

驱动具身智能前进的根本力量不是资本叙事而是人工智能从“数字世界”走向“物理世界”



支点洞见

访谈嘉宾



王金桥

武汉人工智能研究院院长

编者按

8月19日上午,宇树科技正式登陆科创板,成为A股“人形机器人第一股”,备受资本市场瞩目。

8月22日至26日,第二届世界人形机器人运动会在北京举办。推动更多创新技术从比赛走向应用、从赛场走向生产生活,是此项赛事突出特点之一。

人形机器人作为智能机器人的最高形态,是具身智能的核心分支。目前,具身智能发展到了什么阶段?如何看待资本市场投融资火热与行业规模化应用滞后的“温差”?在国家级标准体系框架不断完善的情况下,各地如何依托自身产业基础实现差异化布局? 湖北日报《理论周刊》就此专访武汉人工智能研究院院长王金桥。

□ 谷穗

湖北地处长江中游,历史悠久,文脉绵远,是古人类重要起源地、楚文化发祥地、三国文化富集地,在中华文明的起源和发展中发挥了重要作用。从曾侯乙编钟的金石和鸣到楚绣的丝缕匠心,从屈原的家国情怀到辛亥首义的敢为人先,荆楚文化和合南北、融通东西,承载着中华民族的精神根脉,丰富了中华民族的文化基因。2024年11月,习近平总书记任湖北考察时强调,要“在加强文化资源保护和推动文化创新发展上担当使命”。明确审美导向、坚守价值内核,在传统与现代的碰撞中找到平衡点,是激活传统文化生命力的重要举措,也是培育青年文化自信、助力青年成长的必然要求。

年轻化表达不是“去传统化”,而是以更年轻的方式传扬“传统之魂”,这是荆楚文化年轻化的核心逻辑。年轻化表达的审美导向,本质是“古今共生”的辩证审美,要打破“传统=复古”“年轻=潮流”的二元对立,实现古典美学与当代审美的双向奔赴。数字虚拟空间已成为青年文化消费、社交互动的核心场域,也为荆楚文化年轻化表达提供了多元路径与无限可能。

依托数字技术赋能传播创新。鼓励运用3D建模、虚拟现实、增强现实等技术,开发更多类似数字编钟、虚拟楚绣的沉浸式体验产

一个产业从概念验证到规模化落地有其不可跳过的客观规律

记者:2026年上半年,全球人形机器人总出货量达到1.91万台,同比暴涨275%,其中,中国企业出货量占全球市场份额高达97.2%。对于国内具身智能行业而言,2026年是否是一个标志性的年份?如何评价目前国内具身智能发展的现状?

王金桥:我认为,2026年的标志性意义在于整个行业完成了一次深刻的认知切换,从能不能做出来转向了能不能用起来。

过去两三年,大家提到人形机器人的第一反应是会跳舞、会翻跟头,觉得是个噱头。但今年明显不一样,行业共识正在从讲故事时代转向交付时代。资本市场的关注也从谁的技术更炫转向了谁的产品能真正落地。工信部把2026年定义为应用元年和量产元年,具身智能被写入“十五五”未来产业规划,相关行业标准及配套的测评国标正式实施,这些政策信号叠加在一起,说明顶层设计层面已经完成了战略布局,产业发展的底层逻辑已经清晰了。

但标志性年份不等于成熟年份。一个产业从概念验证到规模化落地,有其不可跳过的客观规律。当前具身智能的产业成熟度,如果用人的成长阶段来打比方,我认为还处在“幼儿园大班”到“小学低年级”之间,能跑能跳了、能完成一些指令性任务了,但离独立生活、创造价值还有距离。

从技术层面看,当前具身智能面临的核心挑战是多层次的:底层是物理世界知识的匮乏,大模型分不清东南西北、时间长短,必须在空间中去具身体验,建立方位感和距离感;中层是数据获取的成本瓶颈,真机数据采集效率低、成本高,虽然无本体采集等新路徑正在探索,但距离大规模工业化还有差距;上层是泛化能力不足,换一个场景、换一个任务就得重新训练,真正意义上的举一反三暂时还未实现。

我认为,2026年是具身智能从“技术验证期”进入“产业爆发前夜”的拐点。当前最需要的是冷静下来把技术做扎实,把场景做清晰。产业的长远发展,靠的不是一时的热度,而是持续的技术积累和务实的场景落地。

任何硬科技产业从0到1阶段都会经历价值博弈期

记者:2026年上半年,国内具身智能赛道融资总额达935亿元,较去年同期提升5倍。有投资人直言:企业估值“水涨船高”,赛道已经产生泡沫了。泡沫是产业发展初期的必经阶段吗?如何将泡沫给行业带来的冲击降至最低,实现具身智能产业健康可持续发展?

