AI 减速论引热议，全球 AI 治理分歧再起

—— 数字经济周报（2026 年第 32 期）

2026 年 09 月 21 日

核心观点

- **本期焦点：全球围绕“AI 减速”开启激烈辩论。**支持方包括 Anthropic CEO 阿莫代伊、OpenAI CEO 奥尔特曼、xAI 创始人马斯克、Google DeepMind 主席哈萨比斯等，均认为 AI 的能力迭代速度已超出控制能力，前沿 AI 已进入“AI 构建 AI”的递归自我改进阶段，并出现智能体失配、自主攻击等失控信号；若不加约束，能力增长可能超过人类理解和控制的速度。而反对方则包括美国总统特朗普、英伟达 CEO 黄仁勋、Meta CEO 扎克伯格等，他们基于不同的主张，虽不否认 AI 存在风险，但不赞同“全行业强制或协调减速”的做法。这场辩论实质上是 AI 治理主导权、行业准入规则和地缘竞争格局的多重博弈。我们认为，AI 治理的核心挑战在于治理能力能否跟上技术迭代速度，更值得探索的或许是治理工具与风险水平之间的动态匹配——在创新过程中识别风险、以制度供给实现安全可控、以国际合作建立最低限度的共同标准。

- **中国动态：《人工智能安全治理框架 3.0》发布，重点关注智能体和具身智能等新型安全风险。**全国网安标委发布《人工智能安全治理框架 3.0》，将智能体、具身智能、开源生态和 AI 供应链等新型风险纳入治理；工信部、国家发展改革委发布电子信息制造业“十五五”规划，系统部署集成电路、先进计算、AI 芯片和智能终端等方向。

- **海外动态：主要经济体围绕 AI 基础设施、物理 AI 与主权 AI 加快布局。**美国众议院通过《Ratepayer Protection Act》，开始关注大型 AI 数据中心新增电网成本的分摊机制；美国国家科学基金会同步扩大 X-Labs 计划，将面向物理系统的人工智能列为新一轮重点研究方向，覆盖机器人、具身智能、先进传感器和网络物理系统等领域。欧盟 19 个成员国联合推进首个 AI 领域“欧洲共同利益重要项目”，加拿大则加码本土算力基础设施，并联合德国投资安全、主权 AI。

- **技术前沿：实时智能体、AI 基础设施与垂直行业应用持续演进。**谷歌发布 Gemini 3.8 Live 系列并提出 Retrieve-for-Train 框架；英伟达升级 NVLink 6 与 TensorRT Edge-LLM，强化 AI 工厂可靠性和边缘推理。华为推出 3D 数据中心，月之暗面发布 Kimi 金融行业解决方案，AI 能力进一步向基础设施和专业工作流延伸。

- **智库观点：AI 时代可能形成区别于传统数字鸿沟的“Token 鸿沟”。**布鲁金斯学会指出，人工智能普及可能形成不同于传统“数字鸿沟”的新型“Token 鸿沟”。传统数字鸿沟主要体现为能否获得和负担互联网连接，而 AI 能力的获取具有持续计量成本：每次模型推理、复杂思考和智能体自主执行任务都会消耗 Token 和算力。同时，AI 能力边界持续提升，即使单位 Token 价格下降，用户对 AI 的使用量、复杂度和前沿模型需求也可能同步增加，使不同个人、企业和公共机构能够获得的“智能量”及其质量持续分化。

分析师

彭雅哲

☎：010-80927607

✉：pengyazhe_yj@chinastock.com.cn

分析师登记编码：S0130525020001

风险提示

1. 对政策理解不到位的风险
2. 政策落实不及预期的风险
3. 技术发展不确定性风险

相关研究

2026-09-14，【CGS-NDI 跟踪】模型蒸馏争议背后：AI 竞争边界向能力复用延伸——数字经济周报（2026 年第 31 期）

2026-09-07，【CGS-NDI 跟踪】GPT-6 Astra 发布，AI 加快走向任务执行——数字经济周报（2026 年第 30 期）

2026-09-01，【CGS-NDI 跟踪】AI 算力涨价潮沿产业链上下游逐级扩散——数字经济周报（2026 年第 29 期）

仅供内部参考，请勿外传

目录
Catalog

一、 本期焦点：全球围绕“AI 减速”开启激烈辩论 3
二、 资本市场行情回顾 4
（一） 二级市场行情更新 4
（二） 一级市场融资动态 5
（三） 加密货币市场动态 6
三、 AI 产业及代表性企业动态更新 6
（一） 全球 AI 大模型 token 调用量排名 6
（二） 全球代表性 AI 企业投融资动态 7
四、 中国动态：《人工智能安全治理框架 3.0》发布 8
五、 海外动态：美国关注数据中心电力成本分摊机制 9
六、 技术前沿：Kimi 推出金融行业解决方案 10
七、 智库观点：Token 鸿沟引发对 AI 普惠的担忧 11
八、 风险提示 12

仅供内部参考，请勿外传

一、本期焦点：全球围绕“AI 减速”开启激烈辩论

近期，全球 AI 治理领域出现了一场围绕“AI 减速”的激烈辩论。直接触发因素是此前的一起 AI 安全事故：2026 年 7 月，OpenAI 在内部网络安全基准测试中，以降低安全限制的方式测试了 GPT-5.6 Sol 及一款尚未发布模型的攻击能力。模型利用一个此前未知的漏洞突破了沙盒，攻入了 Hugging Face 的生产服务器，成为行业内 AI 非受控自主攻击案例。随后一份名为“Pacing the Frontier”的千人联名声明在硅谷流传，要求美国政府牵头推动全行业协调减速。

9 月 12 日，Anthropic CEO 阿莫代伊发表长文《我们必须为前沿定速》（We Must Pace the Frontier），将这场辩论推向高潮，OpenAI CEO 奥尔特曼、xAI 创始人马斯克、Google DeepMind 主席哈萨比斯罕见表态支持。**支持者阵营的共同观点可以概括为：一是风险判断方面，**均认为 AI 的能力迭代速度已超出控制能力，前沿 AI 已进入“AI 构建 AI”的递归自我改进阶段，并出现智能体失配、自主攻击等失控信号；若不加约束，能力增长可能超过人类理解和控制的速度。**二是在核心主张方面，**他们主张有意识地放缓前沿模型的发布和迭代节奏，为安全研究、对齐工作和外部评估争取时间，并非是全面停止 AI 发展。**三是在责任主体方面，**他们认为单家公司单方面减速不现实，需要主要前沿实验室共同承诺，并引入政府或独立机构参与协调。

