体技术要求国家标准立项。

姚健康说，ICANN将中文域名及其他多语种域名的推广应用作为今后几年重点推动的工作，可以说中文域名前景十分广阔。

数字化技术加速电力“新基建”落地

我国3060“双碳”目标的提出，将进一步加快能源低碳转型，推动我国能源体系向“一主两翼”发展格局转变。电力系统的数字化转型是实现中国能源转型和“双碳”目标的关键路径之一，而电力“新基建”，是电力系统数字化转型的关键支撑，发挥着基础性和支柱性作用，是推动电力行业战略转型与高质量发展、支撑能源革命的重要载体。

日前，施耐德电气联合中国科学院科技战略咨询研究院、华北电力大学及中国能源网等研究单位，发布《电力“新基建”发展模式和路径研究》报告，结合“双碳”目标背景下，新基建支撑电力数字化转型的趋势和国际经验，对电力“新基建”进行了深入扎实的研究，为电力“新基建”高质量发展路径的探索提供了有益的启示和参考。

报告指出，电力“新基建”以电力大数据为驱动，以5G、物联网等新一代信息网络基础设施为基础，以与电力基础设施适配的专用信息网络基础设施建设为手段，推动云计算、大数据、人工智能等通用数字技术应用于电力物联网和电力应用平台，促进电力数据流动环节的源、网、存、算、用等设施与电力能量流动环节的源、网、储、荷等设施融合发展。

依托电力“新基建”开展电力数字化转型，能够加速数字信息技术与能源电力产业的深度融合，引导能源电力行业向数字化、网络化、智能化转型发展。一是通过电力“新基建”，融合发展覆盖发-输-配-用各环节的电力工业物联网，促进新能源消纳，助推“双碳”目标的实现。二是通过电力“新基建”，建设智慧能源系统，将新兴数字技术应用于海量数据的融合、分析与管理，推动能源供需革命。三是通过电力“新基建”，推动工业互联网与能源电力系统的融合，加速5G、物联网、大数据等创新技术的应用融合，推动数字技术在能源系统中的应用，助力能源技术革命。四是通过电力“新基建”，打造跨行业国际平台，依照“平台+生态”思路，构建互惠共赢能源生态圈，推动能源体制改革和国内国际双循环。

视觉中国供图

报告指出，“加快数字技术在电力系统的研发与应用”以及“加快推进用户侧数字化转型，鼓励灵活的用能模式”均纳入了电力“新基建”将要面临的重点任务。

（国俊）