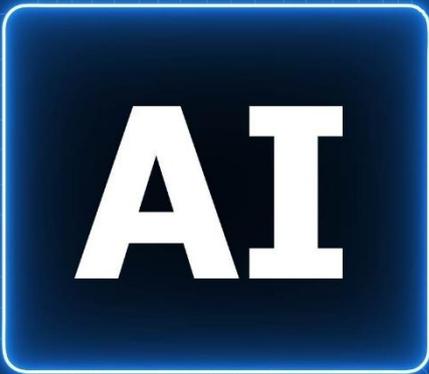


36氪研究院  
36KR RESEARCH

# 36Kr

## 《2025年中国大模型行业发展研究报告》

CBDG四维生态成为新范式，体系化竞争成为关键

A central graphic featuring the letters 'AI' in a bold, white, sans-serif font inside a dark blue square with a glowing white border. This square is surrounded by a complex network of glowing blue lines and dots, resembling a circuit board or a neural network. The background is black, and the overall aesthetic is futuristic and technological.

AI

2025.11

## 报告摘要

- **大模型已成为中国科技创新与产业安全的战略支点，政策体系正加速升级。**
- 中国将大模型视作科技自立自强和产业体系重塑的重要支点。目前，中国大模型产业政策正在形成“创新驱动、基础强化、场景牵引、安全治理”四位一体的发展格局。一方面，通过顶层设计与政策协同，为大模型提供了持续演进的制度环境；另一方面，通过算力、算法、数据、安全等要素的系统推进，建立创新驱动与责任治理并重的发展路径。
- **CBDG四维生态成为中国大模型发展新范式，中国大模型企业正在转向全方位的体系化竞争。**
- 依托中国市场特点与国家政策导向，消费者、企业、设备、政府（CBDG）四大维度构成中国大模型发展新范式。四大维度形成高效运转的飞轮效应，理解CBDG是把握中国大模型市场增长逻辑的关键。以CBDG分析可知，中国大模型企业的竞争格局正在发生深刻转变，从技术和产品单点对决转向全方位的体系化竞争，核心是生态构建、技术研发、行业赋能、商业变现、创新拓展五大能力的全面较量。五大能力共同界定了大模型企业在当前市场的战略地位和未来发展潜力。
- **未来大模型将推动人工智能范式从多模态认知迈向具身化智能，并加速通向超级智能的演进。**
- 未来大模型将引领人工智能从多模态认知迈向具身化智能，成为连接虚拟与物理世界的核心枢纽。技术正从感知驱动向认知推理跃升，智能体化进程加速推进，大模型凭借涌现性特征驱动具身智能系统自我进化，构筑通向超级智能的关键桥梁。

## 案例分析公司

### 科大讯飞

G/B端深耕与抢滩AI终端  
构筑壁垒

### 字节跳动

数据飞轮与C端流量驱动的  
生态闭环

### 阿里

夸克的用户粘性策略与  
差异化突围

## 相关研究报告

36Kr- 新型需供关系驱动  
下的中国AI文旅发展趋势  
报告2025  
(2025.09)

36Kr-2025年中国AI应用  
出海企业发展需求洞察报  
告  
(2025.07)

36Kr-2024年具身智能产  
业发展研究报告  
(2024.09)

# 目录 CONTENTS

## 01 大模型行业发展概况

- 定义与研究范畴
- 发展背景
- 发展现状

## 02 大模型生态新范式与企业竞争力

- CBDG四维生态模型成为新范式
- 中国大模型企业竞争力

## 03 大模型企业典型案例分析

- 科大讯飞：G/B端深耕与抢滩AI终端构筑壁垒
- 字节跳动：数据飞轮与C端流量驱动的生态闭环
- 阿里：夸克的用户粘性策略与差异化突围

## 04 大模型行业未来发展趋势

- 技术趋势
- 产业趋势
- 治理趋势

# 01

## 大模型行业发展概况

- 定义与研究范畴
- 发展背景
- 发展现状

## 大模型是新一代人工智能发展的战略核心与关键引擎，正驱动其实现从单点算法创新到整体智能生态的跨越

- 大模型（人工智能大模型），是指基于深度学习技术，利用海量数据训练而成、具有大规模参数和复杂计算结构的机器学习模型。此类模型能够理解并生成多模态内容，具备逻辑推理和问题求解能力，已推动新一代人工智能发展的核心载体和关键基础设施。
- 大模型的类别体系可从应用、部署与技术三个关键维度予以解析：在应用广度上，分为具备广泛跨领域适应性的通用大模型，与针对金融、医疗、教育、政务等垂直领域进行深度优化的行业大模型，后者在专业知识与业务认知上更具深度。在部署形态上，存在依托云端大规模算力资源、支持高效训练与服务的云端大模型，以及通过轻量化设计在终端本地部署、以敏捷响应安全与实时需求的端侧大模型。在技术路径上，区分为强调商业化闭环与安全可控的闭源大模型，和依托开源社区实现技术普惠与协同创新的开源大模型。
- 本报告聚焦于通用大模型的技术演进与战略价值，并重点洞察行业大模型在重点领域的应用实践。报告将系统梳理对大模型发展的宏观概况、生态格局和典型案例，呈现一幅多层次、全景式的发展蓝图。