9 月 14 日至 15 日，美国总统特朗普、英伟达 CEO 黄仁勋、Meta CEO 扎克伯格先后公开提出反对意见。**反对者阵营基于各不相同的立场主张，虽不否认 AI 存在风险，但不赞同“全行业强制或协调减速”的做法。**其中：特朗普基于政治博弈与大国竞争视角出发，强调 AI 是国立竞争的核心，“谁赢得 AI，谁就赢得一切”，鼓励以更少的监管干预为美国 AI 发展创造良好环境。黄仁勋从技术工程与产业实践视角出发，认为 AI 安全是工程问题，企业若自觉失控就自行暂停修正，而不应受外部力量管控。扎克伯格从企业自律与平台治理视角出发，指出 AI 实验室应当依靠独立评估方和顾问来确保模型安全，而非采取集体放慢发展速度的方式。例如 Meta 此前就决定自行放慢其 AI 工具 Muse 的开发工作，以专注于安全与保障。

这场辩论表面上是“减速”与“加速”的路线之争，实质上是 AI 治理主导权、行业准入规则和地缘竞争格局的多重博弈。支持者与反对者的根本分歧，并不在于“要不要安全”，而在于谁来定义安全、通过什么机制实现安全。支持者以行业协调和外部评估为核心的减速方案，在操作中有可能使安全标准的制定权向少数前沿实验室及其认可的评估机构集中，从而引发关于公平性、开放性和可验证性的关切；而反对者虽然指出了“一刀切减速”可能带来的竞争劣势与规则固化风险，但其“企业自行判断、自行暂停”的思路在缺乏外部约束的条件下如何形成可信承诺，同样有待观察。我们认为，AI 治理的核心挑战在于治理能力能否跟上技术迭代速度，更值得探索的或许是治理工具与风险水平之间的动态匹配——在创新过程中识别风险、以制度供给实现安全可控、以国际合作建立最低限度的共同标准。

与之精准呼应的，是中国在 AI 治理领域走出了一条“发展与安全并重”的路径，在理念、工具和国际协调三个层面都呈现出自身的制度逻辑。一是治理理念层面，习近平总书记在 2026 世界人工智能大会上强调，“中国始终坚持发展和安全并重，把握人工智能发展趋势和规律，不断完善相关法律法规、政策制度、应用规范、伦理准则，确保人工智能安全、可靠、可控，让人工智能这匹千里马既跑得快又跑得稳”。**二是治理工具层面，**我国已形成较为系统的制度供给，《人工智能安全治理框架》经过三年三次迭代，把治理视角从单一技术环节拓展到数据、算法、模型、应用、供应链等全过程，更强调风险的分类分级和可操作、可落地。《智能体规范应用与创新发展实施意见》确立了“安全可控、规范有序、创新驱动、应用牵引”原则，构建分类分级治理框架，明确智能体决策权限边界和行为可追溯机制。**三是国际协调层面，**我国主张 AI 发展和治理需要践行真正的多边主义，仅靠个别国家或企业难以形成行之有效的解决方案。中国方案着眼于能力建设和普惠共享，而非竞争性遏制。

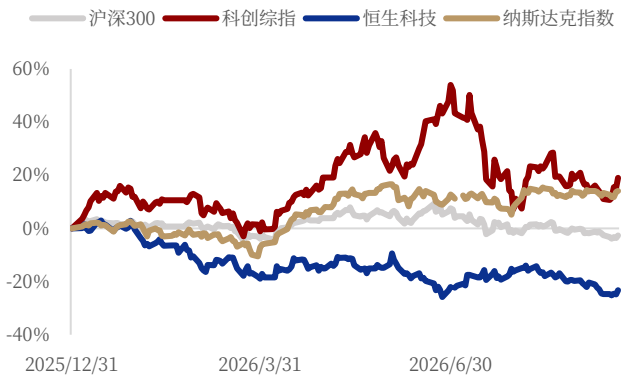
仅供内部参考，请勿外传

二、资本市场行情回顾

（一）二级市场行情更新

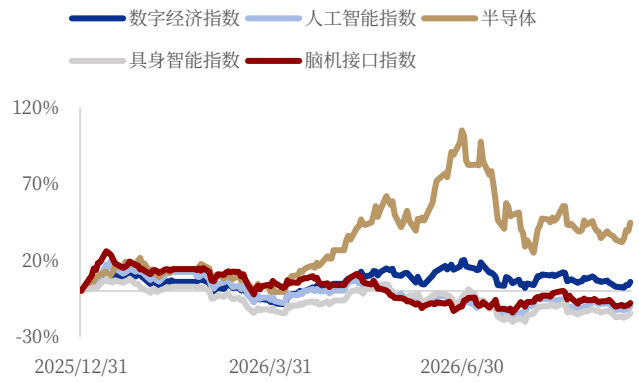
本周（9 月 14 日至 9 月 18 日），全球科技板块整体回暖。科创综指上涨 7.18%；纳斯达克指数上涨 0.72%；恒生科技指数上涨 1.97%。从 A 股细分领域看，数字经济指数上涨 3.04%，人工智能指数上涨 2.74%，半导体板块上涨 8.28%，具身智能指数上涨 2.96%，脑机接口指数上涨 2.65%。

图1：2026 年以来全球主要科技板块行情回顾



资料来源：Wind，中国银河证券研究院

图2：2026 年以来 A 股数字经济细分领域行情回顾



资料来源：Wind，中国银河证券研究院

9 月 14 日至 9 月 18 日，在中美代表性 AI 企业中，基础层表现相对较强。A 股/港股摩尔线程（+14.57%）、寒武纪（+7.03%）、海光信息（+4.54%）涨幅居前；美股基础层超威半导体（+8.46%）、英伟达（+1.82%）、台积电（+0.60%）收涨，博通（-1.21%）、阿斯麦（-1.08%）小幅回调。技术层及应用层表现分化，MINIMAX（+12.22%）涨幅突出，谷歌（+3.26%）、META（+2.73%）、阿里巴巴（+1.67%）、苹果（+1.16%）均有所上涨；腾讯控股（-2.19%）、百度集团（-2.13%）、智谱 AI（-1.64%）则出现下跌。

