

EN 展望 2026

# 具身智能芯片：多元与趋同



本报记者 王信豪

经过数年的技术积累，具身智能正从过去的模糊概念向明确具体的落地应用迈进，渗透至人形机器人、新能源汽车、工业控制、物流运输等多元场景，成为AI与实体经济融合的重要载体。

作为具身智能设备的算力中枢，芯片的重要价值越发凸显。当下，这一赛道呈现出“百花齐放”的分化态势：场景需求多元、技术路径各异、参与玩家林立，尚未形成一套为具身智能打造的专用芯片。而在“分化”的背后，底层技术逻辑、产业发展规律又逐渐凝聚起趋同的共识，一场关于算力、架构与生态的博弈与协同正在悄然上演。

## 多元：场景需求催生的差异化探索

路径分化的根源，在于具身智能扎根的物理世界没有统一场景模板。工业、家庭、医疗、物流等场景特性各异，对芯片的功耗、算力、实时性、安全性需求差异显著，直接催生了芯片领域的差异化探索。

目前，业内常用“大小脑”表示具身智能在不同控制层次的处理方式，“小脑”负责低延迟、高确定性的实时控制，“大脑”侧重高层次感知、规划与学习，进而形成“大小脑融合”与“大小脑分离”的路径分歧。辉羲智能创始人兼CEO徐宁仪向《中国电子报》记者解释：“‘大小脑’架构的选择本质上取决于场景的本质需求。在工业场景中，分离设计可能更利于实时保障；而在家庭服务等动态开放环境中，融合

架构能提升多模态信息的协同效率。当然，我认为融合的智能上限更高，因为人类抽象化的思考行动方式更倾向于融合。”

场景的多元与技术路径的分歧，直接导致行业至今没有能“通吃所有需求”的统一解决方案。对于大多数具身智能终端企业而言，“跨厂商芯片拼搭”成为普遍的算力构建方式：算力密集的感知与规划环节，可能采用智能驾驶芯片（如英伟达Orin）；核心控制与实时运算部分，倾向于选择消费级CPU（如英特尔酷睿系列）；低功耗的边缘处理方面，ASIC的能效比更具优势。这种“按需组合”的模式，正是当前行业缺乏统一标准与万能方案的直接体现。

“当前并非推出机器人专用芯片的最佳时机。”英特尔中国研究院院长宋继强直言，“因为机器人产业规模尚小，且VLA、世界模型等技术路径未统一，机器人本体也没有公认的最优标准，因此通用芯片的适配性更能应对算法迭代的不确定性。”

## 趋同：差异背后的共性逻辑

看似多元的赛道背后，行业已在核心要素与技术架构等方面凝聚起共识。

一方面，数据作为核心生产要素的重要性日益凸显。所有技术探索的起点都源于场景需求，场景定义了数据采集的范围、标准与维度——无论是VLA训练所需的视觉语言数据，还是精密控制所需的动作反馈数据，都必须与具体场景强绑定，进而形成“场景-数据-算法-芯片”的正

向循环。

“多模态数据处理能力的升级是具身智能最为核心的部分。真正的智能必须理解并整合视觉、语音、触觉等不同模态的信息，才能做出接近人类的决策。”徐宁仪向记者表示，“在具身智能从仿真迈向落地的进程中，首要任务就是确立数据闭环，通过实时收集端侧数据并优化模型，为应用场景寻找最为合适的算法，并定义后续的算力需求。”

另一方面，数据与算力的多元需求推动异构计算成为未来重要的架构方向。目前，单一计算单元无法满足“感知-决策-执行”的全链路需求；高层级感知规划需要GPU的并行计算能力，实时动作控制依赖NPU的低功耗与快速响应，精密执行则离不开CPU的高精度浮点运算。宋继强表示，“很难用同一种硬件解决所有问题，底层必然需要异构计算”，这种“多单元各司其职、协同发力”的架构逻辑，已成为行业默认的技术底座。英特尔发布的第三代酷睿Ultra for Edge平台基于异构计算的核心思路，可提供180 TOPS的算力，依托18A制程也能保证较高的能效比。

在CES 2026上，高通面向人形机器人、自主移动机器人和工业机器人推出了高通骁龙IQ10处理器，通过异构计算架构，将集成定制的CPU、GPU、NPU与低功耗传感器中枢深度融合，为机器人等智能终端提供低功耗的AI性能。

## 和而不同：锚定商用落地向前迈进

多元与趋同并非对立，而是

具身智能芯片产业发展的一体两面：多元分化是场景多样化的客观结果，能为产业保留创新弹性；趋同统一是规模化落地的必要前提，为行业划定发展底座。二者在动态平衡中形成的协同生态与规范标准，将成为2026年产业商业化落地的主线。