分类维度	分类方式	典型类型	核心特征与应用举例
应用广度	通用大模型 行业大模型	通用大模型（GPT-4、Claude、讯飞星火等） 行业大模型（金融、医疗、文旅、政务等）	通用大模型具备跨领域泛化能力；行业大模型在垂直场景中深度微调，强化专业知识与业务理解
部署形态	云端大模型 端侧大模型	云端模型（ChatGPT、通义千问） 端侧模型（手机AI助手、企业私有部署模型）	云端模型依托算力中心，支持大规模训练与推理； 端侧模型注重轻量化、安全性与实时响应
技术路径	闭源大模型 开源大模型	闭源模型（OpenAI、Anthropic等） 开源模型（Llama、Mistral、InternLM等）	闭源模式以商业化与安全管控为核心；开源模式强调生态共建与创新扩散
智能层次	知识理解 → 多模态融合 → 联觉交互 → 自主学习 → 自我认知	阶段1：语言生成、问答、检索 阶段2：图像、语音、视频理解 阶段3：联觉交互（跨模态感知与反馈） 阶段4：连续任务执行、策略迭代 阶段5：具身智能、AGI探索	体现AI从“感知、认知、行动、反思”的演进路径，是未来智能进化的主线
价值链位置	模型层、平台层、应用层	模型层（通用/行业模型研发） 平台层（模型微调、部署、评测） 应用层（AI助手、智能体、行业应用）	展现从基础能力到产业落地的全链条生态，是衡量产业成熟度的重要依据

图示：大模型分类体系

参考资料：36氪研究院根据公开资料整理

## 全球大模型行业正从技术竞速转向生态竞合，中国凭借完备产业链与广阔应用场景加速跃迁

- 自2017年Transformer架构问世，全球大模型技术从预训练语言模型迭代到大语言模型，继而拓展至多模态与智能体。其演进呈现出从知识获取到语言理解、再到跨模态融合与自主协作的渐进路径，整体加速迈向通用人工智能。
- 全球大模型产业格局正由少数科技巨头依托算力、数据和资本优势形成的垄断态势，逐步转向多极竞争与生态共建。**随着开源开放降低技术门槛，各国加快算力基础建设、拓展应用领域，竞争主体不断增多，合作网络持续延伸。当前，美国在基础研究和商业化生态方面保持领先，欧洲以开源发展与合规监管构建差异化路径，中国则凭借超大市场规模、丰富应用场景和端侧生态快速迭代，稳居全球发展第一梯队。
- 当前，中国大模型行业的发展进入快速跃升阶段，形成了“技术演进与场景协同并进”的独特路径。**中国厂商已成功从技术跟随转向并行乃至部分引领，在应用广度和特定技术深度上跻身国际前列。我国大模型技术持续突破，参数规模普遍已达千亿级，部分头部模型在文本和多模态内容理解与生成、深度推理等核心能力上已比肩全球顶尖水平。在应用落地上，中国依托庞大的市场规模和丰富的产业场景优势，加快推动大模型在金融、政务、教育、医疗等领域实现规模化部署，形成了从通用能力到行业应用的双轮驱动；行业新兴方向上，端侧大模型与智能体生态快速成长，多智能体协作、多模态融合和复杂任务执行的演进趋势日益显现。

区域	战略定位	发展阶段	技术与产业优势	主要特征与趋势
中国	凭借超大市场规模与丰富场景	全球第一梯队	产业链完整，政策体系完善，应用落地能力突出	产业链完善，落地速度快
美国	聚焦基础模型、算法创新与AGI验证	全球第一梯队	算法与算力领先，私营科技企业主导，商业化路径成熟	创新驱动强，生态集中度高
欧洲	以监管和伦理为导向，强调安全与合规治理	全球第二梯队	优势在标准制定与学术研究深度，公私合作模式稳健	合规导向，发展节奏平稳