王金桥:估值存在泡沫并非具身智能独有的问题,而是任何硬科技产业从0到1阶段都会经历的价值博弈。泡沫的本质,是技术价值被过早、过度地折现,当营销的速度和声量远超技术成熟度时,当投机预期取代真实用户需求成为定价基础,泡沫的产生几乎是必然的。

回顾科技史,这一规律一再被验证。2000年的互联网泡沫、移动互联网时代的“烧钱大战”、新能源汽车早期经历的“内卷式”竞争……每一轮泡沫退去之后,活下来的企业反而走得更稳、更远。具身智能当下的局面与此类似,不少企业通过概念包装推高估值,但真正跑通商业闭环的屈指可数。但泡沫不等于技术方向被证伪,恰恰相反,它标志着产业正进入一个去伪存真、价值回归的必经阶段。

我更倾向于将这一阶段表述为“价值博弈”,而非简单判定为泡沫,因为具身智能的底层技术逻辑是成立的。在2026世界人工智能大会暨人工智能全球治理高级别会议开幕式上,习近平主席强调:“人工智能是世界经济增长的新引擎和新旧动能转换的加速器,正从‘数字世界’走向‘物理世界’。”这是一个不可逆的技术趋势:AI不再只是在电脑里回答问题,而是走进工厂、仓库、医院去完成物理世界中的真实任务。这个方向的确定性不容置疑。具身智能之所以被赋予如此重要的产业地位,在于它连接了模型、芯片、数据、核心部件、制造能力和应用场景,具有极强的产业牵引作用。

比如,目前武汉人工智能研究院依托“紫东太初”多模态大模型,已在北京、武汉、佛山、青岛、长春、贵州等地建成机器人训练场,构建起“一脑多形、全域协同”的产业生态体系。每一个训练场背后都有真实的产业需求在驱动,有的是工业制造场景,有的是能源与特色产业深度融合,有的服务于智能制造转型升级。这些不是概念层面的演示,而是扎扎实实的产业落地。

因此,我认为,短期必然会有调整,部分企业会被淘汰,估值将回归理性,这是一个必要的“挤水分”过程。但长期来看,

具身智能完全具备支撑数十年健康发展的产业基础。中国拥有全球最大的制造业体量、最完整的产业链条、最丰富的应用场景资源,这些都是实打实的土壤。更重要的是,国家层面“人工智能+”行动的战略部署,以及“十五五”规划对智能经济的前瞻布局,都为产业提供了持续的制度保障。

泡沫破灭之后,存活下来的企业会更具竞争力。这个行业不会因为短期的资本波动而停滞,因为驱动它前进的根本力量不是资本叙事,而是人工智能从“数字世界”走向“物理世界”这一不可逆转的历史进程。

具身智能是大模型进入“物理世界”的“代理人”它把大模型的认知能力转化为实际生产力

记者:能不能结合具体案例展开讲讲?

王金桥:过去这些年,AI主要在数字世界里发挥作用,推荐算法、语音助手、内容生成,本质上都是在处理信息、生产内容。但人工智能真正的价值最大化,一定是从“数字世界”走向“物理世界”,去改造真实的产业,去参与真实的社会。这也是为什么具身智能被放到如此重要的战略位置,因为它是连接两个世界的桥梁。