徐宁仪表示，具身智能的落地从来不是单一硬件的比拼，而是软硬件深度协同的系统工程。芯片性能的发挥离不开传感器的精准采集、算法的高效适配，更需要配套的软件工具、算法库与开发套件降低落地门槛。这要求芯片企业成为“场景的深度解读者”，据了解，辉羲智能通过打造异构可扩展的计算架构，凭借较高的易用性和拓展性，可实现对多种应用场景下不同算力与接口的兼容。

针对机器人、高通不仅提供了包含软件、机器学习运维、AI数据飞轮等多种工具套件，还与VinMotion、加速进化Booster、阿加星APLUX和Advantech等商业伙伴广泛展开合作，共同推动机器人创新解决方案的规模化部署。

2025年12月，工业和信息化部人形机器人与具身智能标准化技术委员会成立，将聚焦标准滞后、基础接口不统一、应用缺乏规范等问题开展工作，加快推动人形机器人与具身智能技术熟化和应用落地。

随着开放生态持续完善、行业标准逐步落地，具身智能芯片产业将摆脱碎片化的过渡状态，在多元创新与协同规范的平衡中实现技术价值与产业价值的双向奔赴，为具身智能从概念走向现实、从试点走向普及筑牢算力根基。

## 灵达发布服务器存储控制与高速网络连接系列I/O部件

**本报讯** 记者许子皓报道：近日，光合组织成员单位灵达发布服务器存储控制与高速网络连接系列核心I/O部件。灵达发布的灵可达Linkdata产品采用自研架构，在性能与能效方面实现同步提升。目前，灵可达Linkdata产品已在通用服务器、AI服务器及重点行业应用场景中实现规模化交付。

在服务器产业链中，RAID、HBA及高速互联等核心I/O部件虽然在整机成本中占比有限，但技术壁垒高、生态依赖强，长期由少数国际厂商主导。以全球RAID控制卡市场为例，博通、微芯等厂商占据超过一半的市场份额，并长期作为主流服务器厂商的基础配件来源。在国内市场，本土厂商在上述核心部件领域的市场占有率长期处于较低水平。尤其是在RAID和HBA控制卡方向，过去几年国产产品市场份额仍低于国际品牌。

在金融、电力等对稳定性和可靠性要求极高的行业中，这类核心I/O部件直接关系到系统长期运行能力，因此，补齐这一环节，被业内认为是国内构建完整算力体系的重要前提。

灵达本次发布的灵可达Linkdata产品覆盖HBA卡、RAID卡及10G、25G网卡等核心I/O部件，均基于自研架构设计，实现了100%可控。其中，存储产品全面支

持Tri-Mode规格，可同时兼容SATA、SAS与NVMe三种协议形态，为以NVMe为主的存储架构预留了升级空间。

在典型多盘位服务器及高并发场景下，灵可达Linkdata自研HBA在读取性能优异，自研RAID卡通过硬件与算法优化，在SSD场景中表现出更优的RAID写入性能。多通道下联设计与板载M2热插拔能力，使单卡扩展密度和运维效率同步提升，更贴近金融、电力等关键行业对稳定运行和快速维护的实际需求。

在网络侧，灵可达Linkdata 10G、25G网卡在实现线速吞吐的同时，兼顾延时延与能效表现。相关产品均已完成与多家主流云平台、整机厂及操作系统的兼容适配，不仅支持国内CPU与操作系统环境，也可接入Intel、AMD等主流体系，为不同算力平台提供可复用的I/O方案。

灵达表示，公司将依托光合组织平台，与上下游伙伴在接口标准、开放架构及联合定制等方面进一步深化协作，推动本土硬件向“主动参与规则构建”转变。

据了解，灵可达Linkdata产品已完成与多家主流整机厂的适配部署，覆盖通用服务器、AI服务器及液冷服务器等多种形态，累计出货量超过30万片，并在金融、运营商、互联网及能源等行业实现稳定运行。

（上接第1版）以近期热门的AI Agent Clawdbot为例，传统的AI Agent多依赖云端大模型，而“无涯问知AIPC”通过搭载本地30B大模型并结合Clawdbot，可通过即时通信软件，以自然语言交互的方式指挥笔记本电脑执行联网搜索及本地问答。

