

积极从经典理论汲取人工智能治理智慧

2026年恰逢两个科技里程碑的七十周年。国际上,1956年达特茅斯会议提出“人工智能”学术概念;在中国,我国制定《1956—1967年科学技术发展远景规划纲要》,是新中国第一个科技规划,将无线电电子学、自动化、电子计算机、半导体列为优先发展的四项紧急措施,推动自动控制理论纳入国家布局。钱学森深度参与了这一规划,其以控制论为代表的理论探索正是这一布局的重要思想基础。70多年来,全球人工智能加速发展,人工智能深刻重塑科技范式、产业发展和经济社会,控制论等经典科学理论正为人工智能研究与治理提供新的思想资源。控制论以信息为基础、以反馈为机制、以系统为对象,与卡尔达舍夫文明理论所代表的能量利用维度、萨根信息层级所代表的信息利用维度内在呼应,共同构成审视人工智能发展与人类文明进阶的经典理论视角,积极探索其时代启示,意义重大、正当其时。

中国电子信息产业发展研究院党委副书记 胡国栋
人工智能产业工委秘书长 高宏玲
赛迪智库新型工业化研究所 王高翔

控制论等经典科学理论

提供人工智能研究新视角

当前,大模型技术加速突破,智能体应用持续拓展,机器人产业蓬勃发展,人工智能重塑科技范式、产业发展、经济社会。同时,国际学术界出现运用经典控制论解析人工智能的新研究视角,英国剑桥大学等机构学者探索运用反馈闭环、系统稳态等经典控制论思想,为大模型、智能体搭建底层理论框架。

早在20世纪50年代,钱学森出版《工程控制论》,搭建起面向工程实践的控制理论体系。这些以信息为基础、以反馈为机制、以系统为对象的科学思想,正与今天人工智能发展形成跨越时代呼应。《工程控制论》奠定了工程层面反馈与自动控制的理论基础,其内在预设是:所有受控系统的目标、边界、安全约束均由人类设定,自动化只是实现人类目标的工具手段。在此基础上,钱学森进一步发展提出“人机结合、以人为主”的核心原则。该思想主张,面对日益复杂的人机协作体系,最终决策权必须保留于人类主体,不能完全交由自动化系统。这为当代应对人工智能自主性风险、抵制AI神化叙事、守住人类主体性,提供了源自工程控制论的理念支撑。受此启发,基于“CTRL”的语义框架(“控制”英文缩写),“C5(Cake)、T5(Trend)、R5(Recursion)、L5(Level)”恰好形象概括了科技产业界关于人工智能的讨论。

一是做大“C5”代表的人工智能产业蛋糕。2026年1月,英伟达创始人黄仁勋在达沃斯世界经济论坛上把AI基础设施描述为“五层蛋糕”:自下而上分为能源、芯片、基础设施、模型、应用五个层级。

二是推动“T5”代表的人工智能技术迭代,即从弱人工智能到通用人工智能(AGI),从语言模型到世界模型,从信息智能到物理智能,从机器自主到人机共生,从单点智能到体系智能。

三是把握“R5”代表的人工智能进化趋势。中国科技学者2026年发表《人类制造的最后一个AI:走向真正的递归自我改进》(《The Last AI Built by Humans: Toward Genuine Recursive Self-Improvement》,提出人工智能将从改进执行自主(improvement-execution autonomy)、改进策略自主(improvement-strategy autonomy)、经验获取自主(experience-acquisition autonomy)、环境适应自主(environment-adaptation autonomy)走向递归元改进(recursive meta-improvement),日益实现执行自主、策略自主、经验自主、适应自主、机制自主的演化。

四是把握“L5”代表的人工智能发展阶段,即前沿大模型企业提出的对话者(Conversationalist)、推理者



(Reasoner)、代理者(Agent)、创新者(Innovator)、组织者(Organizer)等五个阶段。

经典控制论思想

对人工智能治理的重要启示

第一,坚持一脉相承,从控制论根基把握人工智能演进逻辑。人工智能并非凭空而来的科学范式,而是经典控制论思想的数智化升级。大模型的闭环推理、智能体的自主迭代,机器人的环境交互,底层运行逻辑并未脱离控制论科学框架。中国香港理工大学等学者2026年发表《基础智能体缺失的科学:智能体控制论》(Agent Cybernetics Is the Missing Science of Foundation Agents),指出当前基础智能体普遍存在底层理论缺失、机理阐释不足的问题,而经典控制论的闭环调控、系统稳态原理,能有效补齐新一代人工智能的理论短板,为智能体发展提供底层科学支撑。

