

中国工程院院士李培根：

智能工厂是AI赋能新型工业化重要抓手



本报记者 宋婧

4月26日，在2024中关村论坛年会期间召开的智能制造创新发展论坛上，中国工程院院士李培根表示，智能工厂建设是人工智能(AI)赋能新型工业化的重要抓手，其核心在于智能机器人、数字孪生、大模型等关键技术在工业场景中的深层次渗透与应用。

第一个核心：智能机器人

发展至今，智能工厂建设已经离不开机器人。李培根指出，现在企业里装配线上的大量机器人都是在封闭环境下工作的，实际上并不是真正的智能机器人。未来制造业里的机器人可以走出仓储和制造环节，进入开放环境，这是今后一个重要方向，该场景的实现主要依靠自主移动技术(AMR)。

目前，车间里的自主引导小车大多是按照固定路径移动的，实际

上真正的自主移动技术没有这些限制，它基于地图定位技术，通过扫描作业环境并自主更新地图，无须辅助固定信标，对工作场地几乎没有改造需求，非常适合部署在复杂、动态的生产场景中。在日本牧野Makino车间中，采用了运用AMR技术的机器人，可以同时照看多台机床，上料、下料、装刀具、卸道具等操作都不在话下。

机器人的应用范围不应局限于

企业内部，而是应该能够应对各种操作条件、人口稠密的空间甚至其他自动化设备。“走出以往的受控仓储和制造环节，机器人将为各行各业带来发展机遇。”李培根说道。

比如，人形机器人具备高度自动化且适用于各种复杂环境，是智能工厂建设的一个重要探索方向。特斯拉计划率先在汽车装配过程中引入人形机器人。奔驰正在和机器人开发商Apptanik合作，

人形机器人具备高度自动化，是智能工厂建设的一个重要探索方向。

希望把人形机器人用到工厂中，用于装配零件。

“通过以类似人类的能力不需要改造环境就能应对复杂的人类环境，使用场景更广，能作为简单、重复、危险的劳动力替代，也能在适应非标准服务场景的同时满足情感需求和交互。”李培根表示。此外，机器人还具备学习功能，越用越聪明，工厂还可以通过机器人实现新老员工的技能传承。

供应链数字孪生体现在企业中，企业不仅要有好的供应链系统，还要有数字化供应链。

整来保证质量，保证效果。而供应链数字孪生体现在企业中，企业不仅要有好的供应链系统，还要有数字化供应链。而且不同于早就应用于供应链中传统、离线独立的仿真，基于数字孪生的仿真是实时的、动态的，能够与物理系统交互连接。

AI大模型对世界高阶认知已经远超人类，要重视其在智能工厂建设过程中的应用。

题、热和动力学误差、机床全生命周期误差控制等问题追根溯源，里面可能会有无数需要我们考虑的复杂的关系，比如机床部件、零件、精度、运动参数、转速等，这些关系在制造过程中很难考虑周全。”李培根坦言道。不过，大模型的出现或许可以帮助智能工厂洞察更为复杂的高阶关联。

一批人工智能与芯片项目集中亮相

在活动现场，一批新技术新产品首发首秀，展示了人工智能与高端芯片技术在汽车、网络安全、金融、法律等多领域的创新应用。

建设高性能芯片测试平台，是优化提升北京市集成电路产业生态的重要举措。在会上，高性能芯片测试平台正式发布，平台落地后，将支撑北京市集成电路企业优化设计，提升研发效率，加强协同创新，为集成电路产业高质量发展注入新动能。

作为支撑数字经济高质量发展的关键力量，下一步，人工智能和芯片产业应该如何发展？北京市经信局副局长顾瑾翔透露，北京将进一步构建算力供给体系、打造大模型行业应用新生态、提升大规模算力集群输出能力，加快推进首都新型工业化建设和全球数字经济标杆城市建设。

视频生成模型“Vidu”发布 满足长视频内容的生成需求

本报讯 记者宋婧报道：4月27日，在2024中关村论坛年会未来人工智能先锋论坛上，生数科技与清华大学联合发布了视频生成模型“Vidu”，其具备多镜头生成、时空一致性高等特点，性能全面对标国际顶尖水平，并在加速迭代提升中。业内人士表示，中国版“Sora”正在加速向我们走来。

