

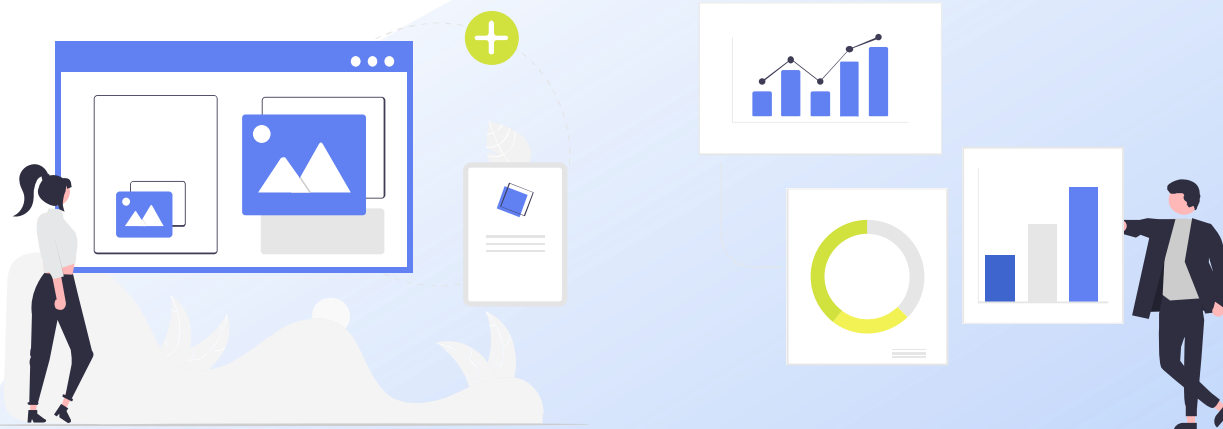


极客邦科技双数研究院

InfoQ 研究中心

# 中国软件技术发展洞察 和趋势预测报告2026

InfoQ 极客传媒





# 研究数据说明——研究方法

## 1

### 桌面研究

通过对行业公开信息进行桌面研究，资料包括但不限于专业机构研究报告、技术领域书籍、技术社区网站、技术企业官网及专家公开演讲内容等；

## 2

### 技术模型

通过技术专利数量、技术诞生时间、技术舆论指数指标，进行技术发展阶段类型模型设计，作为细分技术发展阶段展现；

## 3

### 专家访谈

InfoQ 研究中心针对本次研究定向邀请了多位技术企业管理层、垂直技术领域专家、技术从业者、产业从业者等各方面专家；



# 研究数据说明——数据来源

本次研究通过收集并整理160+细分技术领域的技术专利数量、技术诞生时间、技术舆论指数、技术融资事件作为2025年重要技术发展回顾与技术发展特征解读的基础。

## 01

### 技术专利数量

技术专利数据来自国家知识产权局专利数据库。使用对应技术领域关键词进行检索，截止时间2025年12月31日。

## 02

### 技术诞生时间

考虑到数据可得性与统一性，技术诞生时间使用知网论文库进行相关技术领域论文最早收录年份统计。

## 03

### 技术舆论指数

技术舆论指数数据来源为搜索引擎和主要技术媒体，其中包括百度、InfoQ中文站和CSDN社区等网站，具体数据通过对应技术领域关键词搜索涉及的文章数量获得，指数通过加权获得。