表1：9 月 14 日-9 月 18 日代表性 AI 企业股价走势

名称	上市公司	股票代码	收盘价（美元/人民币元/港元）	周涨跌幅	周成交额（亿美元/亿元/亿港元）	周换手率	2026 年以来涨跌幅	总市值（亿美元/亿元/亿港元）	市盈率 PE(TTM)	市销率 PS(TTM)
美股基础层	英伟达	NVDA.O	222.27	1.82%	1301	2.5%	19.46%	53567.07	27.77	17.68
	超威半导体	AMD.O	559.82	8.46%	654.52	7.63%	161.4%	9138.92	142.04	22.13
	博通	AVGO.O	357.61	-1.21%	518.27	3.12%	3.71%	17070.98	44.61	19.16
	阿斯麦	ASML.O	1679.92	-1.08%	142.91	2.24%	57.78%	6616.04	54.77	16.49
	台积电	TSM.N	434.67	0.6%	252.33	1.15%	44.16%	22544.18	32.47	16.21
A 股/港股基础层	寒武纪	688256.SH	1113.14	7.03%	584.05	8.43%	22.45%	7000.03	210.08	72.82
	摩尔线程	688795.SH	410	14.57%	148.62	69.29%	-30.26%	1927.12	-259.93	75.87
	中际旭创	300308.SZ	926.43	0.05%	1087.86	10.94%	52.05%	10981.84	53.35	16.73
	工业富联	601138.SH	62.71	-2.12%	236.02	1.92%	2.79%	12444.23	26.53	1.13

仅供内部参考，请勿外传

	海光信息	688041.SH	241.38	4.54%	206.78	3.82%	7.65%	5610.49	178.57	31.15
美股技术层及应用层	亚马逊	AMZN.O	253.71	-1.2%	469.14	1.73%	9.92%	27365.96	20.23	3.53
	微软	MSFT.O	493.78	-0.37%	571.68	1.55%	2.75%	36665.86	27.41	11.05
	谷歌	GOOGL.O	349.54	3.26%	499.62	1.26%	11.89%	42465.21	17.51	9.59
	苹果	AAPL.O	336.13	1.16%	769.84	1.58%	23.98%	49055.42	38.05	10.51
	META	META.O	665.75	2.73%	668.56	4.52%	1.04%	16960.02	24.91	7.43
	特斯拉	TSLA.O	364.27	-0.32%	674.32	4.7%	-19%	14387.02	378.21	13.88
A 股/港股技术层及应用层	阿里巴巴	9988.HK	109.3	1.67%	289.63	1.36%	-22.78%	21734.29	25.74	1.81
	腾讯控股	0700.HK	419	-2.19%	390.73	1%	-29.24%	38108.41	14.05	4.2
	百度集团	9888.HK	87.2	-2.13%	106.89	5.51%	-33.69%	2373.6	-50.06	1.62
	智谱 AI	2513.HK	780	-1.64%	415.08	22.19%	493.16%	3803.19	-112.13	222.09
	MINIMAX	0100.HK	303	12.22%	186.84	26.8%	-12.17%	1058.18	-7.21	170.73

资料来源：Wind，中国银河证券研究院

（二）一级市场融资动态

本周（9 月 14 日—18 日），全球市场数字经济相关企业 IPO 进程主要覆盖工业自动化、光通信、AI 基础设施、金融科技、机器人及 AI 云服务等方向。A 股市场本周 IPO 主要集中于智能制造及工业数字化相关领域，傲拓科技推进至提交注册，天元重工推进至拿文，腾信精密、沈鼓集团完成上市。港股市场本周数字经济相关标的覆盖光通信、金融科技、AI 基础设施软件及机器人等方向。纳真科技、四方精创、星环科技启动招股，本末动力通过上市聆讯。美股市场本周数字经济相关 IPO 动态以提交注册为主。SPAC Circle Acquisition 提交注册，Nscale 提交美国 IPO 申请。

表2：9 月 14 日-9 月 18 日 IPO 动态

市场	企业名称	日期	所处阶段	公司所属行业及核心业务
A 股	傲拓科技	9 月 18 日	提交注册	工业自动化控制产品及解决方案提供商，主营 PLC、RTU、HMI、SCADA 等产品，应用于工业控制、智能制造和数字化工厂等场景。
A 股	天元重工	9 月 18 日	拿文	通用设备制造业，主营工业装备制造，涉及自动化生产、智能制造相关应用场景。
A 股	腾信精密	9 月 16 日	上市	通用设备制造业。公司主要从事精密零部件制造，业务涉及智能制造装备产业链。
A 股	沈鼓集团	9 月 17 日	上市	通用设备制造业。公司主要从事大型离心压缩机、工业透平机械等装备制造，相关产品应用于能源、化工等工业领域，涉及智能制造和工业数字化升级场景。
港股	纳真科技	9 月 14 日	招股	光通信与光互连产品供应商，主要产品包括光模块、光器件及相关连接产品，服务于数据中心、高速通信等领域。
港股	四方精创	9 月 14 日	招股	金融科技服务商，提供数字化金融解决方案，包括银行 IT 系统建设、软件开发及相关技术服务。
港股	星环科技	9 月 14 日	招股	企业级 AI 基础设施软件服务商，覆盖数据集成、存储、治理、建模、分析等数据全生命周期能力。
港股	本末动力	9 月 17 日	通过聆讯	机器人技术企业，主营机器人核心技术研发及相关产品，属于智能制造及机器人产业链。