据了解，Vidu采用了创新性的U-ViT架构，该架构融合了Diffusion(扩散)与Transformer技术，能够一键生成长达16秒、分辨率高达1080P的高清视频内容，满足长视频内容的生成需求。

根据现场演示的效果，Vidu能够模拟真实的物理世界，生成细节复杂、并且符合真实物理规律的场景，例如合理的光影效果、细腻的人物表情等。与此同时，它还具有丰富的想象力，能够生成真实世界不存在的虚拟画面，创造出具有深度和复杂性的超现实主义内容，例如在“画室里一艘船正在海浪中驶向镜头”的场景中，船和海浪的逼真效果体现了Vidu的强大生成能力。

此外，Vidu还能生成复杂的动态镜头，不再局限于简单的推、

拉、移等固定镜头，而是能够围绕统一主体在一段画面里就实现远景、近景、中景、特写等不同镜头的切换，包括能直接生成镜头、追焦、转场等效果，给视频注入镜头语言。

“值得一提的是，Vidu采用的是‘一步到位’的生成方式。”清华大学教授、生数科技首席科学家朱军表示，与Sora一样，从文本到视频的转换是直接且连续的，在底层算法的实现上是基于单一模型完全端到端生成，不涉及中间的插帧和其他多步骤的处理。除了在时长方面的突破外，Vidu在视频效果方面实现显著提升，主要体现在模拟真实物理世界、多镜头语言、时空一致性高、理解中国元素等方面。

公开资料显示，生数科技由清华系AI公司瑞莱智慧RealAI、蚂蚁集团和百度发起的BV百度风投联合孵化创立，目前已完成三轮融资，融资金额共计数亿元。其核心技术U-ViT架构由团队于2022年9月提出，早于Sora采用的DiT架构，是全球首个Diffusion与Transformer融合的架构，完全由团队自主研发。

未来人工智能探索涉及 三个关键方向

本报讯 记者张琪玮报道：4月27日，2024中关村论坛年会人工智能主题日未来人工智能先锋论坛在北京举办。在论坛上，微软全球资深副总裁、Microsoft AI亚太区总裁张祺表示，AI“大航海时代”已经到来，AI正在重塑每个企业和组织，赋能千行百业提升生产力和创造力。

张祺指出，AI新时代的到来，源于“用户界面”和“推理引擎”两项重大技术突破的结合。

在大语言模型的赋能下，用户界面已经从晦涩的编程语言发展为自然语言界面。“当下，任何人都可以通过自己的母语同AI进行高带宽的交流，进而延伸到图像、计算机视觉等方式进行自然的输入和输出。AI不仅能执行用户的指令，还可以帮助用户更好地梳理知识、展现创意……这是一个根本性的改变。”

同时，伴随着个人电脑、互联网、移动通信终端和社交平台的陆续诞生，对物理世界的数字化进程越发深入，要从海量数据中准确地定位、筛选、提炼需要的数据，就对推理引擎提出了更高要求。对此，张祺表示：“新的推理引擎可以理解传统搜索引擎的进阶版和智能体，大大缩短了用户为处理信息而耗费的时间，让他们可以专注于更

重要的工作上。”

在谈到人工智能未来的创新趋势时，张祺表示，未来人工智能的探索会涉及认知智能(Cognitive AI)、具身智能(Embodied AI)、科学智能(AI for Science)三个关键方向。

张祺介绍道，认知智能使AI能够协助人类进行更复杂、更智能的任务，甚至与多个AI代理进行协作，不断提高AI的功能表现和所服务对象的生产力；具身智能可以延展AI对物理世界的理解和影响，让世界变成提示符，让AI能够驱动自动驾驶汽车、机器人、混合现实装置等实体硬件，与现实生活产生更多交互；科学智能则旨在用AI促进科学进步。

其中，张祺重点强调了人工智能在科学智能方向的应用。他指出，作为推理引擎，AI能够辅助科研工作者在短时间内理解大量的数据，并排除许多错误的选项，在推动新材料、化学合成、生物医药等领域的科学进程方面起到关键作用。“例如，微软与合作伙伴一起，在AI的帮助下，从3200万种候选材料中发现了一种可以大幅减少电池中锂金属使用量的新材料，仅耗时80个小时。如果没有AI，同样的工作至少需要花费20多年的时间才能完成。”张祺说道。