## 04

### 技术融资事件

技术融资时间数据来源为IT桔子融资事件库和36氪项目库，分别按照不同细分技术领域进行融资时间、融资金额、投资机构的统计。

CONTENT

# 目录

01 | 2025年技术市场回顾

---

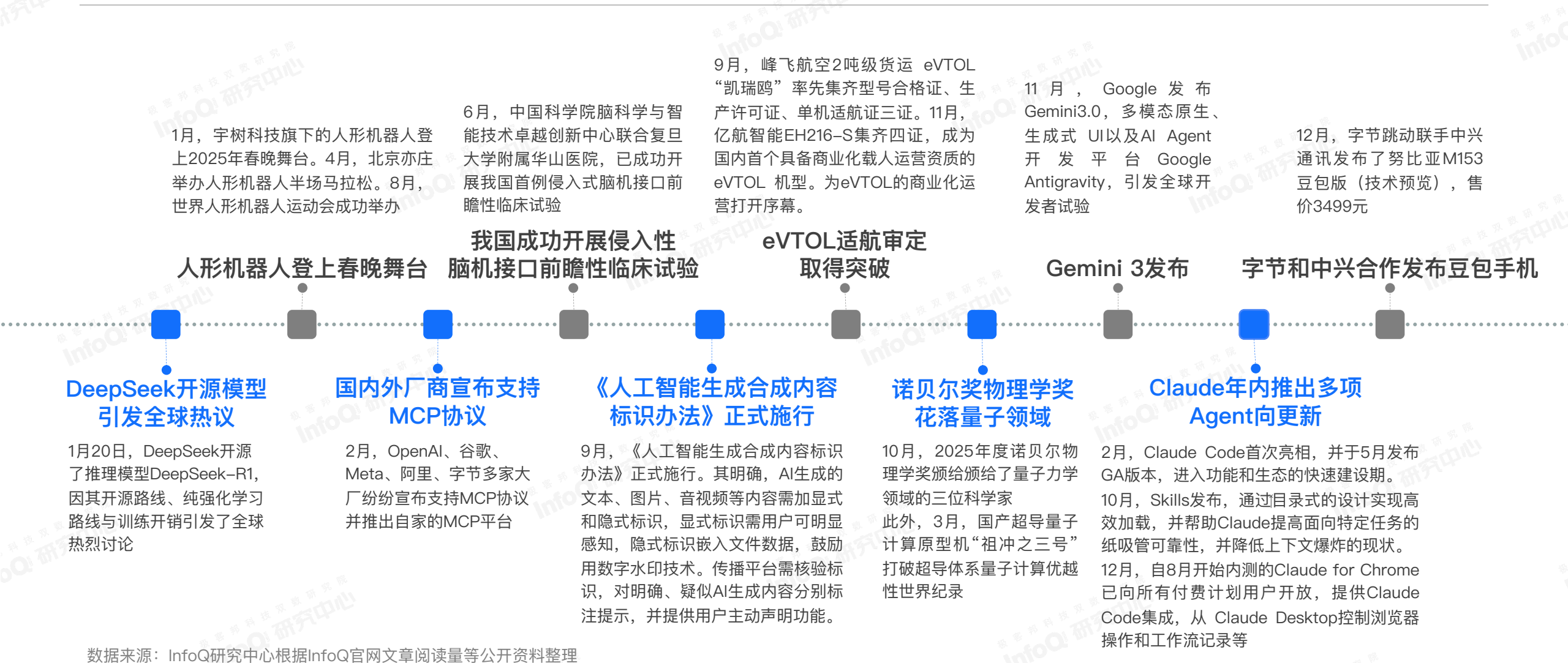
02 | 2026年技术趋势展望

---

01

# 2025年技术市场回顾

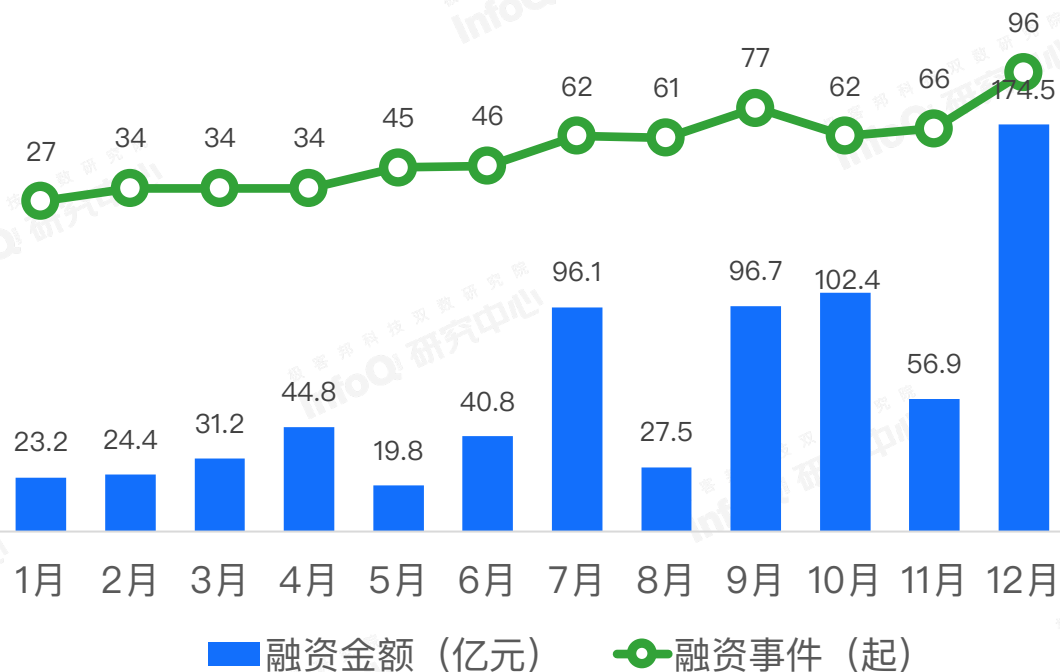
# 2025年技术领域热点新闻：AI、低空飞行、量子技术、人形机器人等



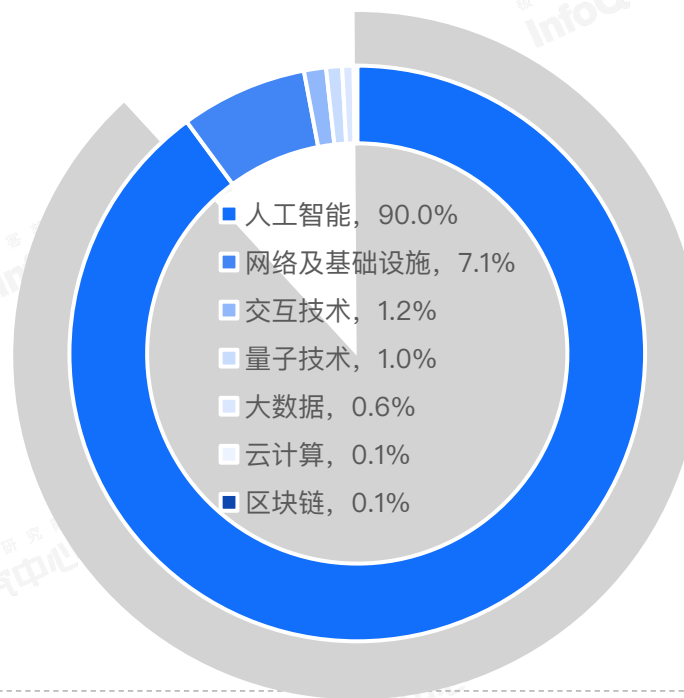
## 投融资：2025年技术领域共融资738.4亿元

- 2025年技术领域融资事件共644起，共融资738.4亿人民币。
- 从细分领域来看，人工智能领域吸纳资金能力继续一骑绝尘。2025年全年，人工智能领域融资金额占据整体技术领域融资金额的88.1%，相较2024年上浮5.3个百分点。

2025年融资金额与事件数量统计



2025年技术领域融资金金融统计



人工智能领域  
融资额占比达  
到90.0%

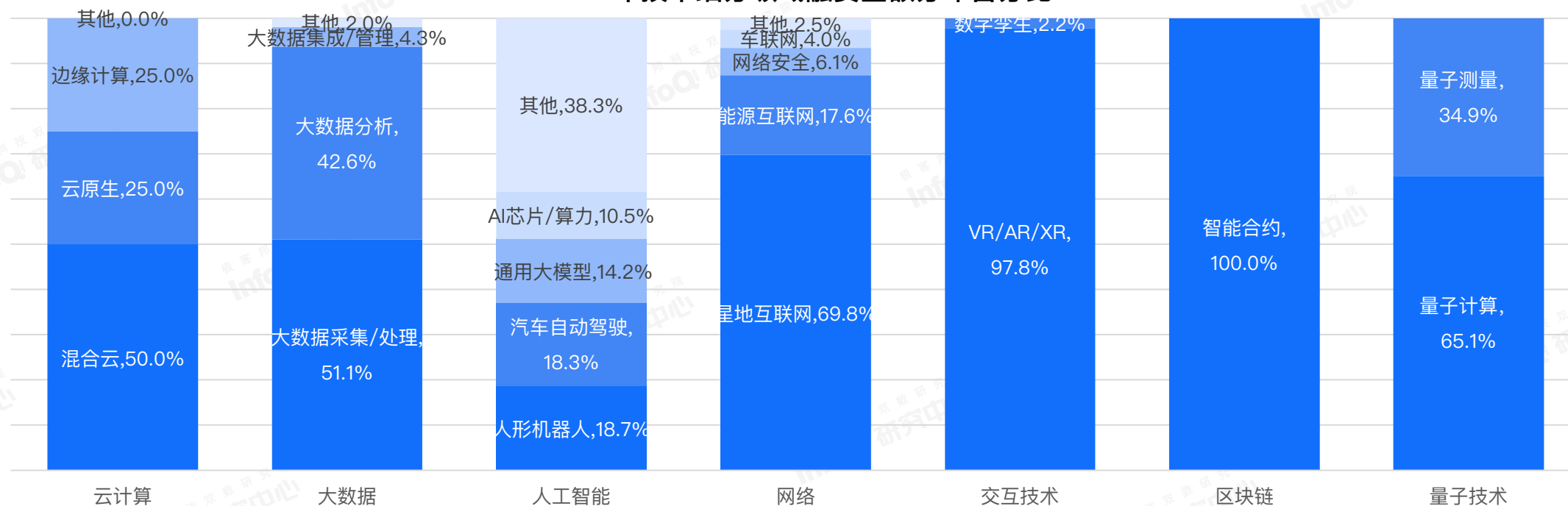
数据来源：InfoQ研究中心根据投融资新闻、资讯等公开资料整理

涉及美元融资金额统一按照2025年全年的平均汇率（6.96）进行换算，港元统一按照2025年全年的平均汇率（0.89）进行换算

## 投融资：人形机器人接棒通用大模型，吸纳技术领域融资额的16.8%

- 2025年，人形机器人成为所有技术领域吸纳资金最多的二级领域，占整体技术领域融资额的16.8%。同样受具身智能概念影响的自动化/机器人领域融资额也有所上升，占整体技术领域融资额的6.0%。
- 通用大模型与多模态模型两大相关细分领域，融资进程虽有所放缓，但仍然是人工智能领域吸金能力靠前的细分领域，合计占整体技术领域融资额的21.6%。

2025年技术细分领域融资金额分布百分比



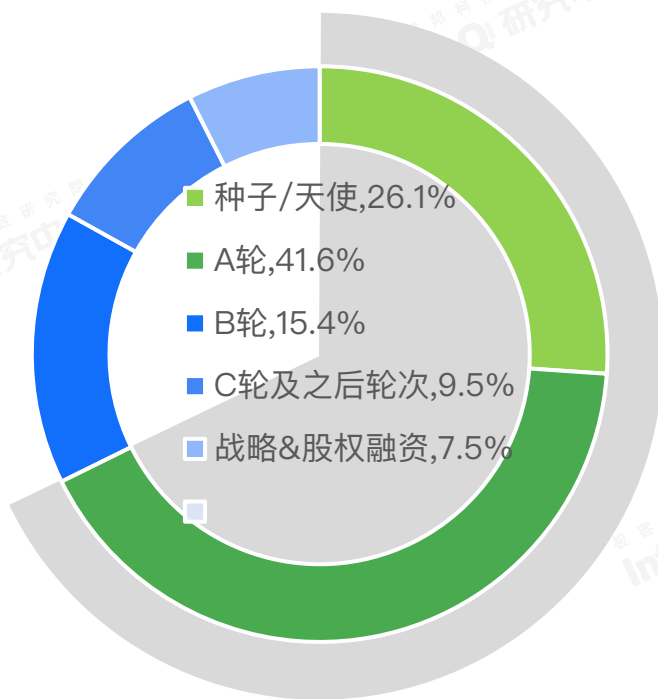
数据来源：InfoQ研究中心根据投融资新闻、资讯等公开资料整理

涉及美元融资金额统一按照2025年全年的平均汇率（6.96）进行换算，港元统一按照2025年全年的平均汇率（0.89）进行换算，最终百分比经过四舍五入取整数处理。

## 投融资：人工智能领域持续诞生新吸金领域

- 人工智能领域在融资方面表现亮眼，不仅总融资额庞大，而且大额融资事件也频频发生。据统计，人工智能领域过亿融资事件达到了212件，占整体过亿融资事件的88.7%。

2025年技术领域融资事件轮次分布情况



2024年技术领域融资金额最多的中国公司 TOP10

| 公司简称    | 成立时间    | 总部 | AI方向    | 最新融资轮次  | 年融资总额  |
|---------|---------|----|---------|---------|--------|
| 智谱AI    | 2019/6  | 北京 | 通用大模型   | 已上市     | 56亿人民币 |
| Minimax | 2021/11 | 上海 | 通用大模型   | 已上市     | 45亿人民币 |
| 月之暗面    | 2023/3  | 北京 | 通用大模型   | C轮      | 35亿人民币 |
| 清微智能    | 2023/3  | 北京 | AI芯片/算力 | C轮      | 20亿人民币 |
| 时空道宇    | 2022/3  | 浙江 | 星地互联网   | 战略投资    | 20亿人民币 |
| 星动纪元    | 2016/8  | 北京 | 人形机器人   | A+轮     | 15亿人民币 |
| 曦智科技    | 2023/5  | 上海 | AI芯片/算力 | C轮      | 15亿人民币 |
| 自变量机器人  | 2021/8  | 山东 | 人形机器人   | A+轮     | 13亿人民币 |
| 千寻智能    | 2014/3  | 浙江 | 人形机器人   | Pre-A+轮 | 11亿人民币 |
| 云深处科技   | 2018/2  | 浙江 | 人形机器人   | C轮      | 11亿人民币 |

数据来源：InfoQ研究中心根据投融资新闻、资讯等公开资料整理

涉及美元融资金额统一按照2025年全年的平均汇率（6.96）进行换算，港元统一按照2025年全年的平均汇率（0.89）进行换算，最终百分比经过四舍五入取整数处理。



# AI正在卷入一切：重构的临界点

# AI本身持续进化：推理模型诞生新技术范式

- ◆ 如同InfoQ研究中心在2025年5月发布的《推理模型综合测评报告》所言，2025年是推理模型的一年：受「推理时计算拓展」与「可验证奖励强化学习（RLVF）」两大技术范式驱动，全球厂商已进入推理模型密集发布期。从2024年Q4到2025年，大模型领域的知名厂商均发布/迭代了至少一款推理模型。在年末，我们看到了更多垂类领域推理模型的发布，例如技能、医疗、编程，同时也看到了混合思考模式已经对模型训练和产品设计带来了诸多影响。



备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合

# AI本身持续进化：多模态模型开启原生模态整合，世界模型概念探索中

- ◆ 长久以来，多模态模型的理解和生成技术发展相对独立，并形成了两种不同的架构探索路径。但2025年，我们看到了以GPT-5、Gemini3、Bagel、VEO等开始探索统一理解和生成底座的多模态模型的迅速发展。同时我们也观察到了图像、语音、文字模态的技术路线的相对成熟，视频模态模型发展仍以视频理解和视频生成相对独立的技术发展路线为主。