图示：中美欧大模型发展优势比较

参考资料：36氪研究院根据公开资料整理

## 大模型是中国人工智能发展的战略引擎，推动科技创新、安全治理与产业转型深度融合

- **大模型已成为国家科技创新体系和产业转型升级的关键环节。**人工智能正成为全球新一轮科技革命与产业变革的核心引擎，大模型是通用智能的重要载体。中国将大模型视作科技自立自强和产业体系重塑的重要支点，其发展不仅关乎创新能力和产业竞争力，更关乎国家安全、数字主权和社会治理能力的全面提升。通过系统布局政策体系、强化技术攻关与安全治理，中国正加快构建自主可控、创新驱动的人工智能产业生态。
- **中国人工智能政策体系正从战略谋划迈向系统构建。**从政策脉络来看，中国逐步构建起以战略导向、基础支撑、应用牵引、安全治理为主线的大模型政策体系。自2017年《新一代人工智能发展规划》（国发〔2017〕35号）首次提出建设世界主要人工智能创新中心目标以来，国家密集出台算力基础设施建设、行业应用推广、数据要素治理、算法安全监管等系列文件，形成了纵向衔接、横向协同的制度框架，持续释放政策协同效应。特别是2025年国务院印发的《关于深入实施“人工智能+”行动的意见》（国发〔2025〕11号），进一步明确以人工智能赋能千行百业、提升经济社会运行效率，标志着我国人工智能政策体系进入以融合创新与安全治理并重的新阶段。



图示：中国人工智能政策体系演进路径图  
参考资料：36氪研究院根据公开资料整理

## 大模型已成为中国科技创新与产业安全的战略支点，政策体系正持续完善与深化

- 总体来看，中国大模型产业政策正在形成“创新驱动、基础强化、场景牵引、安全治理”四位一体的发展格局。一方面，通过顶层设计与政策协同，为大模型提供了持续优化的制度环境；另一方面，通过算力、算法、数据、安全等要素的系统推进，建立创新驱动与责任治理并重的发展路径。未来，随着“人工智能+”行动的全面深化，中国将进一步完善人工智能政策与产业生态，构建以自主创新为核心、安全可控为底线、融合赋能为导向的中国特色大模型发展模式，为全球人工智能治理输出可复制、可推广的“中国方案”。

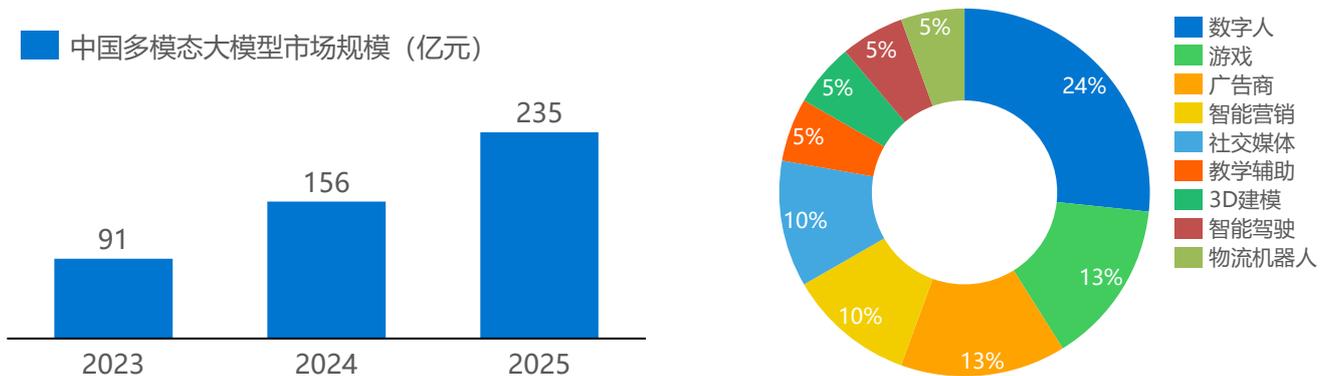
发布时间	政策文件（代表性）	发布单位	阶段层级	核心内容	对大模型发展的作用
2017.07	《新一代人工智能发展规划》（国发〔2017〕35号）	国务院	战略导向	明确人工智能为国家战略，到2030年建成世界主要AI创新中心	奠定国家战略定位与目标
2022.01	《“十四五”数字经济发展规划》（国发〔2021〕29号）	国务院	基础支撑	构建一体化算力网络体系，推进国家枢纽节点与智算中心建设	为大模型训练与推理提供算力底座
2022.03	《互联网信息服务算法推荐管理规定》	国家网信办等四部门	算法治理	规范算法推荐服务行为，建立算法备案、透明与人工干预机制	强化模型可解释与责任管理
2022.08	《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》（国科发规〔2022〕199号）	科技部等六部门	应用牵引	细化机器人应用方向，聚焦“AI+制造业”“AI+政务”“AI+教育”“AI+医疗”等场景创新	推动通用模型向行业化、场景化转化
2023.07	《生成式人工智能服务管理暂行办法》	国家网信办等七部门	安全治理	明确算法合规、安全责任与内容管理要求	建立全球领先的生成式AI治理框架
2023.10	《算力基础设施高质量发展行动计划》（工信部联通信〔2023〕180号）	工信部等六部门	基础支撑	推动算力、网络、存储、应用协同发展，完善智算体系布局	夯实大模型算力和网络支撑能力
2024.07	《国家人工智能产业综合标准化体系建设指南（2024版）》（工信部联科〔2024〕113号）	工信部等四部门	标准体系	构建基础共性、关键技术、安全治理等七大标准体系	推动产业标准化与伦理治理体系化
2025.08	《关于深入实施“人工智能+”行动的意见》（国发〔2025〕11号）	国务院	行动纲领	明确“人工智能+”为新质生产力核心引擎，推动融合创新	标志AI由技术突破迈向生产力变革