生成式人工智能给 Web3.0带来三项挑战

本报讯 记者卢梦琪报道：4月28日，在2024中关村论坛年会——互联网3.0：未来互联网产业发展论坛上，清华大学特聘教授、清华大学人工智能研究院视觉智能研究中心主任邓志东表示，生成式人工智能给Web3.0产业发展带来三项挑战。

一是可能会产生幻觉、偏见与歧视。以大型语言模型作为基础的生成式人工智能，不同于Web3.0的去中心化，普遍存在着所谓的幻觉问题。

邓志东表示，生成式大语言模型及通用世界模型出现模拟偏差是不可避免的。大模型生成或模拟的多模态内容通常准确性不高，甚至出现事实性错误与偏见，训练数据可能存在种族与性别歧视等，这些或许会给Web3.0产业的安全性带来挑战。例如，在Sora生成的持续时间较长的视频中，偶尔会发生不连贯，或跳出一些不相干的物体，有时也存在空间左右不分的情况。

二是可能会带来知识产权、隐私与合规性问题。

邓志东表示，生成式人工智能的超大规模自监督预训练，主要依靠人类存储在互联网上的各种真实多模态数据，包括全世界各个语种的文本语料、程序代码、音频、图片和视频等，也涉及人类交互与反馈的价值观数据、下游的垂域微调数据等。

此外，大模型从网页抓取的数据可能涉及数据源的知识产权与隐私问题。由于生成的多模态内容是综合表达的结果，无法指明具体出处，因此也会存在合规性问题。

三是可能会形成新型网络诈骗与威胁。AIGC的文本到语音和图像生成功能，以及由Sora生成的视频，可能会传播虚假信息或社会谣言，生成式人工智能的滥用可能会带来有违法律、伦理、价值导向等有害信息。

邓志东表示，巨量的生成文本、图片与视频，可能会严重污染互联网空间，而利用AI复活获得的“永生”，也涉及伦理问题，因此也会带来新型网络诈骗、威胁与犯罪等各种挑战。

北京力争2030年形成脑机接口产业发展聚集区

本报讯 记者王伟报道：记者从4月26日举行的2024中关村论坛年会脑机接口创新发展与应用论坛上获悉，北京市将加快打造脑机接口产业创新高地，致力于在2030年形成脑机接口产业发展聚集区，培育全球影响力生态型企业和百家中小企业，实现典型应用场景和成熟产品落地。

据了解，脑机接口是集神经科学、电子技术、信息技术、先进制造、材料科学等多学科于一体的前沿技术，是新质生产力的典型代表，也是培育未来产业发展的重要方向。

在脑机接口创新发展与应用论坛上，北京市经信局正式发布了《加快北京市脑机接口产业发展行动方案(2024—2030)》(征求意见稿)(以下简称《行动方案》)。《行动方案》聚焦产业发展的关键环节和突出问题，提出一系列具有针对

性的政策措施，将通过构建产业生态体系、培育形成重点应用场景和推广应用示范、搭建共性支撑服务基础和完备标准测评体系“三大工程”，推动北京市脑机接口产业创新与突破。

《行动方案》明确了北京市发展脑机接口产业在2026年和2030年两个时间节点的发展目标。具体来说，北京市致力于到2026年实现核心技术突破、培育龙头企业、制定产业图谱和创新产品名录、推动医疗器械启动临床试验、推动典型应用场景、建成特色体验和展示中心、形成产教融合平台等发展目标；致力于到2030年实现技术体系完善自主、培育全球影响力生态型企业、培育百家中小企业、形成产业发展集聚区，吸引海外在京设立研发中心、打造产业发展集聚区、实现典型应用场景和成熟产品落地。

面向先进计算基础设施，摩尔线程“夸娥”智算集群以全功能GPU为底座，服务数字经济的各种应用场景；聚焦大模型底座，百川智能以其超大规模的参数量、卓越的语境理解与生成能力，以及对中文语言的把握，成为驱动众多应用场景创新与落地的强大引擎；依托全性能北斗芯，凯芯科技专注于平台研发，提供面向于智能驾驶、智能无人设备等领域的产品和服务，让万物感知更精准的时空……

本报讯 记者齐旭报道：在2024中关村论坛年会期间，中关村国际技术交易大会高精尖技术产品首发会人工智能与高端芯片专场活动在中关村软件园召开，一批北京人工智能与高端芯片产业前沿引领性项目首发亮相，为加速技术革新及行业规模应用按下快进键。