## 以扩散架构为主的多模态模型路线

- **技术关键：**以扩散机制为基础，通过迭代去噪过程实现高质量图像生成，同时融入多模态上下文（如文本、图像嵌入）实现理解能力。
- **优点：**擅长生成，生成图像质量高、细节丰富；可根据提示词进行多样化创作（风格、编辑等）；训练过程相对稳定。
- **缺点：**推理速度慢、部分模型训练时学习信号稀疏，对不同长度的输出适配不好；部分模型依赖外部框架，开源支持有限。
- **2025年发布的代表模型：**MMaDA（PU&PKU、2025-05）、FUDOKI（HKU&华为，2025-05）、Muddit（PKU&中国典型&NUS&PU、2025-05）

## 以自回归架构为主的多模态模型路线

- **技术关键：**基于 LLM 的自回归架构，将图像转化为序列Token，通过预测下一个Token的目标统一建模文本与视觉模态。
- **优点：**和大语言模型结构相通，能灵活进行跨模态推理；支持图文交错生成（比如边写文字边插图片）；部分模型用连续令牌，不会丢失图像原始信息。
- **缺点：**模态对齐依赖令牌器质量，设计难度高。
- **2025年发布的代表模型：**TokLIP（腾讯 ARC Lab&中科院自动化所等、2025-05）、Selftok（华为、2025-05）、UniTok（字节&CUHK、2025-02）、UniFork（上海AI实验室、2025-06）、OmniGen2（北京人工智能研究院、2025-06）、Qwen-Image（阿里、2025-08）、Ming-Omni（蚂蚁、2025-06）、Skywork UniPic（昆仑万维、2025-08）

## 扩散与自回归混合统一的多模态模型路线

- **技术关键：**融合自回归的序列推理优势与扩散的视觉生成优势，文本令牌自回归生成，图像令牌多步去噪生成，通过双向注意力或共享骨干网络实现跨模态融合，平衡文本语义控制与图像视觉保真度。
- **优点：**兼顾文本语义可控性与图像生成质量；支持复杂任务
- **缺点：**架构复杂，训练与推理成本高；模态融合难度大，易出现文本 - 图像对齐偏差；
- **2025年发布的代表模型：**Mogao（字节、2025-05）、Bagel（字节、2025-05）

备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合

# AI Infra竞争升维：围绕提高算力利用率的推理优化成为竞争力的重要一环

- ◆ 2025年初，围绕DeepSeek的火热讨论，除了关于开源、纯强化学习路线和训练开销外，其一系列DeepEP、DeepGEMM、FlashMLA、EPLB 等关键组件的开源，让大家的目光将转向了推理优化。
- ◆ 另一方面，Token消耗量以比大家预期要快的多的速度迅速增加，这进一步驱动了大家围绕算力利用率而展开的Infra讨论。
- ◆ 技术圈逐渐意识到，沿用传统训练架构来承载推理任务，是无法让算力、显存等AI infra资源同时达到最优配置的，于是在2025内，我们看到了围绕提高算力利用率而展开的一系列探索成果。从训练优先到训推一体并最终转向推理专用的架构解耦与优化，构成了2025年大模型基础设施演进的一条主线。



备注：InfoQ研究中心根据官网信息、媒体公开报道、相关论文和技术报告信息整合



# MCP、A2A、Skills等生态标准逐渐加速Agent发展

◆ 虽然MCP面临诸多挑战，但在2025年中，其仍是共识度最高的协议。

## 2024年

### MCP正式发布

Anthropic发布MCP开放生态协议，为大模型应用与外部数据源和工具之间的集成提供一种标准，同步发布Claude Desktop客户端

## 2025年3-5月

### 国际巨头宣布支持

3月，OpenAI Agent SDK正式支持MCP服务  
4月，Google宣布将为Gemini模型及其SDK新增对MCP的支持

### 国内大模型平台同步跟进

3月，百度千帆大模型平台宣布接入集成MCP，支持通过AppBuilder SDK开发的组件无缝转化为MCP Server  
4月，阿里云百炼上线首个“全生命周期MCP服务”，接入高德、无影等产品，支持5分钟完成工具链集成  
4月，腾讯云发布AI开发套件，支持MCP插件托管服务  
5月，华为云ModelArts Studio支持MCP服务  
5月，火山引擎发布大模型生态广场 MCP Servers

## 2025年

### 行业应用厂商持续贡献MCP服务

编程工具：Cursor在其0.45.6版本中加入一键MCP功能，Windsurf发布支持MCP的Wave 3，VS Code 1.101集成MCP协议，JetBrains IDEs在2025.2版本中支持MCP  
支付宝、百度地图、Bilibili、Cloudflare、ElevenLabs、Minimax、Notion、Shopify、Unity、WPS等软件先后宣布发布官方MCP服务  
2026年1月，MCPServers已收录63410项MCP服务，MCP.so已收录14456项MCP服务，Aibase已收录MCP服务13736项

### MCP协议持续更新

伴随着MCP真正进入企业环境，MCP在安全、上下文膨胀等方面面临挑战，MCP进行了多轮关键更新  
围绕授权体系的安全更新：基于DNS和HTTPS信任体系的白名单机制、最小特权与渐进式授权、工具命名攻击的防御机制、URL 模式形成的物理隔离、本地安全防护等  
上下文膨胀：ToolSearch、Skills等解决机制

2025年12月，Anthropic宣布将MCP正式捐赠给Linux基金会旗下新成立的Agentic AI Foundation (AAIF)。

备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合

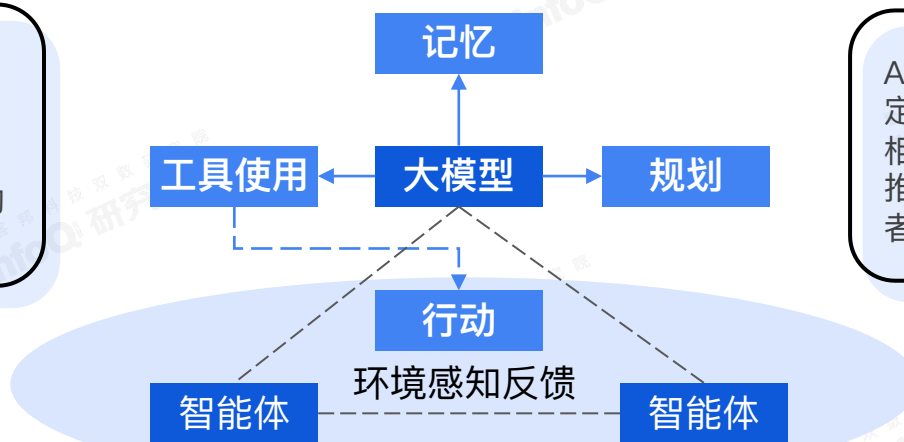
# AI Agent：仍不足以长出自主的头脑和手脚，但在持续进化中.....

- ◆ AI Agent这个概念，从Devin后，2025年3月，一款叫Manus的产品，又一次掀起了技术圈对Agent产品的讨论热潮。
- ◆ 在InfoQ研究中心发布《2024 中国AI Agent应用研究报告》中提到，现阶段的智能体应用，仍未长出足够的自主思考能力以应对复杂场景。时过1年半，这句判断依然成立，但记忆、推理、工具使用都有了一些新的新的工程化的进展，以及也有模型开始探索从模型架构重构思考和记忆的方式，例如DeepSeek提出的Engram架构）。
- ◆ 技术以外，我们也看到了DeepResearch、Computer Use等特定领域的智能体应用形态的出现。

- MCP协议：Anthropic提出的让基于LLM的智能体能通过标准化接口，安全、一致地与外部服务和数据源交互。
- Skills：应对MCP带来的上下文爆炸的一种解决方案

- 外部记忆机制：Manus、OWL、OpenManus等利用外部文件系统存储中间结果与历史数据
- 持续学习（Continuous Learning）：核心目标是让模型在任务序列中不断吸收新知识，同时维持甚至提升旧任务的表现