图示：中国人工智能与大模型发展相关核心政策梳理（2017—2025）

参考资料：36氪研究院根据公开资料整理

## 大模型由规模驱动转向结构创新，多模态融合与智能体演进成为中国市场竞争焦点

- 进入2025年，全球通用大模型的基座能力持续演进。语言生成、长文本处理、多模态理解等关键技术加速突破，主流厂商普遍将迭代周期维持在3至4个月。随着Transformer架构及RLHF/SFT微调机制的广泛应用，模型泛化能力、任务迁移效率显著提升。根据艾媒咨询数据，2024年中国大模型市场规模约为294.16亿元，预计2026年将突破700亿元，行业正处于爆发式发展阶段。
- 多模态能力正在成为大模型在中国的新一轮竞争核心。**传统的文本处理正在向图像、语音、视频、代码等多模态交互拓展，跨模态融合推动大模型由感知智能向认知智能演进。根据智研咨询数据，2024年我国多模态大模型市场规模为156.3亿元，其中，数字人领域占比24%，游戏与广告商拍各占13%，智能营销、社交媒体分别占10%，教学辅助、3D建模等领域合计占30%。
- 长文本处理能力与智能体（Agent）进程正同步深化。**随着上下文窗口持续拓展，长文本生成、结构化推理和流程化任务执行成为热点。作为大模型落地的重要承载，智能体通过任务分解、自主决策、工具调用与环境交互，显著拓宽了模型的任务边界。在此进程中，国内开源与闭源路径并行发展：以DeepSeek为代表的开源力量加速崛起，带动中国开源生态快速扩张，而闭源巨头则依托算力、专有数据和产业链整合构筑壁垒。总体来看，核心基座能力日趋收敛，竞争中心从功能实现转向质量与应用效果的比拼。



图示：中国多模态大模型市场规模及应用场景分布  
数据来源：智研咨询，36氪研究院整理

## 大模型商业化进入体系化阶段，企业通过API、授权、方案与服务构建价值闭环

- 大模型产业正从规模化扩张迈向体系化商业落地的新阶段。主流厂商普遍确立了以API调用、模型授权、行业解决方案与增值服务为核心的四大变现路径，推动大模型从技术产品向综合服务体系演进。
- 在基础服务层面，API调用成为最主流的商业化模式。** 凭借其通用性和规模化优势，企业通过标准化接口提供模型能力，实现“按调用付费+订阅制”的持续收益。根据Grand View Research数据，2024年中国人工智能API市场收入约35.6亿美元，预计2030年将达194亿美元，年均复合增长率 32.7%。
- 在行业应用层面，模型授权与行业级解决方案是价值拓展主要方向。** 金融、政务、制造等领域对安全性与定制化需求日益增长，厂商通过模型微调、私有化部署与垂直优化，构建起模型即服务（MaaS）与行业智能体等商业模式。以科大讯飞为代表的行业头部企业，正在探索“模型底座+行业知识库+业务插件”复合型结构，推动商业模式从单一授权向体系化能力输出转变。
- 在生态拓展层面，增值服务成为大模型企业新的盈利增长点。** 企业通过提供数据治理、RAG增强检索、模型监控与评测、智能体开发平台等配套服务，构建开放生态体系，实现从技术能力输出向长期运营价值的深化。随着企业用户规模扩大，模型管理与安全审计等服务的商业价值将持续提升。