比如,宇树科技和DeepSeek双方将在通用人工智能、高性能通用机器人、AI大模型三方面开展重点合作,就是一次“大脑”与“身体”的战略性对接。

宇树科技在硬件本体和运动控制上积累了多年经验,该公司的机器狗和人形机器人已经进入了数十个国家,在硬件工程化能力上有很强的实战经验。DeepSeek在大模型推理能力上做到了领先的水平,尤其在长上下文和复杂推理任务上表现突出。两者结合,正好优势互补。这个合作反映了一个更深层的产业趋势:人工智能正在从“单一能力竞争”走向“系统能力竞争”。以前做AI的公司拼算法、拼算力、拼数据就足够了,但今天你要让AI真正走进“物理世界”,光有算法远远不够。你得懂硬件工程、懂运动控制、懂传感器融合、懂场景适配。没有哪一家公司能独立完成所有事情,所以跨领域、跨技术的深度协同会成为常态。

从我们研发“紫东太初”大模型的经验来看,感受最深的一点是:大模型的能力不能只停留在“对话”和“生成”上,它必须能“动手”。但技术突破是一回事,真正让机器人在“物理世界”中稳定工作是另一回事。因此我们推出了“一脑多形”具身智能训练平台,同一个大脑,可以适配不同形态的机器人本体,目前已经支持超过50个品类的机器人,训练了超过300个具身原于技能。

具身智能在人工智能走向“物理世界”的进程中扮演的角色,我是这么理解的:

它是大模型的“物理接口”。大模型在云端的能力再强,如果它不能感知物理世界、不能与环境交互、不能执行具体任务,那它的价值就局限在屏幕里。具身智能无论是人形机器人、机械臂还是智能车,都是大模型进入“物理世界”的“代理人”。

它是真实数据的“采集器”。大模型要理解物理世界,靠互联网上的文本数据远远不够,它需要的是和物理世界交互的第一手数据,如拿起一个杯子的触感、避开一个障碍物的路径规划、识别一个物体的材质和重量。这些数据只能通过具身智能体在真实环境中采集,在真实场景中积累经验。

它是产业升级的“能力载体”。人工智能对实体经济的赋能,最终要落到具体的产品和服务上。具身智能就是这个“落地”的关键环节,它把大模型的认知能力转化为实际生产力。

当然,这条路还很长,但方向是对的,进步的速度也比很多人想象的要快。人工智能从“数字世界”走向“物理世界”,不是一句口号,而是整个人工智能产业正在经历的一场深刻变革。具身智能就是这场变革的排头兵。

能干活和能规模化地干活中间隔着一条不小的鸿沟

记者:目前,大部分的人形机器人卖给了大学、研究机构和个人开发者,真正跑通商业闭环的企业仍是少数。具身智能距离真正的规模化商业落地还有多远?当前要突破的最关键瓶颈是什么?

王金桥:我们要区分两个概念:能干活和能规模化地干活,中间隔着一条不小的鸿沟。目前大部分机器人销往高校、研究机构和开发者社区,这本身说明产业仍处于“技术验证→场景适配”的早期阶段,而非“产品交付→价值创造”的成熟阶段。真正跑通商业闭环的企业屈指可数,这背后反映的是技术成熟度与市场需求之间的结构性错配。

验中培育高尚的道德情操、树立正确的价值观念。

深入挖掘荆楚文化的当代价值。荆楚文化中,屈原“路漫漫其修远兮,吾将上下而求索”的执着追求,体现着家国情怀与责任担当;楚人的“敢为人先”,彰显着创新进取的时代精神;非遗传承人的匠心坚守,传递着精益求精的价值追求;辛亥首义的革命精神,蕴含着救亡图存的爱国情怀;茶圣陆羽、医圣李时珍的探索实践,彰显着求真务实的治学态度。这些精神内核,既是荆楚文化的灵魂所在,也是审美年轻化表达的核心导向。这种“以美育人、以文化人”的浸润式引导,远胜于生硬的说教,能让历史人文的印痕真正融入青年血脉。文化育人不是“药到病除”的即时疗法,而是“润物无声”的长期浸润。正如屈原的求索精神,其力量不在于给出了答案,而在于教会了青年如何面对没有答案的困境。