Agentic Workflow：通过预先定义的代码路径来编排大模型和相应的工具，是现阶段面对模型推理能力不足下的一种妥协，或者说，是一种过渡形态。



- A2A协议：谷歌提出的面向多智能体之间的通信协议

备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合



# Coding领域：Copilot与Agent两脚走路，氛围编程贯穿全年

• 编码领域一直是大模型以及Agent领域最核心、也最先发生质变的探索阵地。

• 在2025年，在代码补全能力和体验持续提升的同时，我们看到了直接接管开发任务、自主规划调试和结果验证的Coding Agent的崛起。

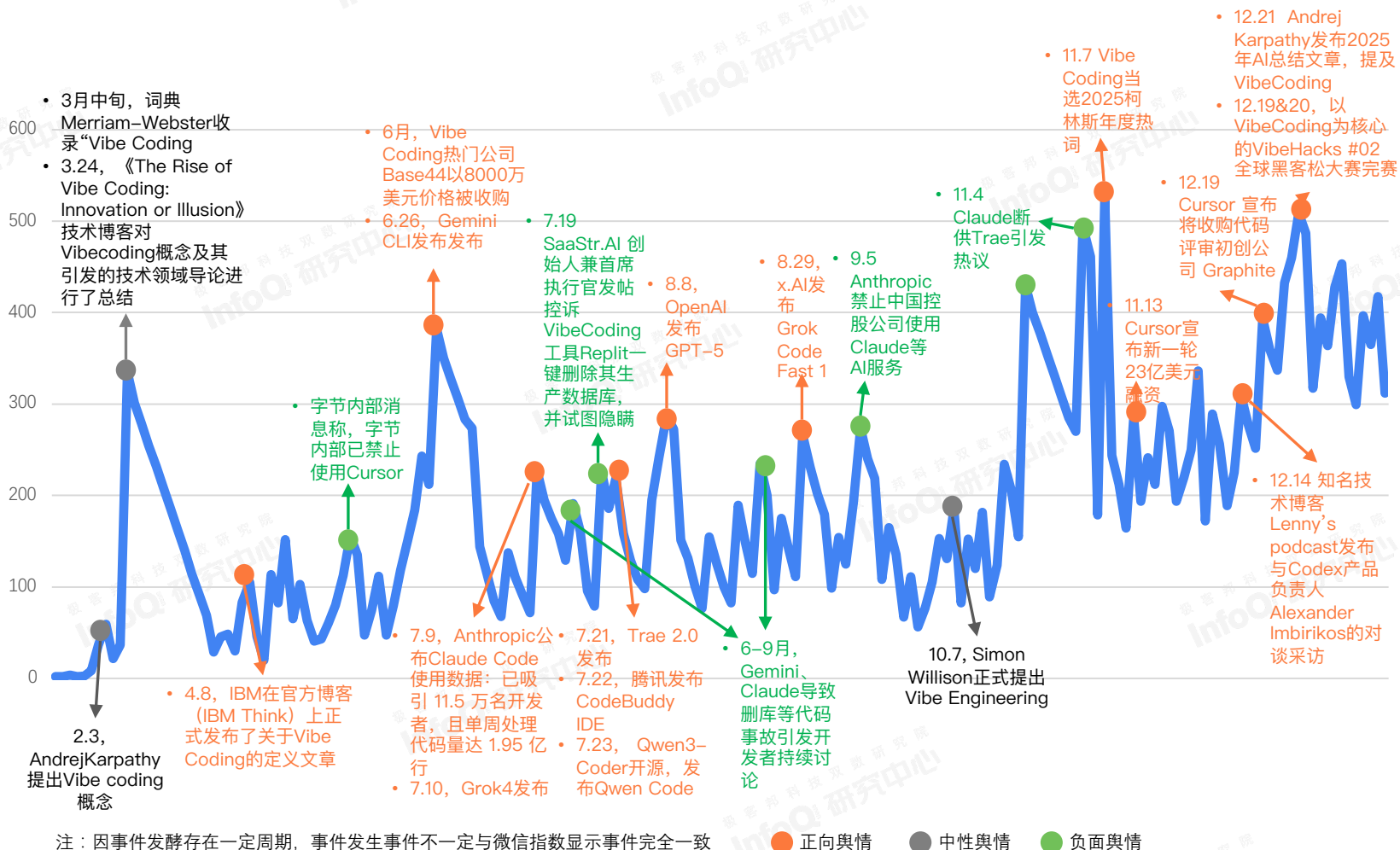
• 以Claude Code、Codex、Gemini CLI、Qwen Code为代表的Coding Agent，正在把代码能力从模型能力，推向可持续工作的工程能力。

• 在能力提升的同时，因Token消耗量激增而带来的成本与产品定价，尤其以Cursor一年内连续多轮的涨价与功能削减，也引发了开发者的广泛讨论。

• Vibe Coding，成为贯穿 2025 年代码领域的关键词。一方面，部分开发者开始积极分享基于Vibe Coding的实践成果，展示个人或小团队在极短时间内完成原型构建、产品上线的案例，也催生了Loveable、Replit等新产品、新公司，推动人人可开发的边界进一步前移。

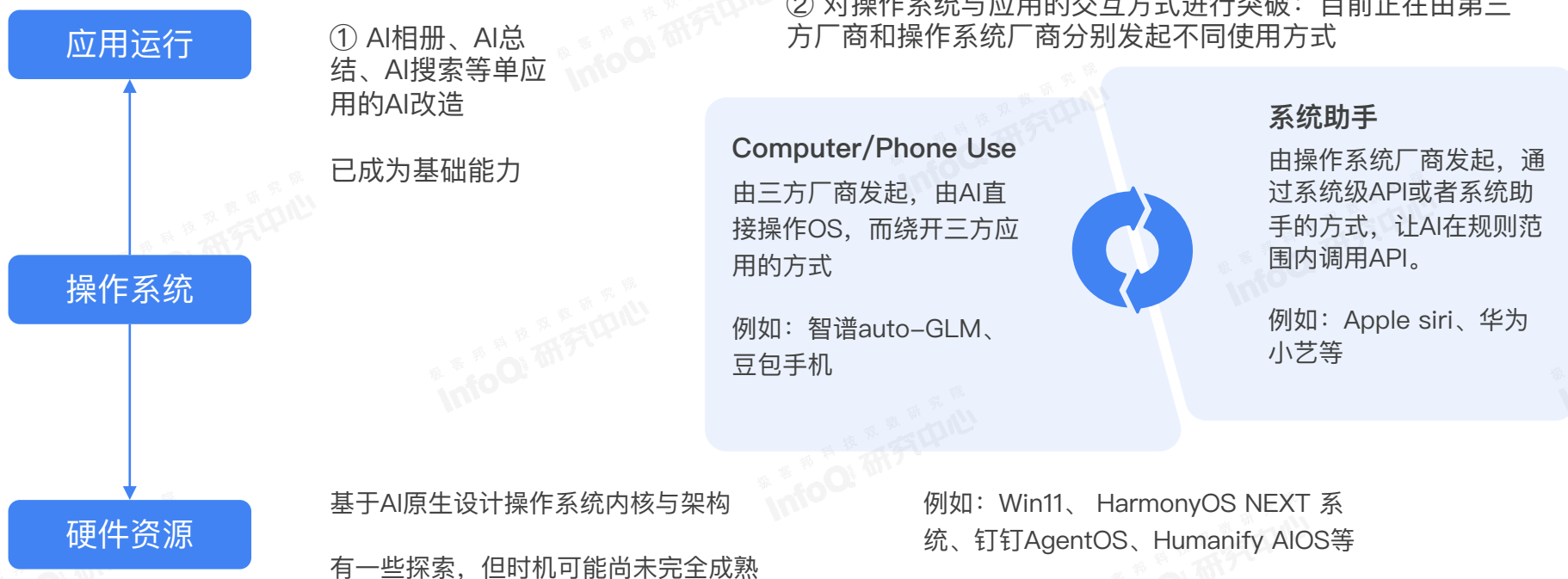
• 但另一方面，随着Vibe Coding从实验走向真实生产环境，其带来的安全、稳定性与运维复杂性问题也开始集中暴露：代码可控性下降、依赖链不透明、隐性缺陷难以审计，以及对长期维护和责任归属的挑战，成为企业级应用无法回避的新课题。

2025年Vibe Coding领域热点事件舆情分析图



# AI OS：应用层探索更近一步，内核与架构AI原生设计为时尚早

- ◆ InfoQ研究中心在2年前的年度报告中，提及手机和电脑厂商探索将大模型能力搭载在端侧。彼时的探索更多在总结摘要和图像处理上，可视为AI向终端渗透的初步尝试。两年过去，这种针对单应用的AI赋能已演进为基础能力。行业焦点正从单点功能的实现，转向对操作系统底层的深度重构，即AIOS的探索迈出了实质性一步。
- ◆ 在2025年我们看到了autoGLM和三星的合作、看见了豆包手机的发布。同时，我们也看见了App厂商、硬件终端厂商、大模型技术公司之间，围绕生态主导权与用户入口的博弈。



备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合



## 生态开放度：中国拥抱开源，海外坚守闭源

- ◆ 2025年，中国市场中，开源队伍逐渐壮大，除了一直以来坚持开源路线的阿里Qwen、DeepSeek等，百度、Kimi、Minimax、美团等均经历了从闭源为主转向开源的过程。
- ◆ 国际层面则呈现了相反的局面，以Llama系列开源模型的Meta，在Llama4之后宣布后续模型不默认开源，而国际主力军OpenAI、Anthropic、Google则仍然坚守闭源路线。



备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合



# 面向未来的探索：正在迈向现实



# 星地互联网：加速组网、设备普及和新标准新技术融合成果

- ◆ 中国在5G网络发展中所形成的建用并举、市场驱动、标准引领的中国特色路径，正在星地互联网领域开启新的篇章。2025年，我国星地互联网发展实现系统性突破：星座建设迈入规模化部署新阶段，直连卫星终端生态初步形成并在消费端快速普及，同时在6G星地融合国际标准制定中取得主导权，并于星地量子通信等前沿技术领域实现全球引领。

## 卫星规模化组网与发射里程碑

- **中国星网（GW星座）**：累计发射17组卫星，合计125颗，其中2025年份发射量占82.4%
- **千帆星座（G60）**：累计发射6组卫星，合计108颗，其中2025年份发射量占50.0%

## 卫星通信终端初局，普及应用新台阶

- **市场准入**：2025年9月，工业和信息化部向中国联通、中国移动颁发卫星移动通信业务经营许可，加之此前已获许可的中国电信，三大基础电信企业全面进入卫星通信市场。
- **车机突破**：华为发布车载卫星通信系统——星河，推动卫星能力上车。
- **手机普及**：中国电信联合华为、荣耀、小米、OPPO、vivo、中兴、三星等品牌厂商已推出约40款直连卫星手机，相关终端累计出货超2500万台；