第二,坚守安全底线,以鲁棒约束筑牢发展前提。人工智能发展的首要前提是系统稳定性与鲁棒约束。经典控制论强调抑制扰动、维持系统稳态,这是当前国际人工智能治理研究的核心焦点。2025年,图灵奖得主、深度学习三巨头之一约书亚·本吉奥在论文《超级智能体构成灾难性风险:科学AI能否提供更安全路径?》中提出,通用自主智能体若无底层安全约束,极易出现欺骗、目标偏移、行为不可预测等失控风险。前沿学者普遍认为人工

智能自进化必须建立约束边界,无约束的自主迭代会引发系统偏离,稳定性是智能持续进化的底层前提。

第三,强化系统观念,以多元协同引领人工智能发展。从单点智能走向体系智能,本质是控制论多子系统耦合思想的体现。世界经济论坛2025年白皮书《行动中的AI智能体:评估与治理基础》提出,多智能体生态加速形成,多个LLM智能体通过分工协商、分布式协同处理复杂跨领域任务,治理框架必须从单模型评估转向多智能体系统层面的风险管控;现代经济学、复杂系统科学前沿也证实,高级智能、高级经济系统核心特征是协同共生而非单点最优,这与经典控制论学派长期坚持的系统协同思维高度一致。

第四,汲取历史智慧,以自组织治理思想指引智能未来。智能体自组织演进,呼应控制论学者格卢什科夫基于控制论提出的国民经济自动化治理的超前构想。格卢什科夫1962年提出OGAS全国自动化经济管理系统构想,并在1975年专著《宏观经济模型与OGAS建设原理》中,运用控制论研究大规模社会经济复杂系统的信息采集、分层调控与自组织。他设想依托全域信息闭环,实现国民经济动态协同治理。如今智能体逐步从工具成长为组织者、创新者、协同者,正是当年这套复杂社会经济系统控制思想在人工智能时代的回响,也成为当下复杂系统与智能治理领域学者重新挖掘的经典思想源头。

总之,经典控制论思想启示我们,人工智能治理既要立足技术演进规律、把握发展逻辑,也要坚守安全底线、强化系统观念、汲取历史智慧。唯有以科学理论筑牢治

灵巧手量产按下加速键

(上接第1版)

“从行业层面来看,万台级规模化出货,印证了国产灵巧手的工程化落地能力和供应链体系成熟度,为行业后续降本增效、大规模场景应用奠定了核心的规模基础。”因时机器人CMO房海南告诉《中国电子报》记者。从核心零部件采购、产品装配到质量检测,灵巧手正逐步告别以手工调试为主的小批量试制,向标准化、规模化制造迈进。

随着量产能力的逐步成熟,市场检验的重点开始发生变化。灵巧智能创始人兼董事长许春山博士在接受《中国电子报》记者采访时表示:“当前行业存在‘重参数、轻能力’的倾向,许多产品评价仍集中在自由度、重量、负载等指标上。但随着产品逐步进入场景验证阶段,客户的选择标准正在发生深刻变化。真实工况下的稳定性、场景适配度和全栈交付能力已成为下游客户的核心关切。”

从“卷参数”到“拼交付”,灵巧手市场正跨越从实验室验证到规模化应用的关键节点。不过,交付规模的扩大只回答了灵巧手能否批量“造出来”,尚未回答造出来后能不能真正“用起来”。

万只灵巧手究竟流向哪些客户,其中又有多少进入真实作业场景并稳定创造生产价值,成为需要追问的下一个关键问题。

工业场景导入仍处早期

客户首先要算经济账

灵巧手的高速增长由多元场景共同支撑。“当前推动销量的三大需求主要是,其一人形机器人本体配套需求;其二科研实验与数据采集需求;其三工业场景的初步导入,如汽车制造、3C电子、精密装配、电力巡检等领域开始小批量部署。”许春山告诉《中国电子报》记者。

据MIR睿工业今年9月发布的《2026年中国灵巧手行业发展白皮书》,2025年灵巧手应用场景中,科研教育和数据采集合计占比为74.4%,娱乐表演占8.3%,工业场景占比最低,仅为3.5%。

近3/4的科研教育和数据采集应用与不到4%的工厂应用,暴露出灵巧手供给结构性不足的问题。许春山指出:“通用型、低成本的灵巧手供给相对充足,但真正能在工业场景中长期稳定运行的高可靠性产品仍然稀缺。行业不缺‘能做演示的手’,缺的是‘能稳定干活的手’。”