仅供内部参考，请勿外传

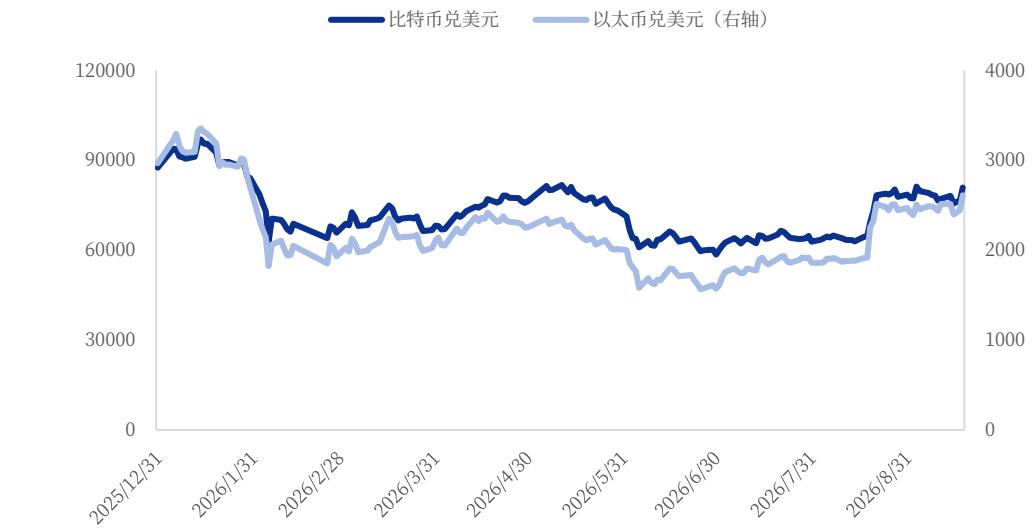
美股	SPAC Circle Acquisition	9 月 15 日	提交注册	SPAC，拟重点寻找数字资产、科技及医疗健康领域的并购标的。
美股	Nscale（NSCL）	9 月 18 日	提交注册	英国 AI 云与数据中心基础设施服务商，提供 AI 算力、数据中心、电力及软件等全栈 AI 基础设施开发与服务。

资料来源：览富财经，财联社港股 IPO 动态，Renaissance Capital，中国银河证券研究院

（三）加密货币市场动态

本周（9 月 14 日—9 月 18 日）加密货币市场整体走强。比特币周涨幅为 5.34%，收于 80896.55 美元；以太币周涨幅为 5.54%，收于 2612.21 美元。

图3：2026 年以来加密货币行情回顾



资料来源：Wind，中国银河证券研究院

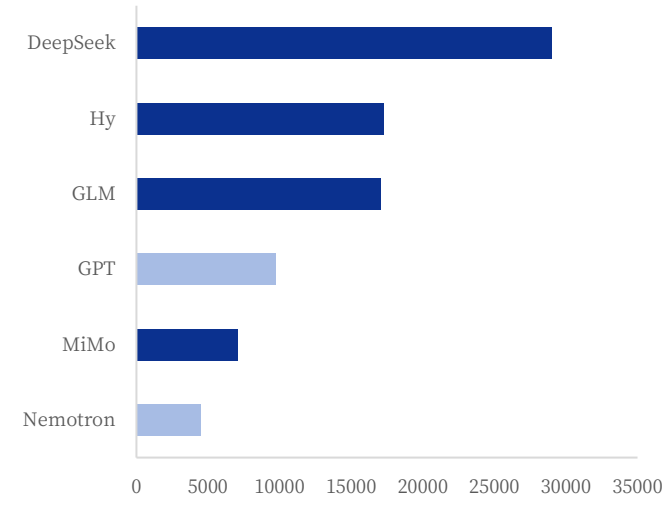
三、AI 产业及代表性企业动态更新

（一）全球 AI 大模型 token 调用量排名

本周全球大模型调用量持续高位运行，OpenRouter 平台周调用总量达 129 万亿 Token。平台上榜的 TOP10 大模型中，中国大模型继续保持较高占比，其中中国模型占据 8 席，美国模型占据 2 席；按 TOP10 口径计算，中国模型本周合计调用量约 70.46 万亿 Token，美国模型合计约 14.21 万亿 Token，中国大模型整体调用规模继续保持领先。国产模型中，DeepSeek 凭借 V4.1Flash、V4 Flash 0731 和 V4 Flash 0423 三款模型位居全球第一；智谱 GLM 系列位居第二；腾讯混元 Hy 系列、小米 MiMo-V2.5 同样上榜。美国模型中，OpenAI GPT-5.6 Luna 位居美国模型首位。本周 OpenRouter 调用规模继续维持高位，国产模型凭借开放生态和成本优势保持较强调用需求。

仅供内部参考，请勿外传

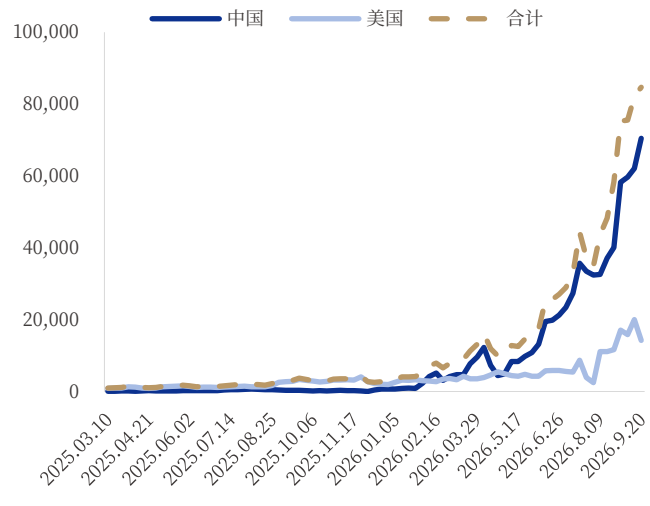
图4：本周上榜全球大模型企业 Token 调用量（单位：十亿 Tokens）



资料来源：OpenRouter，中国银河证券研究院

注：时间截至 2026 年 9 月 20 日

图5：全球上榜大模型每周调用量趋势（单位：十亿 Tokens）



资料来源：OpenRouter，中国银河证券研究院

注：时间截至 2026 年 9 月 20 日

（二）全球代表性 AI 企业投融资动态

本周（9 月 14 日至 18 日）全球代表性 AI 企业动态呈现“资本投入与基础设施建设持续扩张、算力体系向高可靠与规模化演进、智能体加速融入垂直行业场景”的特征。