## 星地融合新技术标准与量子通信技术突破

- **通信标准新突破**：2025年3月关于6G的标准研讨会上，全球卫星运营商协会等50多家公司支持地面移动通信和卫星通信标准统一设计，卫星通信成为6G标准第1版本的重要组成部分。
- **中国团队牵头立项通信国际标准**：由中国信科、大唐移动、信通院、宸芯科技、北邮组成的中国团队代表中国首次在3GPP实现5G NTN标准牵头立项，累计牵头22项立项，占比1/3，是全球立项最多的单位，全面提升了我国在卫星通信国际标准的话语权和竞争力。
- **星地量子通信**：3月，中科大基于“济南一号”构建卫星组网实现跨越12900公里的星地量子密钥分发，支持北京—南非实时加密通信，这是中国首次实现上万公里星地量子通信，为实用化卫星量子通信组网铺平了道路

备注：InfoQ研究中心根据官网信息、媒体公开报道信息整合，相关数据截至2025年12月



# 卫星向通信+计算转变，太空算力布局缓解AI算力压力



## 首个太空计算卫星成功发射

2025年5月，国星宇航与之江实验室成功发射“三体计算星座”首批12颗计算卫星。该星座搭载80亿参数的天基模型，可对L0-L4级卫星数据进行在轨处理。首发单星最高算力744 TOPS，星座总算力5 POPS，总存储容量30TB，星间激光通信速率最大100Gbps。

太空计算卫星

算力、模型上天

天算星座

空天计算

## 星算计划全面启动

北邮牵头，联合长沙天仪研究院有限公司发起的“天算星座”计划，定位为开放开源的空天计算在轨试验平台。2025年，作为二期计划的首批卫星北邮二号、三号卫星顺利进入预定轨道。



## 太空数据中心建设工作推进会成功召开

2025年11月，北京市科学技术委员会等组织召开太空数据中心建设工作推进会，宣布拟在700-800公里晨昏轨道建设运营超过千兆瓦功率的集中式大型数据中心系统，以实现将大规模AI算力搬上太空。数据中心系统由空间算力、中继传输和地面管控分系统组成，空间算力计划部署多座太空数据中心，每座功率约1GW，可容纳百万卡级别的服务器集群，开展天基数据中继传输和计算服务。

太空数据中心

空间算力



# 量子技术多点突破，产业链协同发展，国际竞争优势明显

## 量子计算

基于量子比特，利用量子叠加等构建新型指数级加速并行计算的计算体系，以推动算力跨越式发展。

### 2025年多项技术路线获得突破，整体位于世界前列水平

- **超导路线**：3月，中国科学技术大学联合国内多家机构研制的超导量子计算原型机“祖冲之三号”正式发布，创造超导体系量子计算优越性的新纪录。
- **离子阱路线**：4月，华翊量子发布了百比特离子阱量子计算原型机；7月，合肥么正量子成功完成国内首台4K低温QCCD芯片型离子阱量子计算系统的组装调试。
- **中性原子路线**：8月，中国科大中性原子量子计算构建了2024个原子的无缺陷二维/三维阵列，单比特门保真度达99.97%，刷新世界纪录。
- **光子路线**：2月，中国“九章三号”光子量子计算机实现商用，特定场景运算效率超传统超算千亿倍；7月，玻色量子展出研发中的1000量子比特的相干光子量子计算机真机模型。
- **量子纠错**：12月，中国科学技术大学基于“祖冲之3.2号”，提出并实现了全新的“全微波量子态泄漏抑制架构”，实现了码距为7的表面码逻辑比特，逻辑错误率随码距增加显著下降

## 下游应用化学、药物筛选先行

- **生物医药筛选**：深圳湾实验室量子制药工厂宣布建成，其是全球首个基于量子计算的药物分子筛选中心，实现高精度量子化学模拟加速新药研发。广州量子计算+生物医药创新研究战略合作正式发布，合作方包括量子计算领军企业玻色量子、中山大学附属第三医院、广州市红十字会医院、AI生命科学平台百图生科、广州移动。

## 量子通信

利用量子比特作为信息载体来进行信息交互的通信技术，主要探索方向有物理技术的量子保密通信和数学技术的抗量子密码。

### 独创量子安全直接通信，积极探索QKD与PDC技术

- **量子密钥分发（QKD）**：3月，中科大基于“济南一号”构建卫星组网实现跨越12900公里的星地量子密钥分发，支持北京—南非实时加密通信；4月，北京量子院联合清华大学、中山大学在薄膜铌酸锂基光子集成平台上成功构建QKD系统，并在25公里传输距离下，达到了11Mbps的安全成码率；12月，北京量子院公布一项高性能实用化QKD系统研究成果，实现了2dB信道损耗下15.2 Mb/s的高安全密钥率（SKR），量子误码率（QBER）低至1.2%–1.3%，刷新了系统性能纪录。
- **量子安全直接通信**：2月，北京量子院与清华大学、北方工业大学，联合提出单项量子直接通信理论，并创造2.38kps@104.8km@168小时的长距离稳定传输世界记录。
- **抗量子密码（PDC）**：西交利物浦大学数学物理学院院长丁津泰领衔的研究团队在国际公开 Darmstadt 格最短向量（SVP）挑战赛中破译200维难题，刷新全球纪录。
- **QKD+PQC融合路线**：4月，中电信量子推出QKD+PQC融合分布式密码系统，并完成实战验证，成功实现跨境量子密信通话。

## 量子城域网、金融专网成果斐然

- **金融量子安全**：中行、四川农商行、华夏银行等先后宣布北京、四川、深圳、广西开通量子加密金融网络/专线。
- **量子城域网**：目前合肥、上海、北京、广州等重点城市已开始建设量子城域网。

## 量子测量

利用量子特性实现对物理量的超高精度测量，以显著突破传统测量仪器的精度限制

### 精密测量精度与稳定性持续突破

- 北京大学研制皮米级分辨率位移测量迈克尔逊激光器，较国际最优迈克尔逊激光干涉仪指标提升1个量级，
- 7月，中国科学院国家授时中心在87Sr光晶格原子钟系统领域取得了突破进展，使我国在光晶格钟主要性能上达到世界领先水平，成为第二个实现光晶格钟频率不确定度和频率不稳定性均优于 $2 \times 10^{-18}$ （相当于连续运行160亿年误差一秒）的国家。
- 7月，中科大团队高精度原子重力仪山区连续稳定运行150天，证实了原子重力仪在地震研究领域的应用价值
- 8月，由北京航空航天大学开发的全新原子自旋传感器成功实现对弱磁场的超高灵敏度与可溯源精准测量

## 电力、矿产勘探已形成一定应用规模

- **电力能源**：1月，由南方电网牵头，中国科学技术大学、中国电气装备西安西电高压开关有限责任公司等单位联合研制的全球首套 $\pm 800$ 千伏特高压直流量子电流传感器成功落地，标志着量子测量技术在电力系统实现应用。
- **矿产勘探**：湖南量子测量产业创新联合体成立，由长沙量子测量产业技术研究院牵头，联合中南大学、湖南师范大学、中电长城科技、湖南省地震局、湖南省地质调查所、达晨财智等18家单位共同发起。
- **医疗成像**：GE医疗Expert X 新量子CT亮相，验证量子技术（如Apex量子平台）是支撑高精度成像的关键



# 量子技术与AI协同创新，酝酿量子人工智能（QAI）新学科

- ◆ 而在量子技术2025年发展的背后，也离不开AI的身影，AI for Quantum和Quantum for AI都在迅速发展，同时QAI新交叉学科也正在酝酿发展中。

## 人工智能 For 量子技术

- 量子计算芯片设计与量子纠错算法设计
- 量子系统优化、设备检测与校准、数据后处理精度提升

- 量子通信系统模拟与训练、协议设计、参数预测、系统校准
- 量子通信网络的攻击智能检测和资源分配

- 量子测量系统的参数优化与异常检测、实时优化和设备校准
- 后数据处理与智能分析

2025年8月，中科大潘建伟团队利用AI技术，在60毫秒内成功构建出多达2024个原子的无缺陷二维和三维原子阵列，创造了当时中性原子体系无缺陷原子阵列规模的世界纪录。此外，该系统单比特门保真度达99.97%、双比特门保真度达99.5%、探测保真度达99.92%，关键指标达到国际领先水平，并入选《物理》杂志2025年国际物理学领域九项重大进展

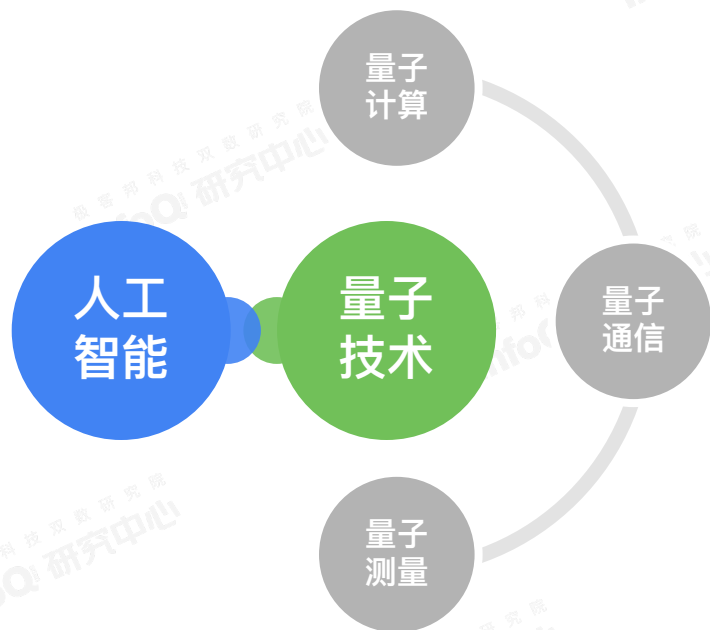
## 量子技术 For 人工智能

- 量子计算形成的并行计算优势，满足人工智能算力需求，加速机器学习训练过程，降低计算成本，增强模型表达能力例如QML（量子机器学习）

- 量子通信，为人工智能的数据安全与隐私安全以及云边端的传输安全提供强有力的保障

- 量子测量作为AI的新型数据采集与特征提取机制，以提供较传统传感器数据精度更高的量子态数据，以探索之前无法捕捉的微观特征

2025年1月，英矽智能与多伦多大学研究团队共同发布一项将量子计算模型与生成式人工智能相结合的药物发现成果，提出了一种由量子变分生成模型和长短期记忆网络相结合的量子-经典混合框架模型，将量子计算与经典计算方法结合起来设计新分子，并发现了靶向“不可成药”的癌症驱动蛋白KRAS的新颖分子。