产品仍处于早期产业化阶段,工业场景中的刚性需求尚未充分释放。因时机器人2025年灵巧手交付量超1万台,据GGII数据,公司当年的灵巧手出货量位居中国灵巧手企业首位。但具体到场景细分,工业端的应用进度仍相对缓慢。房海南表示:“现阶段因时机器人客户以机器人本体厂商、科研教育及算法数据采集客户为主,工业场景客户处于POC验证渗透阶段。”

与此同时,受场景复杂度、算法适配、产线改造等多重因素影响,房海南指出:“灵巧手从送样、POC验证到正式产线落地的周期差异非常大,现在行业尚未形成标准化、固定化的落地周期。”

自由度、触觉点等技术参数,最终都要落实到具体场景中的工作表现。帕西尼灵巧手技术产品负责人向《中国电子报》记者介绍:“真正进入实际应用后,客户更关注的是产品在真实复杂场景中的综合表现。比如面对不同材质、形状和重量的物体,能否实现稳定抓取和精准力控;在高频重复作业中,性能是否能够保持一致;同时,产品是否易于集成、维护和二次开发。这些问题实际上都指向同一个核心:灵巧手不仅要‘能完成一次任务’,更能够在真实环境中持续、稳定地完成任务。”

跨越运行稳定性的门槛后,客户还要考量一个核心指标——投资回报率(ROI)。房海南坦言:“在因时对接的工业验证客户中,完成POC验证后实现复购及小批量采购的客户占比有限。核心也是要算ROI,所以尚未形成规模化采购意愿。”

对生产端客户而言,灵巧手作为一种新型“生产资料”,“划不划算”将直接影响甚至决定其采购决策。企业需要从长短期效益等多个层面出发,对比灵巧手和熟练工人之间的综合成本和实际效益。房海南认为,对于生产场景来说,灵巧手目前在产品层面还没有完全达到成熟状态。

迈向规模应用

工程化闭环能力仍是关键

“制约灵巧手大规模交付的本质在于,灵巧手是目前精密机械电控与感知领域‘体积受限最严重、集成度最高、动态环境交互最频繁’的极端系统。”临界点CTO熊坤告诉《中国电子报》记者。

要在“巴掌大”的空间内放置几十个微型关节和驱动、传动、传感及控制单元,还要兼顾散热、布线与结构强度,这对企业的系统集成与工程化能力提出了极高的要求。与此同时,自由度和触觉等性能的提升也会增加工艺复杂度,进而对长期运行稳定性和成本控制提出更高要求。

“高自由度触觉灵巧手的市场价格为几万元到十几万元,但在工业连续作业下实际使用寿命有限,客户面临‘买不起、也不耐用’的两难困境。”许春山告诉记者。性能、可靠性与成本,构成了灵巧手绕不开的“不可能三角”。

除了硬件工程化的难题之外,软硬件之间的磨合同样面临挑战。一方面,“有手无脑”的瓶颈仍然突出。许春山告诉记者:“当前具身智能大模型在驱动高自由度灵巧手完成复杂精细作业方面仍面临较大挑战,开源的灵巧操作真实数据也相对有限。”

另一方面,能够支撑模型研发、被研究者广泛采用的成熟硬件产品仍然缺口。Alphaist Partners联合创始人陈哲在此前的一场对话中曾表示,灵巧手市场也存在着一个对标“宇树G1”的生态位,即应该有一款产品能够成为中美所有灵巧手研究者的首选方案,进而形成事实上的行业基准。

但目前这个“宝座”仍虚位以待。为此,部分灵巧手企业将从硬件到模型的全栈自研视为突破路径。Sharpa在自研灵巧手的基础上,开发了融合视觉、触觉和语言信息的CraftNet模型。今年9月28日,公司在美国匹兹堡举行的智能机器人与系统国际会议(IROS 2026)上发布全触觉轻量化灵巧手W02,并同步推出外骨骼触感数据手套