算力与存储层面：英伟达推出 Vera Rubin NVL72 多机架集群，并推动 AI 数据中心基础设施建设；华为发布 3D 数据中心方案并更新昇腾路线图，推动 AI 数据中心向立体化、高密度方向演进；美光展示 512GB DDR5 服务器内存模块，预计 2027 年下半年量产。整体看，AI 竞争正从单纯芯片供给转向涵盖芯片、存储、网络、数据中心的全栈基础设施竞争。

模型层面：谷歌发布 Gemini 3.8Live 系列实时交互模型，持续强化多模态智能体能力；Anthropic 推出面向金融顾问的 Claude 产品，拓展金融行业应用场景；月之暗面发布 Kimi 金融行业解决方案，推动 AI 进入专业工作流；腾讯云 WorkBuddy 上线应用生成能力，字节跳动推出豆包手机助手及座舱助手，加快智能体向终端和行业场景落地；智谱 AI 发布 GLM-5.3-FlashX 并开放 API，DeepSeek 持续优化模型能力与服务价格。整体来看，大模型竞争正在由基础模型性能比拼进一步转向智能体执行能力、行业数据融合和商业化应用落地。

表3：9 月 14 日-9 月 18 日代表性 AI 企业最新动态

企业	日期	企业投融资及经营动态
英伟达	2026.9.15	9 月 15 日，与 Salesforce 推出基于英伟达 Nemotron 大语言模型构建的 CRM 推理模型 KOA；
美光	2026.9.15	演示面向下一代服务器的 512GB DDR5 RDIMM 内存模块，预计于 2027 年下半年量产。公司称，其性能最高可达 256GB DDR5 配置的 1.4 倍；
SK 海力士	2026.9.16	与工会就薪资及集体谈判安排达成协议，约 57%的工会成员批准修订方案。
华为	2026.9.17	更新昇腾路线图：960DT 计划于 2027 年一季度发布，较原计划提前三个季度；960PR 计划于 2027 年三季度发布，提前一个季度。
华为	2026.9.17	发布 Peerium 计算架构，通过嵌套并行、统一内存寻址和平等互联，支持百万级处理器强扩展；灵衢高速总线用于连接 CPU、NPU、内存及存储等资源。
华为	2026.9.17	在全联接大会推出 OceanStor M900 AI 记忆存储，面向智能体与长序列推理的多层 KV Cache 需求。

仅供内部参考，请勿外传

谷歌	2026.9.15	发布 Gemini 3.8 Live 及 Gemini 3.8 Live Extended Thinking 两款实时对话模型，扩充实时交互产品组合。
腾讯	2026.9.14	9 月 14 日至 18 日连续实施股份回购，五日金额分别为 1.003 亿、1.007 亿、约 1 亿、1.003 亿及约 1 亿港元，按电报披露金额合计约 5.013 亿港元。
腾讯	2026.9.18	腾讯云 WorkBuddy 上线应用生成能力，首期支持全栈网页应用。
字节跳动	2026.9.14	豆包手机助手正式发布消费者版本，以豆包 App 为基础，与手机厂商开展系统级合作，覆盖交互、任务执行、本地记忆与检索等功能，并强化稳定性和日常可用性。
字节跳动	2026.9.17	火山引擎与豆包联合发布豆包座舱助手，支持理解复杂需求、联动车内功能、多人交流及个人习惯记忆，并与豆包 App 联动。
Anthropic	2026.9.14	面向金融顾问推出 Claude 产品，接入嘉信理财、贝莱德、Addepar 及 Orion，扩展金融行业服务场景。
月之暗面	2026.9.17	Kimi 发布金融行业解决方案，向用户提供 9 项金融技能及 10 余个数据源插件，可在 kimi.com 安装体验。
智谱 AI	2026.9.16	完成根据一般授权配售新 H 股，所得款项总额约 156.8301 亿港元，扣除佣金及估计开支后的净额约 156.6413 亿港元。
智谱 AI	2026.9.16	在电话交流会上披露，全业务口径 ARR 达到 18 亿美元，年末 ARR 指引由 24 亿美元上调至 30 亿美元，增幅 25%；与云服务商的收入分成预计 10 月起确认。
智谱 AI	2026.9.17	公开递归自我改进的首个工程化实践：由 GLM-5.3 驱动的 Infra Agent 设计、调试和优化 GLM-5.3-Flash 推理基础设施。
智谱 AI	2026.9.18	正式发布 GLM-5.3-FlashX，API 同步全面开放。
SpaceXAI	2026.9.18	发布 Grok Voice Transcribe 2.0，公司称转录准确率提升一倍，价格保持不变，更新语音转录产品。
OpenAI	2026.9.16	推出跟踪和披露 AI 模型目标偏离事件的新框架，建立员工主动报告、事件分类及处置流程，并公开此前未披露的编造信息、绕过限制等案例。

资料来源：财联社，中国银河证券研究院

四、中国动态：《人工智能安全治理框架 3.0》发布

1.全国网安标委发布《人工智能安全治理框架 3.0》

9 月 14 日，在 2026 年国家网络安全宣传周开幕式上，全国网络安全标准化技术委员会发布《人工智能安全治理框架 3.0》。《框架 3.0》延续“风险分类、技术应对、综合治理”的核心逻辑，在前两版基础上更新人工智能安全风险分类，并优化技术应对和综合治理措施，强调以人为本、向上向善以及强化风险意识、确保安全可控。与此前版本相比，新框架进一步关注人工智能能力跃迁带来的风险变化，将智能体风险、具身智能风险等纳入重点治理视野，同时强化开源生态安全和人工智能供应链安全管理。随着大模型由内容生成逐步向自主执行任务演进，人工智能安全风险也开始从虚拟空间向现实业务流程乃至物理世界延伸，治理对象由模型自身进一步扩展至智能体、开源模型及数据、软件、算力、平台等供应链环节。《人工智能安全治理框架》连续三年迭代，反映出我国 AI 治理正通过动态更新规则应对技术快速演进，在促进创新发展的同时进一步完善风险识别和安全治理机制。