## 智能终端

### 如何成就AI Agent?

“更懂用户”正不约而同地成为几乎所有智能终端的设计目标。

面对“千人千面”的用户群体，在处理复杂任务时，终端设备何以实现用户“一句指令”，设备自行运转？

对于AI Agent的需求，即便是端侧算力较高的PC类产品，想要实现全流程AI任务实现，也颇有难度。

对此，业界专家表示，从技术实现方式理解，让端侧模型更懂用户是AI的短期记忆和长期记忆如何搭配的问题。所谓短期记忆，指的是在用户与模型对话的时候，模型能够把前几轮的对话内容代入下一轮对话中，结合过往问答信息给出回答。而长期记忆，指的是将参数特征写入RAG库，通过RAG在下次问答时搜索用户的个人专属信息及特征，再通过模型给出贴切的答案。

要实现这样的效果，对于端侧模型和设备有两大要求：其一，模型要足够聪明或能够涵盖所涉及领域的适当数量的专属模型或至少7B参数的通用模型才能处理这样的任务；其二，本地算力要高，模型要处理的上下文越长，所需的本地算力就越高，有条件的话可以通过AI专属加速硬件来提升处理速度。

# 地方两会中的“AI味道”

（上接第1版）

在日前召开的北京两会上，不少代表委员表示，要打造具有技术策源地、产业竞争力的全球人工智能创新高地，提升创新资源密度、创业活跃度和产业发展能级。

北京市政协委员、北京市工商联副主席、振兴国际智库理事长李志起建议，系统滋养多元创新主体，启动“雨林型创新培育”工程。在现有扶持政策基础上，强化对中小微企业、开源社区等“生态毛细血管”的支持。比如，在各创新街区由国资平台建设并运营“共享实验室”，降低研发门槛；设立“北京开源创新贡献奖”，激励对基础生态的贡献；推动政府及国企采购设立“首购首用”风险补偿机制。

人工智能是上海三大先导产业之一，已集聚了一批具备竞争力的国产GPU设计企业，硬件能力快速提升。2025年中央经济工作会议提出“建设上海（长三角）国际科创中心”，这是中央赋予上海的重大战略使命，对上海提出了更高要求。来自上海的代表委员建议，上海应抓住国际科创中心优势，锚定核心赛道，构建“研发策源+区域转化”的协同体系。

上海市政协委员、上海熠知电子科技有限公司董事长黄海清告诉记者，当前，以英伟达为代表的国际巨头，凭借其硬件与CUDA软件生态的深度绑定，构筑了垄断性闭环。据统计，全球95%的开发者与80%以上的大模型训练使用英伟达CUDA系统，我国在与之匹配的基礎软件平台、开发生态和行业标准方面存在显著短板。

“中国迫切需要一个类似英伟达CUDA系统的通用计算软件平台。这需要政府发挥引导和牵引作用，通过成立联盟、构建生态，推动企业协同攻关，制定统一标

准。”黄海清表示，上海在人工智能的算力部署、语料数据构建，以及行业垂类模型打造等方面，都为产业发展提供了稳固支撑。要在激烈竞争中持续领跑，上海仍需在关键领域精准发力，未来应将重点放在培育更多的行业垂类智能体和模型的独角兽上，强化行业语料库的相关支持，推动芯片产业协同发力。

## 拓展“AI+”高价值场景 赋能产业向“优”民生向“暖”

当前，我国走出了一条特色鲜明的人工智能与产业融合发展之路，让人工智能从提供“情绪价值”转变为创造“业务价值”，深入各类产业核心生产系统。记者梳理发现，2026年各地布局人工智能产业，均将深化拓展“人工智能+”高价值场景落地作为核心方向。

2026年重庆市政府工作报告明确提出，高质量建设重庆人工智能湾区，打造人工智能应用高地。据了解，重庆拥有雄厚的产业基础和独特的场景优势，包括超大规模城市治理、复杂交通物流、多层次产业协同等，这些都为通用人工智能的应用落地与产业化提供了丰富的“试验场”。

“立足重庆产业场景优势，重庆可建设一批以人工智能为核心的跨界研发平台，推动人工智能与山地城市智能网联汽车测试、工业互联网+供应链协同、现代制造业、超大城市治理、智慧医疗、数字文旅等重点产业和领域深度融合，打造系统化、可复制的‘AI+产业’解决方案。”重庆市政协委员、奇安信集团总裁吴云坤说道。