# 低空飞行产品原型频繁发布，但面临技术、基建、盈利模式的制约

## 低空空域管理规范出台，提供有序管理和发展的基石

- 《无人驾驶航空器飞行管理暂行条例》在2025年全面施行
- 12月，新修订的《民用航空法》通过，明确提出优化低空空域资源配置，推动建设民用低空飞行服务与监管平台，建立适应低空经济发展的适航审定和飞行管理制度。
- 12月，中央空管办发布《国家级和省、市级低空飞行综合监管服务平台功能要求（1.0版）》和《国家级和省、市级低空飞行综合监管服务平台信息交互规范（1.0版）》。

## eVTOL新品频繁发布，适航审定取得突破

- 峰飞航空2吨级货运 eVTOL “凯瑞鸥”率先集齐型号合格证、生产许可证、单机适航证三证。
- 亿航智能EH216-S集齐四证，成为国内首个具备商业化载人运营资质的 eVTOL 机型，为eVTOL的商业化运营打开序幕。
- 以小鹏汇天、沃飞长空、沃兰特等为代表的企业，在2025年发布了多旋翼、倾转旋翼、复合翼等多种技术路线并存的产品。

## 人才培养起步期，高校和职业教育双步走

- 高等教育：2025年，教育部更新发布了《普通高等学校本科专业目录（2025年）》，增列29种新专业，纳入2025年高考招生。其中包括增设的航空运动等专业；瞄准低空经济快速发展需要，突破集中申报限制，指导北京航空航天大学等6所高校增设低空技术与工程专业。
- 职业教育：全国职业院校低空经济技能竞赛成功举办，职业院校增设6个相关专业，构建多层次人才培养体系。

## 低空飞行的2025 有彩虹也有风雨

## 动力系统、自主导航、低空通信等技术制约低空飞行器的应用场景拓展

- 电池能量密度、循环寿命、充放电倍率限制动力系统续航与运营能力
- 现有自主导航系统无法满足雨雾、雨林等复杂环境以及多机编队协同的技术需要
- 5G、卫星通信等对低空空域覆盖率不足，频谱和无线电频率等资源的碎片化影响低空飞行行业的实际运营和管理能力

## 低空飞行相关基础设施建设尚未完善

- 起降场地和停机坪、充（换）电等能源设施等硬基建仍需持续建设，监管和运营维护缺乏统一标准
- 城市空管信息系统、空域管理辅助系统、飞行服务站系统等软基建数据分属不同部门，统一服务平台建设刚刚起步

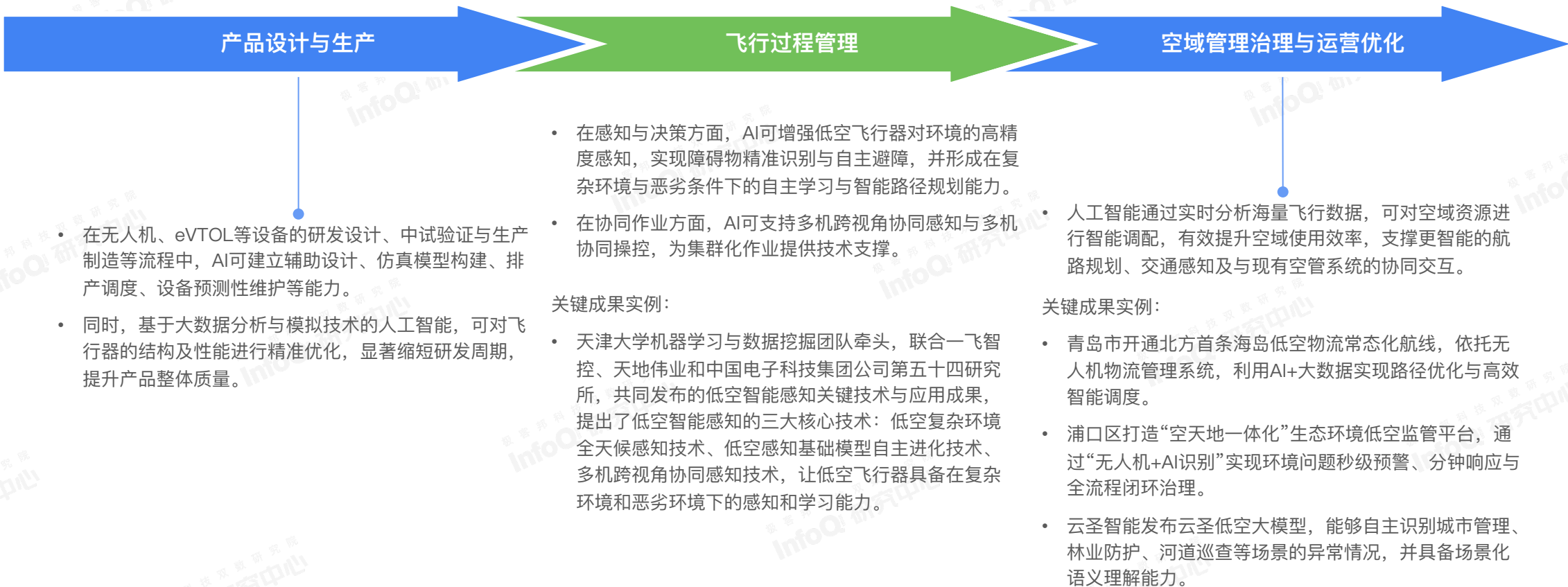
## 应用场景与配套产业链需要协同发展，以拓展运营盈利模式

- 从技术原型到盈利运营，除了产品设计生产外，应用场景的探索，包括保险、金融、数据服务等配套产业的不成熟，均推高了低空飞行整体行业的盈利发展



# AI助力低空飞行：智能感知与协同决策已成为技术落地焦点

- ◆ 2025年低空飞行产业发展中，AI也在持续发力。例如目前在低空障碍物识别与主动避障中发挥重要作用的智能感知与协同决策。除此之外，智能空域管理和运营调度，也是现阶段探索技术结合的重点。



# AI for Science：开启新科研范式探索，筹备系统化转变

- ◆ 8月，国务院印发《关于深入实施“人工智能+”行动的意见》，提出加快实施“人工智能+”科学技术行动，加快科学发现进程，率先建立基于AI的新型科研和研发范式，推动我国科技创新走在世界前列。
- ◆ 年内，一系列科学基础大模型的集中发布，标志着“AI for Science”（AI4S）正在完成从关键能力到系统化基础设施的关键部署，为我国加速科学发现、引领科技创新提供了核心支撑。

## 集成电路设计

中国科学院计算技术研究所与软件研究所联合研发的启蒙2号系统，实现了全自动设计的RISC-V CPU芯片，规模达1700万个逻辑门，性能对标ARM Cortex-A53。

该系统能够自动配置操作系统、编译程序及高性能算子库，完成从芯片硬件到系统软件的全链条自动化设计，其设计效率与成果质量已超越人类专家水平。

## 新材料研发

方大炭素与晶泰科技达成合作，聚焦炭素新材料的智能化研发。

依托晶泰科技的AI算法、机器人自动化实验平台与方大炭素的产业化经验，双方共同推进高性能炭基材料的快速筛选、工艺优化与智能制造升级，显著提升研发效率。

## 药物筛选

中国创新中心推出AI驱动的药物虚拟筛选平台DrugCLIP，首次完成人类基因组规模的药物筛选。

该平台突破传统药物研发的周期与成本瓶颈，后续将通过与瑞金长三角健康研究院、Enamine中国等机构合作，推动小分子药物智能筛选技术的产业化应用。

## 疾病预测

华大生命科学研究院与江之实验室联合发布百亿参数人类基因组通用基础模型Genos。

该模型支持百万碱基对级别的超长上下文分析，具备单碱基分辨率的精准识别能力，为遗传病研究、癌症早筛等提供底层工具支持。

## 航天发动机模拟

DeepFlame Rocket平台实现火箭发动机亿级网格高精度仿真，以“朱雀2号”为例，完成从燃料喷注到羽流场的全流程模拟。

基于AI的数值模拟方法在保持工业级精度的同时，较传统方案加速超千倍，为推进系统优化提供新范式。

## 有机液体预测

中石化石油化工科学研究院基于Uni-Mol分子表示学习模型，实现对有机液体电学、力学、热学性质的高效预测。

该模型从600余万种化合物中筛选出400余种高性能浸没式冷却液候选材料，并完成4种体系的实验验证，大幅缩短研发周期



# 2026年技术趋势展望

# 2026年十大技术趋势

01

收敛已久的Transformer架构，即将迎来分化与创新新阶段

02

RLVR范式应用扩展与持续演进，经验学习等新范式正在路上

03

原生多模态成为默认能力，原生全模态加速成型，世界模型技术路线迎来首轮技术收敛周期

04

AI推理基础设施凸显战略价值，系统化工程决定长期竞争力

05

Agent迈向结果交付，Agent Infra从算力基础演进为风险可控、可验证、可托付的业务级支撑

06

C端应用，记忆机制与生态整合成为核心壁垒

07

AI硬件持续在垂类场景破局，手机仍是核心管理与交互中心

08

有竞争就有动力，中国继续以开源撬动世界影响力

09

AI for Science推动科研生态升级，科学伦理面临深刻变革

10

前沿技术交融，智能协作开启新格局，系统级能力强化科技与战略话语权

## 趋势一：收敛已久的Transformer架构，即将迎来分化与创新新阶段

- ◆ 在大模型研发最核心的架构层，曾长期高度收敛的技术路径，正在出现清晰的分化与重组迹象。自2021年起，围绕非Transformer架构的持续探索不断积累有效成果，从状态空间模型到递归与条件计算，这些方向逐步走出单点验证阶段，并进一步催生出多种混合架构的演进形态。
- ◆ 在当下，Transformer仍稳居通用模型的事实标准，其工程成熟度和生态优势短期内难以被取代。或许我们即将看到下一代模型架构的共识了。