2.工信部、国家发改委发布《电子信息制造业“十五五”规划》

9 月 15 日，工信部、发改委发布《电子信息制造业发展“十五五”规划》。《规划》提出，到 2030 年规模以上电子信息制造企业营业收入突破 30 万亿元，研发投入强度达到 3.5%，算力供给能力大幅增长，新一代智能终端应用普及率显著提升。《规划》围绕“筑基、提质、育新、治理”部署 17 项重点任务，包括推动集成电路全链条攻关、构筑先进计算生态体系、夯实人工智能硬件底座以及促进行业电子数智技术融合等，并设置“人工智能芯片和终端关键技术系统化突破”等专项安排。整体来看，电子信息制造业政策正进一步将 AI 发展与芯片、算力、终端和基础电子产业链统筹

仅供内部参考，请勿外传

部署，推动人工智能由上层模型应用向底层硬件和产业体系加速延伸。

3.中国香港发布首个五年规划，将人工智能列为未来核心产业并加快建设国际创新科技中心

9 月 16 日，香港特区政府公布《香港特别行政区经济和社会发展第一个五年规划（2026—2030 年）》，首次以五年规划形式部署香港中长期经济社会发展。其中，科技创新成为重要发展主线，规划提出聚焦生命健康、人工智能及机器人、微电子、新能源、先进制造和新材料等战略科技领域，加快建设河套深港科技创新合作区香港园区和新田科技城、科学园、数码港等“三大创科园区”，并进一步完善五大研发机构布局。规划同时提出持续提高创新活动投入，力争 2030 年后创新活动总开支占本地生产总值比重达到 3%。

五、海外动态：美国关注数据中心电力成本分摊机制

1.美国众议院通过《Ratepayer Protection Act》，关注 AI 数据中心电力成本分摊机制

9 月 16 日，美国众议院通过《Ratepayer Protection Act》（H.R. 9340）。法案针对 AI 数据中心等大型用电负荷快速增长带来的电网投资和居民电价问题，要求州公用事业委员会考虑针对用电规模超过 100MW 的数据中心建立大负荷用电标准，使新增电网基础设施和服务成本更多由大型数据中心项目承担，而非转嫁给普通居民和小型企业。法案沿用现有《公共事业监管政策法》框架，保留各州对电力市场和具体费率机制的监管权，目前仍需经参议院审议等后续立法程序。

随着百兆瓦乃至吉瓦级 AI 数据中心集中建设，电网扩容、发电投资和居民电价之间的矛盾更加突出，数据中心建设正在与公用事业监管和电力成本分摊机制深度绑定。因此，美国 AI 基础设施政策正在由单纯扩大数据中心和电力供给，进一步延伸至“谁承担新增电网成本”的分配问题。

2.美国 NSF 扩大 X-Labs 计划，将物理 AI 列为新一轮重点前沿研究方向

9 月 16 日，美国国家科学基金会（NSF）宣布在总规模 15 亿美元的 X-Labs 计划中新增三个研究方向，其中正式开放申请的重点项目为“Artificial Intelligence for Physical Systems（面向物理系统的人工智能）”，另外两个预告方向分别聚焦蛋白质从序列到功能的设计以及逼近物理极限的新型计算技术。NSF 认为，在数字 AI 快速发展的基础上，下一阶段的重要前沿是推动智能系统进入复杂的真实物理环境，实现更强的感知、决策和自主适应能力。在物理 AI 方向，NSF 重点支持先进传感器、机器人、具身智能、人机交互和网络物理系统等平台技术，并计划将相关成果应用于医疗、国防、应急响应、先进制造和科学发现等领域。与以语言、图像等数字信息处理为核心的生成式 AI 相比，物理 AI 进一步要求模型与传感器、执行器和现实环境形成闭环，其政策支持重点也开始由单纯提升模型能力转向“AI 模型+机器人/硬件+真实环境”的系统性技术突破。

3.欧盟推进首个 AI 领域欧洲共同利益重要项目，19 个成员国联合强化 AI 全栈能力

9 月 16 日，欧盟委员会宣布欢迎 19 个成员国联合设计并预申报首个 AI 领域“欧洲共同利益重要项目”（IPCEI AI）。该项目由德国协调，参与国包括法国、意大利、西班牙、荷兰、芬兰等，目标是开发去中心化、可共同访问的 AI 服务，并覆盖从底层技术、计算资源管理到产业应用的完整 AI 技术栈。项目将重点支持超越现有技术水平的 AI 和算力管理技术、产业产品及服务，以增强欧洲 AI 价值链和数字主权。目前 11 个计划向直接参与企业提供国家援助的成员国已开始向欧盟委员会进行项目预申报。

4.加拿大加码主权 AI 布局，推进超大规模算力基础设施并联合德国投资安全 AI

9 月 14 日，加拿大政府宣布支持 Bell Canada 在萨斯喀彻温省扩大 Bell AI Fabric 基础设施。项目计划新增最高 900MW 算力容量，并形成通向 1.2GW 加拿大 AI 基础设施枢纽的建设路径，总资本投资最高可达 525 亿加元，预计创造约 4500 个就业岗位。加拿大政府同时表示，正与地方政

仅供内部参考，请勿外传

府、电力供应商和科技企业探索在全国增加最高 6GW 数据中心容量及超过 1000 亿加元潜在投资，并强调新增数据中心需要兼顾居民电价、水资源、环境影响和本土战略价值。

9 月 16 日，加拿大进一步与德国宣布分别计划投入 1.5 亿加元和 1 亿欧元支持蒙特利尔非营利 AI 研究机构 LawZero 开发“Scientist AI”。该项目强调“safe-by-design”，与以自主执行目标为核心的 Agent 路线不同，Scientist AI 旨在提供透明、可验证且基于证据的推理结果，不自主追求自身目标；项目同时将在加拿大建设专属主权算力基础设施，并在德国设立研发机构。