理根基,方能让人工智能在服务人类发展中行稳致远、向善普惠。

进一步应用经典理论

探索人工智能的未来

如果说控制论为人工智能研究与治理提供了底层理论框架,那么以卡尔达舍夫理论为代表的经典文明演进理论,则从人类文明进阶的高度,拓展了人工智能治理与发展的宏观视野。

人类文明进阶评价体系有着深厚的经典天体物理理论渊源。1964年,天文学家卡尔达舍夫在布拉卡恩会议上提出经典文明等级理论,以宇宙能量利用能力为唯一标尺,构建了行星级、恒星级、星系级的三阶文明演化框架。该理论长期引领人类对文明发展上限的认知,近年来再度成为全球前沿科技界重点研究范式。以马斯克为代表的硅谷前沿科研与产业主体明确提出,人类迈向卡尔达舍夫Ⅱ级恒星级文明,必须突破地球能源边界、构建月球生产体系与太空太阳能利用体系。总体来看,传统卡尔达舍夫理论建立了能量维度的文明评价标准,但受时代局限,尚未纳入信息智能维度,难以适配人工智能、大模型算力革命驱动的当代文明演进特征。

科技学者认为,高阶文明进阶必须实现物质能量开发与数智信息开发双轮驱动。随着算力网络、大模型、智能体技术快速迭代,信息生成、处理、推演、决策能力成为衡量文明高度的新维度。美国学者萨根在原卡尔达舍夫能量分级之外,增设A至Z的信息层级,当前人类综合能量与信息发展水平仅处于0.7H初级文明阶段。学界认为,未来文明迭代不再单纯依赖能源规模扩张,而更多依赖数智技术对宇宙信息的解析、建模、重构与预判,通过能量维度与信息维度协同突破,构建更加完整、科学的人类文明进阶新体系。

从控制论的视角看,卡尔达舍夫理论与萨根理论并非彼此割裂,而是内在呼应、相互补充。控制论以信息为基础、以反馈为机制、以系统为对象,恰为文明评价从单一能量标尺走向能量与信息双维标尺提供了理论依据:文明越是高阶,对信息采集、传递、处理与反馈调控的能力要求越高,文明进阶本质上是复杂大系统调控能力的系统性跃升;能量利用与信息利用作为文明演进的双重动力,唯有在系统层面协同调控、动态平衡,方能支撑文明持续向更高阶段跃升。

回望经典控制论的思想脉络,人类文明进阶的每一步,都伴随着信息获取、系统调控与反馈优化能力的跃升。展望未来,我们更应以系统观念把握人工智能发展全局,以鲁棒约束筑牢安全底线,以多元协同凝聚治理合力,让人工智能这一控制论思想的数智化结晶,在服务人类共同福祉中持续释放正能量,推动人类文明由初级层级持续向高阶文明迭代跃升。

AE01,进一步完善执行硬件与操作数据采集的配套布局。不过,从行业协同层面来看,企业内部的纵向整合仍难以消除不同手型、接口和数据体系之间的差异,模型与数据的跨平台复用仍有赖于行业标准与开放生态建设。

此外,随着灵巧手进入真实工作环境,场景适配问题也摆在了企业面前。

熊坤告诉《中国电子报》记者,从临界点目前对接的客户来看,不同领域客户的需求侧重不同:“偏学术科研的客户可能更加注重灵巧手自由度、感知能力等性能参数指标以及二次开发接口的友好程度;工业和商业客户则更加看重可靠性、使用寿命、性价比以及MTBF(平均无故障时间)与可维护性。”

许春山指出:“灵巧手不是标准标配产品,更多的时候需要针对场景做深度适配。”一只将性能推向极致的灵巧手不等于能够在实际工作中把效率拉到极致,在选购时,客户往往需要在负载、精度、速度、可靠性和成本等指标之间作出相应的取舍,企业也需要具备成熟的方案定制和全栈交付能力。

除此之外,真正对接需求时,企业还面临与客户缺少“共同语言”的困境。“关于规模化复制的条件,当前最关键的制约因素在于,行业尚未形成统一的灵巧操作能力标准和评价体系,不同场景对‘能干多少活’缺乏可量化的衡量尺度。”这种情况不仅增加了客户横向比较、筛选产品和现场验证的成本,也使供应商难以清晰证明产品在具体工况中的真实能力,无形中增加了供需双方的沟通成本。

人形机器人真正“上手”干活,仍需要跨越这最近却也最远的“最后一厘米”。从参数竞赛到能力比拼、从产品交付到全栈服务、从通用方案到场景定义,将是行业在未来竞争的三个核心转向;而相关评价标准的补齐,将会在规范市场竞争、促进供需对接方面为行业发展提供重要支撑。

关于行业未来的发展方向,帕西尼灵巧手技术产品负责人判断:“未来2~3年,行业竞争必将朝‘算法、成本、性能’三者的动态平衡方向演进。真正能建立壁垒的产品,不一定是参数最激进的,但一定是最能实现规模化落地的。”