### 对Transformer架构的改良路线

通过 ① 改良注意力机制：引入低秩或核近似、稀疏化连接方式，  
② 对位置编码、残差结构和前馈网络的重构，让模型在保持表达能力的同时显著降低了计算与存储成本。

代表：Longformer、Performer

与现有生态高度兼容

精度回退、调参成本上升

### 非Transformer的创新架构路线

通过引入状态空间模型（SSM）、长卷积（LCM）、递归计算或频域建模等机制，重构序列建模的核心算子，用显式状态更新或线性复杂度计算替代全局自注意力，以提升长序列建模能力和整体计算效率。

代表：Mamba、RWKV

长序列友好，算力效率高

定制算子等工程门槛高

### 混合架构路线

将 Transformer与CNN、RNN、SSM、MoE等结构进行组合，通过模块化和分层设计发挥不同算子的互补优势，在表达能力、效率和扩展性之间寻求更优平衡。

代表：Hyena、MOR

在特定任务和多模态场景中更具灵活性

系统复杂性明显提升

## 趋势二：RLVR范式应用扩展与持续演进，经验学习等新范式正在路上

- ◆ 如果说2025年以DeepSeek为代表的模型成功验证了RLVR新范式，那么2026年就是RLVR范式走向扩展与深化的一年。
- ◆ 除此之外，经验学习、持续学习等新范式也将在2026年获得进一步探索。

### 01.

#### 夯实RLVR拓展至更多模型的范式验证

RLVR 正在加速扩展至多模态模型、机器人模型和世界模型等更复杂系统之中。

### 02.

#### 围绕RLVR的改进

训练稳定性、奖励设计、样本效率和可扩展性成为新的关注重点，更多工作开始尝试将过程级信号、结构约束和自动评估机制引入RLVR体系，

### 03.

#### 新训练范式的探索

经验学习、持续学习、模仿学习等方向逐渐受到关注，核心目标是让模型能够在使用过程中积累经验、调整行为，并在长期运行中保持能力演进。这类探索尚处早期，但已经显露出对静态大模型假设的突破潜力。



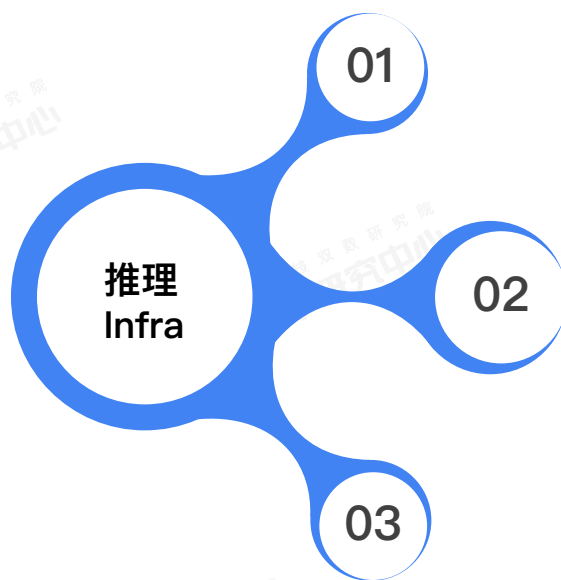
## 趋势三：原生多模态成为默认能力，原生全模态加速成型，世界模型技术路线迎来首轮技术收敛周期

- ◆ 2026年，原生多模态能力成为AI的标配，原生全模态模型加速落地，多模态理解与生成逐步融合。
- ◆ 世界模型技术路线迎来首轮收敛，跨模态统一底座开始形成，为具身智能、自动驾驶等应用的认知、推理与预测提供系统化基础。



## 趋势四：AI推理基础设施凸显战略价值，系统化工程决定长期竞争力

- ◆ 2025年开始，AI基础设施的重心发生了一个很关键的迁移。应用数量的增长和场景复杂度的提升叠加在一起，让 Token 消耗呈现出双重放大效应。一方面是总请求量暴涨，另一方面是单次请求变得越来越长、越来越复杂。这意味着，算力瓶颈开始更多出现在推理侧，而不是训练侧。这不是一个短期波动，而是应用进入真实生产环境之后的必然结果。
- ◆ 这也将大家过往对训练集群的关注重点，转向了推理相关的系统架构，围绕推理芯片适配 — 算力调度 — 推理优化 — 推理加速 — 系统稳定性等系统化工程。追求训练、推理在流程、团队上共同合作的训推一体化趋势也是推理重要性提升的重要表现。
- ◆ 对于2026年，推理优化趋势仍将持续，面向推理的基础设施是推理优化效果的重要保证，也是通过面向推理的专用基础设施，持续放大算力效率差异，并最终转化为平台和生态的长期优势。



### 系统稳定性和服务能力

- 吞吐、延迟、SLA
- 稳定、易用、成本可控的模型API

### 推理优化与加速

- KV Cache、并行策略、算子优化、PD分离、AF分离等

### 芯片与算力

- 异构芯片适配
- 算力资源管理与调度等



## 趋势五：Agent迈向结果交付，Agent Infra从算力基础演进为风险可控、可验证、可托付的业务级支撑

- ◆ 2025年，Coding Agent已率先迈入面向结果的生产力阶段，而2026年，其他业务领域也将加速跟进。Agent记忆模块取得核心突破，使智能体能够更好地积累、调用和复用经验，从而提升跨任务的持续执行能力。Agent Infra承担风险兜底、过程治理与规模化支撑，为Agent自主运行划定安全边界，使行为可观测、过程可控、结果可验证。MCP等生态协议加速多工具、多平台协同，形成统一标准与接口，为智能体在真实业务中稳定运行并持续交付结果提供基础保障。





## 趋势六：C端应用，记忆机制与生态整合成为核心壁垒

- ◆ 2026年，C端应用可能将仍然以大模型即产品为主形态，用户感知到的核心价值依然集中在对话、生成和陪伴感上，入口高度集中，产品形态看起来相似度很高。在这个前提下，真正拉开差距的点，已经从模型参数规模，切换到了长期使用过程中能不能记住用户、理解用户、持续适配用户。即模型能力决定产品能不能被试用，记忆能力决定用户会不会被留下。
- ◆ 除此之外，各大厂商自研的C端应用将进一步尝试将自家产品生态进行深度整合，最大程度探索生态内的产品融合，以进一步巩固自己在互联网网领域的护城河。



记忆机制正在成为 C 端产品最现实的壁垒来源，因为它直接影响用户迁移成本。一旦模型能够稳定记住用户的偏好、语气、长期目标和历史决策习惯，体验就会从一次次重新教模型认识我，变成它天然就懂我在说什么。这种连续性体验是单次能力堆参数很难替代的。



更重要的是，记忆并不仅仅是技术能力，也绑定了产品层面的设计选择，包括哪些信息值得被记住，如何在不打扰用户的情况下持续更新，如何在隐私、安全和可控性之间取得平衡。这些一旦形成体系，就会沉淀为产品级能力，而不是简单的模型功能。



## 趋势七：AI硬件持续在垂类场景破局，手机仍是核心管理与交互中心

- ◆ 如果说趋势六中的AI应用仍然以通用AI为主，那么AI硬件则是通过垂类场景（健身、教育、睡眠等）作为破局核心，但大多数AI硬件，仍将与手机的连接作为设置和管理的首要入口。手势操控方式仍需要一定的推广和用户习惯教育周期。

1

**AI硬件持续垂类场景破局思路，以建立差异化竞争力**

厂商开始以垂直场景寻找突破口，例如健身、教育、睡眠、会议等，通过场景化能力增强用户粘性和差异化竞争力。

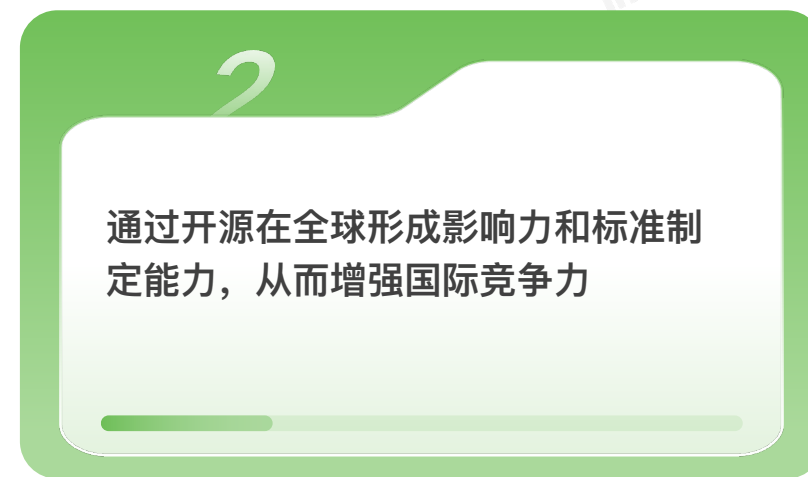
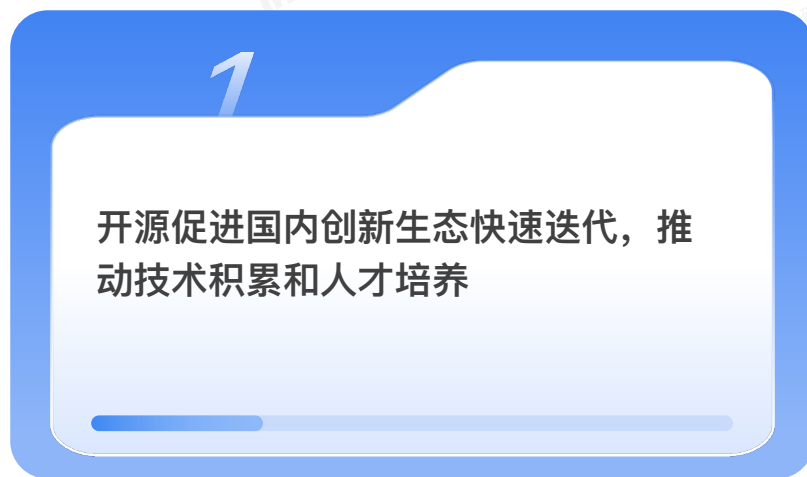
2