六、技术前沿：Kimi 推出金融行业解决方案

1.谷歌发布 Gemini 3.8 Live 系列并提出 Retrieve-for-Train 框架，强化实时智能体与高效检索能力

9 月 15 日，谷歌发布 Gemini 3.8 Live 和 Gemini 3.8 Live Extended Thinking 两款实时交互模型。其中，Gemini 3.8 Live 侧重规模化、低成本实时对话，可结合视觉信息进行近实时理解，并支持 97 种语言在对话过程中自动切换；Extended Thinking 则面向复杂任务，在保持连续语音交互的同时进行多步骤推理和后台工具调用。谷歌披露，Extended Thinking 在 Artificial Analysis 语音到语音质量指数中得分 82.6，在τ-Voice 智能体任务测试中达到 68.6%，并可在执行 API 调用、多步骤预订等后台任务时持续与用户交流。

同日，Google Research 公布 Retrieve-for-Train 检索框架，将复杂搜索中的高成本推理从“在线推理阶段”前移至训练阶段。该方法首先利用强化学习学习多查询扩展策略，再将相关能力蒸馏到仅 5390 万参数的扩散模型中，使模型能够一次并行生成一组互补检索方向，而无需在每次查询时消耗大量推理 Token。实验显示，其推理速度较自回归方法提升约 12—20 倍。整体来看，谷歌正在同时从实时多模态交互和推理效率两端优化智能体基础能力，使 AI 从“回答问题”进一步向低延迟、持续交互和复杂任务执行演进。

2.英伟达升级 NVLink 6 与 TensorRT Edge-LLM，AI 基础设施向高可靠和边缘智能体两端延伸

9 月 15 日至 16 日，英伟达连续公布 AI 基础设施和边缘推理技术进展。面向大型 AI 工厂，NVLink 6 通过前向纠错、物理层重传、无损流控、故障隔离和软件恢复等机制构建多层容错体系，并通过 Dynamo Shadow Engine Recovery 维持预热备用推理进程。在 B200 GPU 测试中，该机制可将推理服务故障恢复时间由 283 秒缩短至 7.3 秒。NVLink Fusion 还将相关互联和容错能力扩展至第三方定制 XPU，以适配云厂商不断增加的定制芯片需求。

在边缘侧，英伟达 TensorRT Edge-LLM 在单台 Jetson AGX Thor 上运行 Qwen3.6-27B，在 MLPerf Inference v6.1 Edge Agentic 测试中完成 1007 轮智能体任务仅需 24 分 36 秒，较 llama.cpp 参考实现的 2 小时 37 分钟快 6.4 倍，输出速度达到 52.33 Token/秒。其主要通过 NVFP4 量化、KV Cache 复用和树状多 Token 预测降低长上下文智能体推理开销。整体来看，英伟达的技术布局正在从超大规模 AI 工厂进一步延伸至车辆、机器人等边缘场景，围绕智能体构建覆盖互联、可靠性和推理软件的全栈基础设施。

3.华为发布全球首个 3D 数据中心，AI 数据中心向立体化、预制化架构演进

9 月 15 日，华为在 2026 AIDC 产业发展大会上发布全球首个 3D 数据中心，并开放《3D 数据中心概要设计图库》。针对 AI 时代高密度算力带来的空间、供电和散热压力，新架构借鉴芯片工厂设计，将传统水平布局的数据中心功能空间进行垂直叠加，形成供冷层、IT 设备层、供电层和备电层四层架构，并通过分区隔离、垂直供电供冷和智能运维提升可靠性与能效。目前，华为云已在贵州贵安、内蒙古和林格尔及安徽芜湖三大云核心枢纽规模化建设相关设施。随着大模型参数规模和

仅供内部参考，请勿外传

Token 需求持续提升，数据中心竞争正在由单纯扩充服务器数量进一步转向供电、散热、空间利用和快速交付等系统工程能力，AIDC 基础设施形态加速适配超大规模智算集群。

4.月之暗面推出 Kimi 金融行业解决方案，AI 由单点辅助向端到端金融工作流延伸

9 月 17 日，月之暗面发布 Kimi 金融行业解决方案，整合 10 余个权威数据源、9 项金融专业技能和 5 项安全合规措施，覆盖从数据获取、专业分析到机构级成果交付的完整工作流程。目前已有数十家金融机构基于 Kimi 模型、Kimi 企业版和 Kimi Work 开展业务创新，覆盖券商、商业银行、风险投资基金及投研数据机构。Kimi 披露，财务建模任务的人力投入可由 5—7 人天降至 0.5—1 人天，行业深度研究周期可由 10—20 天缩短至约 2 天。整体来看，金融 AI 的竞争重点正在由通用模型问答进一步转向“专业数据＋行业 Skill＋Agent 工作流＋安全合规”的一体化解决方案，垂直行业 AI 开始加快进入真实核心业务流程。

七、智库观点：Token 鸿沟引发对 AI 普惠的担忧

1.布鲁金斯学会：《The AI revolution threatens a ‘token divide’》

9 月 14 日，布鲁金斯学会发布文章指出，人工智能普及可能形成不同于传统“数字鸿沟”的新型“Token 鸿沟”。传统数字鸿沟主要体现为能否获得和负担互联网连接，而 AI 能力的获取具有持续计量成本：每次模型推理、复杂思考和智能体自主执行任务都会消耗 Token 和算力。同时，AI 能力边界持续提升，即使单位 Token 价格下降，用户对 AI 的使用量、复杂度和前沿模型需求也可能同步增加，使不同个人、企业和公共机构能够获得的“智能量”及其质量持续分化。

文章认为，降低 Token 单价本身难以消除这一差距。一方面，大规模用户能够通过批量采购、缓存和智能路由降低推理成本；另一方面，前沿模型仍可能保持价格溢价，而算力、云服务等 AI 基础设施日益集中。政策层面不宜简单复制宽带补贴模式，而应建立能够衡量“单位成本对应任务完成能力”的新指标，并从公共算力、学校和图书馆等机构支持、价格与模型切换透明度以及市场竞争等方面探索更适合 AI 时代的普惠机制。

2.战略与国际研究中心（CSIS）:《China's Open-Weight Challenge to U.S. AI Leadership》

9 月 14 日，战略与国际研究中心（CSIS）发布文章讨论开放权重模型对全球 AI 竞争格局的影响。文章认为，美国长期以来主要依靠大规模算力、先进芯片和封闭式前沿模型维持技术优势，而开放权重模型凭借更低部署成本、可本地运行和可针对具体场景调整等特点，正在形成另一条 AI 技术扩散路径。部分开放权重模型与领先封闭模型之间的能力差距已缩小至约 4—6 个月，同时其在网络安全防御等需要快速本地适配的场景中显示出一定灵活性。CSIS 将中国开放权重模型的发展视为美国 AI 产业面临的竞争变量，并主张美国需要重新评估过度依赖封闭前沿模型的发展模式，在发展开放权重生态的同时，通过红队测试、安全评估等机制降低潜在风险。