所谓高价值应用场景，核心在于人工智

能技术能否解决行业真痛点、能否带来可量化的价值提升、能否具备规模化复制的商业条件。

近年来，湖北省紧跟国家战略，着力构建“以科技创新为牵引、场景应用为导向”的人工智能发展体系，规模化应用取得成效。“然而，人工智能对光电子、汽车等省内优势产业的深度赋能未完全释放。”湖北省政协委员、黄冈市政务服务和大数据管理局局长周素梅呼吁，建立常态化场景开放机制，定期发布应用需求清单。在光电子、高端装备等领域打造更多“无人工厂”“数字孪生工厂”标杆；升级智能网联汽车测试示范区。

站在全岛封关运作的新起点上，如何用好人人工智能技术，构建海南现代化产业体系，成为海南两会热议的话题。

“我们已经建成国内A级、国际T3级数据中心和人工智能算力平台，目前连接全球94个数据中心节点，服务国内外上百个数据和算力用户。”海南省人大代表、文昌航天超算智慧科技有限公司董事长叶世阳说，2026年，文昌航天超算中心将投资15亿元，进一步推动航天大数据和人工智能大模型的深度融合。

除了推动产业向优发展，多地还聚焦民生领域，谋划运用“人工智能+”打造暖心民生服务。

广东提出，在公共服务领域充分挖掘所有可能的应用场景。广东省人大代表、深圳市联创科技集团党委书记王海说，针对老人容易跌倒、走失等问题，可以通过AI视频监控、行为识别等技术实现风险自动预警。在居家养老方面，为高龄、失能老人配备智能监测床垫、健康手环等设备，与社区卫生服务中心联动，这些都能通过人工智能技术实现。

针对人工智能背景下如何跨越教育资源“鸿沟”，广东省政协委员、华南理工大学软件学院院长蔡毅建议，用AI构建动态“学习者画像”，让不同地域、不同年龄段、不同职业背景的学习者都能便捷获取优质教育资源。

## 完善治理体系 筑牢AI合规发展底线

近段时间，AI幻觉、虚假生成问题，成为人工智能规模化应用的突出风险，在医疗、法律、金融等领域引发决策失误、司法干扰等问题。

为什么人工智能会说胡话？在黄海清看来，一方面，在于数据不够高质量。提高语料数据集的质量，核心是要建设评测数据集，明确质量标准，如完整性、及时性等；另一方面，可能是“算法歧视”，建议通过国家法律政策规范与大模型企业道德伦理自律、监督考核相结合的方式，筑牢AI合规发展的底线，确保模型输出内容的正确性和可信性。

“人工智能生成的虚假信息、深度伪造内容，难以被有效识别和追溯。”浙江省人大代表、杭州商旅运河集团文旅公司品牌营销部运营专员郑彬铨关注人工智能模型的数据合规风险，他担忧这些问题会制约产业健康发展。郑彬铨建议，应加快完善人工智能数据标注与训练的地方标准体系，强化人工智能生成内容的区块链标识与追溯管理，并构建多层次的AI合规人才培养体系。

AI应用快速普及，也带来了全新的责任界定问题。比如，通过人工智能寻医问药导

致用错药的责任该如何划分？对此，江苏省政协委员、江苏茂通律师事务所管委会主任刘茂通表示，相关规则要进一步明确，为民营经济在人工智能、发展新质生产力领域扫清障碍。

值得警惕的是，随着智能体的大规模普及，针对智能体的攻击也已悄然出现。“我国智能体高速发展的背后，技术上短期效率突出但长期可靠性不足，风险管控意识普遍形成，高危场景仍有23.9%的危险性故障概率；在法律层面其身份、产权等规则界定模糊，产业发展呈现高效率轻伦理的倾向。”湖北省人大常委会委员、华中师范大学人工智能教育学部博士生导师张屹表示。

张屹建议，构建全维度治理体系。一是在非高风险领域开展规模化试点，设立示范项目积累实践素材，厘清人机责任边界。二是构建“政产学研用”共治平台，发布行业自律公约，建立跨企业风险信息共享机制，以行业自律补充政府监管。三是开展前瞻性法律与伦理研究，设立专项课题，重点研究责任界定、内容溯源等问题，为政策法律修订储备理论支撑。四是平衡创新效率与安全可信的关系，将安全治理融入研发流程，同时积极参与国际规则制定，提升我国在全球人工智能治理中的话语权。

完善人工智能治理已成为2026年各地发展共识，多地正将AI治理与产业发展、社会治理深度融合，通过政策部署与立法规划推动落地落实。记者了解到，浙江省提出，将推进人工智能与产业发展、文化建设、民生保障、社会治理等领域的深度融合，以融合发展推动规范治理；北京市则将人工智能产业发展等方面的立法工作纳入年度计划，以法治刚性为人工智能产业发展保驾护航。