**多数AI硬件仍依赖成熟的手机生态，提升体验流畅性**

大多数AI硬件仍依赖成熟的手机生态完成用户连接和管理，从而降低用户使用门槛，提升体验流畅性。无论是设置、数据同步、还是日常管理，手机仍是用户接触和控制这些设备的主要渠道。

## 趋势八：有竞争就有动力，中国继续以开源撬动世界影响力

- ◆ 2026，中国仍然继续以开源持续构建自身在AI领域的竞争力。但除了模型的开源本身，中国技术力量也将更加注重生态建设与应用落地，通过在AI infra、应用等方面的技术成果积累，推动技术积累快速转化为产业能力。
- ◆ 除此之外，相传将在2026年春节前后更新的DeepSeek，已经以MODEL1代码，展现了一些可能的细节，例如创新的KV缓存布局、针对FP8数据格式的优化解码、跨GPU架构的灵活适配，再结合近期发布的Engram论文，业内推测下一代模型可能会存在架构上的颠覆性设计。这也意味着以DeepSeek、Qwen为代表的中国大模型开源力量，将进一步创新突破，引领发展。

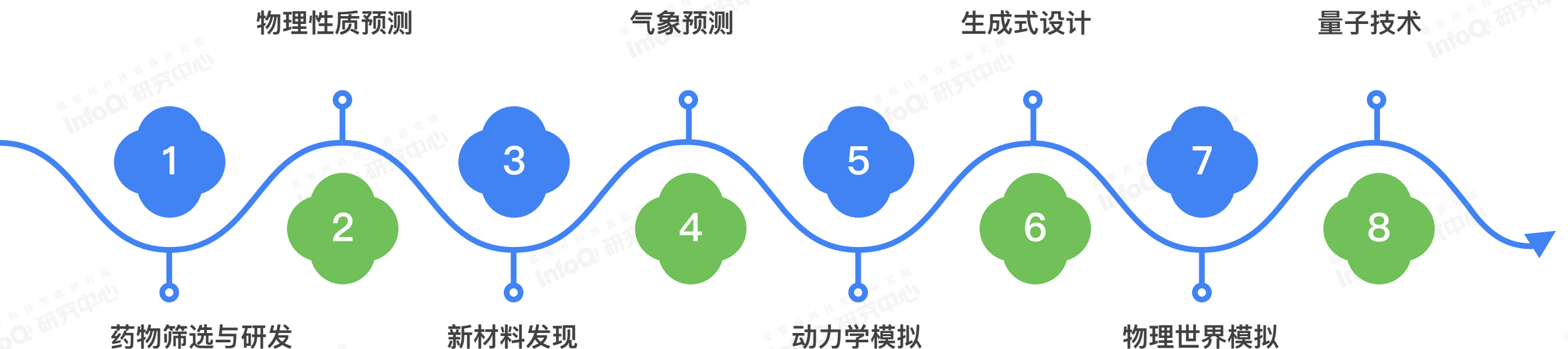


### 围绕AI的开源生态建设

注重生态建设与应用落地，通过在AI infra、应用等方面的技术成果积累，推动技术积累快速转化为产业能力。

## 趋势九：AI for Science推动科研生态升级，科学伦理面临深刻变革

- ◆ 026 年，AI for Science正在驱动科研生态升级与重构。人工智能在新材料设计、基因组分析、药物筛选、航天发动机模拟等科研场景中的应用，正在从单点工具化向系统化基础设施转变，为科研提供端到端智能支持。
- ◆ 与此同时，随着 AI 在科研决策、数据处理和自动化实验中的广泛参与，科学伦理面临重大变革：科研过程的透明性、数据可信性、实验责任归属以及人工智能的决策边界都将受到前所未有的挑战，科研伦理规则和监管体系需要同步升级。





## 趋势十：前沿技术交融，智能协作开启新格局，系统级能力强化科技与战略话语权

- ◆ 2026年，量子、星地互联网、低空飞行与AI等前沿技术加速交融，形成多域协同的智能闭环，推动无人机、卫星与地面网络的高效协作。系统级能力的提升不仅加速产业创新，也在持续增强国家在科技前沿与战略安全领域的话语权。

### 量子技术 + 低空飞行

低空飞行安全通信、高精度量子导航与定位

### AI + 量子技术

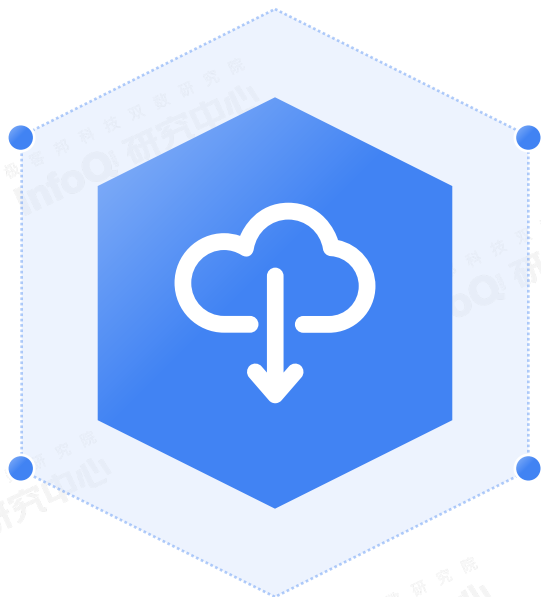
量子计算加速AI训练与推理、AI优化量子系统与量子纠错

### AI + 低空飞行 + 星地互联网

全球低空通信网络覆盖、空域资源智能化利用、复杂场景和恶劣天气下的飞行任务执行

### 量子通信 + 星地互联网

全球安全通信网络、跨空域安全协作与应急通信



极客邦科技双数研究院



InfoQ 研究中心隶属于极客邦科技双数研究院，秉承客观、深度的内容原则，追求研究扎实、观点鲜明、生态互动的目标，聚焦创新技术与科技行业，围绕数字经济观察、数字人才发展进行研究。

InfoQ 研究中心主要聚焦在前沿科技领域、数字化产业应用和数字人才三方面，旨在加速创新技术的孵化、落地与传播，服务相关产业与更广阔的市场、投资机构，C-level 人士、架构师/高阶工程师等行业观察者，为全行业架设沟通与理解的桥梁，跨越从认知到决策的信息鸿沟。

## 技术市场趋势洞察



- 市场份额追踪
- 细分市场分析
- 市场规模预测
- 市场分析模型输出

## 技术市场用户分析



- 用户规模评估
- 用户认知分析
- 用户决策分析
- 用户行为分析

## 数字化实践趋势分析



- 技术需求洞察
- 技术实践分析
- 应用规划建议
- 发展趋势研判

内容咨询: [researchcenter@geekbang.com](mailto:researchcenter@geekbang.com)

商务合作: [hezuo@geekbang.com](mailto:hezuo@geekbang.com)

- 极客邦科技，以“推动数字人才全面发展”为己任，致力于为技术从业者提供全面的、高质量的资讯、课程、会议、培训等服务。极客邦科技的核心是独特的专家网络和优质内容生产体系，为企业、个人提供其成功所必需的技能 and 思想。
- 极客邦科技自 2007 年开展业务至今，已建设线上全球软件开发知识与创新社区 InfoQ，发起并成立技术领导者社区 TGO 鲲鹏会，连续多年举办业界知名技术峰会（如 QCon、ArchSummit 等），自主研发数字人才在线学习产品极客时间 App，以及企业级一站式数字技术学习 SaaS 平台，在技术人群、科技驱动型企业、数字化产业当中具有广泛的影响力。
- 2022年成立双数研究院，专注于数字经济观察与数字人才发展研究，原创发布了数字人才粮仓模型，以此核心整合极客邦科技专业的优质资源，通过 KaaS模式助力数字人才系统化学习进阶，以及企业数字人才体系搭建。
- 公司业务遍布中国大陆主要城市、港澳台地区，以及美国硅谷等。十余年间已经为全球千万技术人，数万家企业提供服务。



促进数字技术领域知识与创新的传播



科技领导者同侪学习社区



数字人才的移动知识库



一站式数字技术学习 SaaS 平台

# 洞察技术创新趋势，推动数字化商业升级



InfoQ 公众号



AI前线 公众号



InfoQ 视频号

---

内容咨询：[researchcenter@geekbang.com](mailto:researchcenter@geekbang.com)

商务合作：[hezuo@geekbang.com](mailto:hezuo@geekbang.com)

---