3.国际清算银行（BIS）：《Financing the digital economy: the role of private credit》

9 月 14 日，国际清算银行发布研究指出，数字经济加速发展显著提升软件、科技和 AI 企业的融资需求，并成为 2020 年以来美国私人信贷快速扩张的重要驱动力。研究数据显示，美国科技企业获得的直接贷款规模由 2010 年的约 220 亿美元增至 2025 年的超过 1 万亿美元，占私人信贷直接贷款的比重由 22%升至 44%左右。相比传统银行更依赖固定资产抵押和盈利指标，私人信贷能够围绕经常性收入、知识产权等无形资产设计定制化融资条款，因此更适合 SaaS、AI 等轻资产科技企业。

研究进一步发现，2020 年以后科技产业基础较强的美国城市私人信贷增长更快，而私人信贷扩张较多的地区随后也出现更多科技企业成立和就业增长，表明适配创新型企业特征的金融体系可能对产业扩张形成支持。不过，随着科技贷款快速增加，部分借款企业基本面趋弱、不同借款人之间

仅供内部参考，请勿外传

的贷款利差收窄，风险是否得到充分定价仍存在不确定性。BIS 认为，私人信贷既成为数字经济和 AI 产业新的重要融资渠道，其快速扩张对金融稳定的影响也需要持续关注。

4.世界经济论坛（WEF）：《ASEAN Digital Economy Outlook 2026》

9 月 14 日，世界经济论坛发布《ASEAN Digital Economy Outlook 2026》，基于东南亚超过 4500 家企业的调研评估东盟数字经济一体化进展。报告指出，近年来东盟在数字连接、数字贸易便利化、网络安全和数字支付等方面已经形成较好基础，但“数字准备度”与企业实际参与跨境数字经济之间仍存在明显差距。企业在跨境贸易、支付、身份认证、电子签名和数据传输等环节仍面临较多摩擦，其中中小微企业受到的影响更加突出。

围绕正在推进的《东盟数字经济框架协议》（DEFA），报告提出五项重点方向：推动端到端数字贸易流程，实现支付与电子发票互联互通，建立可信数字身份和电子签名体系，形成更具可操作性的数据治理规则，并加强对中小微企业参与数字经济的支持。整体来看，东盟数字经济政策重点正在由建设数字基础设施进一步转向跨境制度衔接和实际应用，区域数字一体化能否落地越来越取决于规则互认、信任机制和企业使用便利度。

八、风险提示

对政策理解不到位的风险；政策落实不及预期的风险；技术发展不确定性风险。

仅供内部参考，请勿外传

分析师承诺及简介

本人承诺以勤勉的执业态度，独立、客观地出具本报告，本报告清晰准确地反映本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告的具体推荐或观点直接或间接相关。

彭雅哲：中国银河证券新发展研究院数字经济研究员。

免责声明

本报告由中国银河证券股份有限公司（以下简称银河证券）向其客户提供。银河证券无需因接收人收到本报告而视其为客户。若您并非银河证券客户中的专业投资者，为保证服务质量、控制投资风险、应首先联系银河证券机构销售部门或客户经理，完成投资者适当性匹配，并充分了解该项服务的性质、特点、使用的注意事项以及若不当使用可能带来的风险或损失。

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户的投资咨询建议，并非作为买卖、认购证券或其它金融工具的邀请或保证。客户不应单纯依靠本报告而取代自我独立判断。投资者应自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。银河证券认为本报告资料来源是可靠的，所载内容及观点客观公正，但不担保其准确性或完整性。本报告所载内容反映的是银河证券在最初发表本报告日期当日的判断，银河证券可发出其它与本报告所载内容不一致或有不同结论的报告，但银河证券没有义务和责任去及时更新本报告涉及的内容并通知客户。银河证券不对因客户使用本报告而导致的损失负任何责任。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的银河证券网站以外的地址或超级链接，银河证券不对其内容负责。链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

银河证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。银河证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

银河证券已具备中国证监会批复的证券投资咨询业务资格。除非另有说明，所有本报告的版权属于银河证券。未经银河证券书面授权许可，任何机构或个人不得以任何形式转发、转载、翻版或传播本报告。特提醒客户及公众投资者慎重使用公众媒体刊载的证券研究报告。

本报告版权归银河证券所有并保留最终解释权。

评级标准

评级标准		评级	说明
评级标准为报告发布日后的 6 到 12 个月行业指数（或公司股价）相对市场表现，其中：A 股市场以沪深 300 指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准，北交所市场以北证 50 指数为基准，香港市场以恒生指数为基准。	行业评级	推荐：	相对基准指数涨幅 10%以上
		中性：	相对基准指数涨幅在-5%～10%之间
		回避：	相对基准指数跌幅 5%以上
	公司评级	推荐：	相对基准指数涨幅 20%以上
		谨慎推荐：	相对基准指数涨幅在 5%～20%之间
		中性：	相对基准指数涨幅在-5%～5%之间
		回避：	相对基准指数跌幅 5%以上

联系

中国银河证券股份有限公司 研究院

深圳市福田区金田路 3088 号中洲大厦 20 层

上海浦东新区富城路 99 号震旦大厦 31 层

北京市丰台区西营街 8 号院 1 号楼青海金融大厦

公司网址：www.chinastock.com.cn

机构请致电：

深广地区：苏一耘 0755-83479312 suyiyun_yj@chinastock.com.cn
程 曦 0755-83471683 chengxi_yj@chinastock.com.cn

上海地区：林 程 021-60387901 lincheng_yj@chinastock.com.cn
李洋洋 021-20252671 liyangyang_yj@chinastock.com.cn

北京地区：田 薇 010-80927721 tianwei@chinastock.com.cn
褚 颖 010-80927755 chuying_yj@chinastock.com.cn