将价值引领融入传播全过程。通过楚剧年轻化改编、荆楚故事情景演绎、先进典型宣

距离规模化商业落地还有多远?我的判断是:局部场景在未来2至3年内有望实现商业化突破,但全行业真正意义上的规模化落地,还需要5到10年的持续攻关。这不是一个靠资本加速就能缩短的时间周期,因为它涉及的不只是资金投入,而是一系列硬核技术难题的系统性突破。

当前最关键的瓶颈,我认为可以概括为“三堵墙”:

数据墙。数字AI的成功,很大程度上得益于互联网积累了海量的文本、图像数据。但具身智能需要的是真实物理世界中的交互数据,机器人抓取一个物体的力反馈数据、行走在不同地面的平衡数据、操作柔性物体的形变数据,这些数据极度稀缺,获取成本极高。数据从能用用到够用,从够用到好用,还有很长的路要走。

表征墙。当前的机器人训练存在严重的碎片化问题,不同品牌机器人的传感器配置、数据格式、控制接口互不兼容,A厂数据训练不了B厂机器人,轮式经验无法迁移至双足。跨任务、跨场景、跨本体的统一物理表征尚未形成。这导致我们无法像训练大语言模型那样,用海量异构数据训练出一个通用的机器人基础模型。武汉人工智能研究院正在推动具身智能“一脑多形”技术路线和标准体系上的双线布局,例如在团体标准层面牵头编制全国首个《具身智能机器人训练场建设要求》,在国家标准层面牵头起草我国首个面向具身智能训练基地的数据采集与模型训练规范,试图从技术架构和标准规范两个维度打破数据孤岛与品牌壁垒。但标准的落地、推广和互认,还需要全行业共同努力。

闭环墙。真实世界的试错代价极其昂贵。在数字世界里,AI可以模拟亿万次对局来提升棋力,但在物理世界里,机器人摔一次就可能损坏硬件,操作失误可能造成生产事故。这使得物理世界的反馈循环缓慢而昂贵,严重制约了模型能力的快速迭代。要突破这道墙,需要高保真仿真环境的持续进化、高效的数据飞轮机制,以及从仿真到真实的零样本迁移能力。

除了这三堵技术层面的墙,还有一个不容忽视的现实瓶颈:产业生态尚未形成闭环。当前很多企业卖机器人的模式是卖硬件、送软件,但客户买到手之后,缺乏配套的技能训练服务、场景适配方案和持续迭代支持。一个训练场如果仅靠一次性出售裸数据收回投资,财务模型很难成立。具身智能的商业化,不能只靠卖本体赚钱,而必须构建“数据采集—模型训练—技能交付—持续迭代”的完整服务链条。

既需要战略上的长期主义也需要执行上的系统性工程能力

记者:具身智能是一整套涉及硬件和软件的系统产业链。对各地来说,具备什么样的优势可能更适合发展具身智能?

王金桥:跟风只会制造短期泡沫,深耕才能沉淀长期价值。每个地方的产业基础、科教资源、应用场景等各不相同。千城一面,不是发展,是内卷。

以湖北为例,光电子信息、汽车、现代化工、大健康等都是万亿级产业,是人工智能可以直接赋能、深度改造的真实场景。如何把人工智能与湖北优势产业结合起来,在智能制造、智慧医疗、智能驾驶等领域推进深度落地,这不是凭主观判断,而是基于产业需求做出的选择。只有回到产业本身的逻辑,才能找到AI落地最有效的路径。

错位发展。各地要根据自身产业禀赋选择主攻方向,形成差异化竞争格局,而非同质化竞争。比如,湖北聚焦智能制造和汽车领域的具身智能应用,长三角聚焦消费级和服务型机器人,珠三角聚焦3C电子与物流场景。各展所长,各取所需,而不是在单一赛道上过度拥挤。

标准先行。今年2月发布的《人形机器人与具身智能标准体系(2026版)》是我国首个覆盖全产业链、全生命周期的标准顶层设计,涵盖基础共性、脑智与智算、整机与系统、安全伦理等六大板块。标准的意义不仅是画红线,更是搭底座,只有通过标准统一接口和协议,海量真实数据才能有效汇聚,反哺模型训练,形成“数据—训练—迭代”的良性循环。

开放协同。具身智能的产业链条极长,没有一家企业、一个地区能独立完成所有事情。构建开放协同的生态,是降低重复建设、提升资源配置效率最有效的组织方式。当资源在更大范围内实现流动与共享时,低水平内耗自然失去生存空间。

发展具身智能,既需要战略上的长期主义,也需要执行上的系统性工程能力。产业方向已经明确,真正的考验在于能否在技术迭代中持续积累认知壁垒、在场景验证中形成可复用的方法论、在生态构建中保持战略耐心。

(湖北日报全媒记者 常少华)

传等形式,让青年在审美愉悦中感受文化魅力,在潜移默化中接受价值滋养。要坚决抵制为博取流量而过度娱乐化、消解文化内核的错误倾向,引导青年明辨是非、坚守正道,自觉抵制低俗化、碎片化审美,形成正确的审美认知与价值判断,让荆楚文化成为培育青年高尚情操、坚定理想信念的重要载体。

发挥文化的育人功能。当下,青年群体面临着学业、就业等多重压力,荆楚文化所蕴含的审美力量与人文底蕴,正是抚慰心灵、滋养精神的重要养分。当青年在挫折中迷失方向时,屈原的家国情怀有助于引导其跳出个人困境,树立远大理想;当面对社会不良风气时,荆楚文化中“崇俭尚廉”“诚实守信”的传统美德,能帮助其坚守道德底线,涵养浩然正气;当在纷繁复杂的矛盾中陷入困惑时,荆楚文化所蕴含的道家辩证思维与人文智慧,能引导其理性看待问题,培育端正的世界观、人生观、价值观。

荆楚文化作为湖北最深厚的文化软实力,是推动高质量发展的重要精神支撑。坚持“审美创新不丢魂、价值坚守不僵化”的辩证思维,让荆楚文化的美学魅力适配当代青年的审美需求,让荆楚文化的精神内核滋养当代青年的心灵世界,为湖北加快建设文化强省、铸就社会主义文化新辉煌注入源源不断的力量。

(作者单位:湖北民族大学音乐舞蹈学院)

传统文化的年轻化表达并非脱离本质的“标新立异”

品,降低传统文化的体验门槛,让青年在互动中感受楚文化的魅力。《楚国八百年》纪录片及衍生短视频产品,通过数字化还原楚国历史场景,在B站等青年平台获得过亿播放量,其核心经验就是将专业的历史叙事转化为青年喜闻乐见的视听语言,让文物从“隔着玻璃看”变为“走进故事里”。

突出青年的传播主体地位。鼓励青年群体以短视频、微短剧、文创设计等形式自主解读荆楚文化,支持青年创作团队开发以荆楚文化为背景的动漫、游戏、数字藏品等产品,让青年从文化的接收者转变为传播者、创作者。宜昌“屈原故里端午文化节”近年推出的“端午文创设计大赛”,吸引了全国近千名青年设计师参与,产出了一批兼具文化内涵与时代气息的文创产品,实现了文化传播与青年参与的双赢。更深层的意义在于,青年从“被投喂”的受众转变为“在场”的共创者,这一身份转换本身就是文化自信最生动的注脚。年轻人主动穿上楚风汉服、在短视频里讲述楚辞故事,文化

的传承就不再是单向的“灌输”,而是自发的“流淌”。

推动文化业态深度融合。搭建多元传播平台,推动荆楚文化与数字动漫、沉浸式展演、文旅消费等新业态深度融合,让审美表达既有传统底蕴,又有时代气息。武汉“昙华林”历史文化街区引入楚绣、汉剧等非遗体验项目,结合青年喜爱的咖啡馆、文创市集等业态,成为年轻人感受荆楚文化的网红打卡地,实现了“传统基因+现代表达”的有机统一。事实证明,文化业态融合越深,传统与现代的“化学反应”越强烈,对青年的吸引力就越持久。

形式创新是路径,价值坚守是灵魂。坚守价值,不是墨守成规、故步自封,而是要在辩证统一中把握“变”与“不变”的中国智慧。“变”的是传播方式与表达形式,“不变”的是荆楚文化的精神内核。审美表达的年轻化,绝非脱离文化本质的“标新立异”,而是要在创新形式的同时,深挖精神内涵,引导青年在审美体