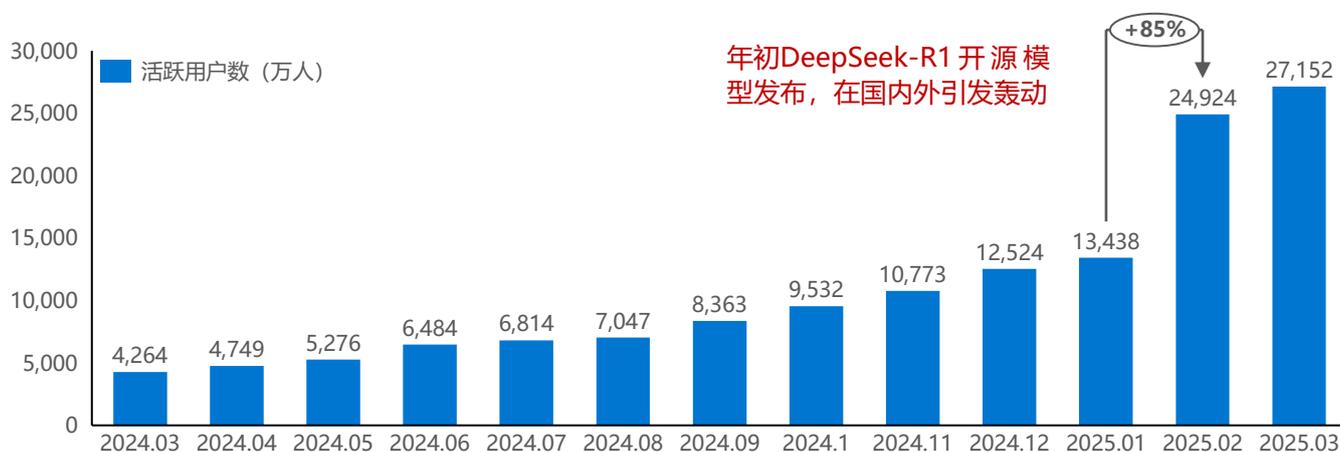
模式类型	主要特征	典型代表企业	盈利机制	发展趋势与特点
API调用	以标准接口形式开放模型能力，按调用量计费，适用于开发者与企业级客户	OpenAI、百度智能云、阿里云、讯飞星火开放平台	订阅制、按量付费制	通用化强、生态扩张快，成为基础变现模式
模型授权	通过模型授权或嵌入式集成，提供企业专属模型能力	Anthropic、商汤科技、智谱AI	授权费、长期使用费	满足高安全场景，增强客户粘性
解决方案	针对行业痛点提供“模型+场景+应用”一体化方案	科大讯飞（教育、政务）、百度文心（能源、交通）、腾讯混元（文旅）	项目制收入、解决方案交付费	场景渗透深化，成为商业落地主流
增值服务	提供RAG增强、智能体开发、MLOps管理等服务	华为昇腾、阿里通义千问、MiniMax	订阅费、运维费、服务佣金	多主体协作，共建产业生态

图示：大模型产业探索多元商业模式

数据来源：36氪研究院根据公开资料整理

## 大模型应用正从试点验证迈向规模化深耕，C端创新持续活跃，B端与G端成为价值实现的主阵地

- 大模型应用正由试点示范向规模化落地加速转变。随着技术门槛降低与生态资源整合，企业端（B端）和政务端（G端）逐渐成为价值重心。
- 在消费端（C端），AI原生应用快速普及，内容与娱乐类场景成为创新重点。**根据QuestMobile数据，截止到2025年3月，中国AI原生应用月活用户已突破2.7亿，月人均使用时长达123.6分钟。随着写作助手、绘画生成、视频创作与虚拟助理等产品不断迭代，AI正逐步成为内容创作与交互体验的基础工具，产业发展的重心也转向用户习惯培育和新型使用场景的构建。
- 在企业端（B端），大模型正成为数字化转型的关键引擎。**根据IDC预测，到2026年，全球三分之一以上组织将实现从试点到规模化部署的跨越，届时AI驱动的任务自动化与流程优化将覆盖60%以上企业工作负载。与此同时，科大讯飞、百度等企业基于自研模型，已推出金融智能体、政务助手、制造智脑等行业解决方案，推动产业智能化从局部创新迈向系统性赋能的新阶段。
- 在政务端（G端），大模型已成为数字政府建设和治理现代化的重要技术支撑。**各地政府正依托政务大模型，深化政策问答、审批流转、文稿生成、信息汇聚、民意分析等功能应用，服务范围已覆盖政务咨询、政策生成、民意分析、文档校对等高频行政场景，进一步助力政务数据要素体系建设。

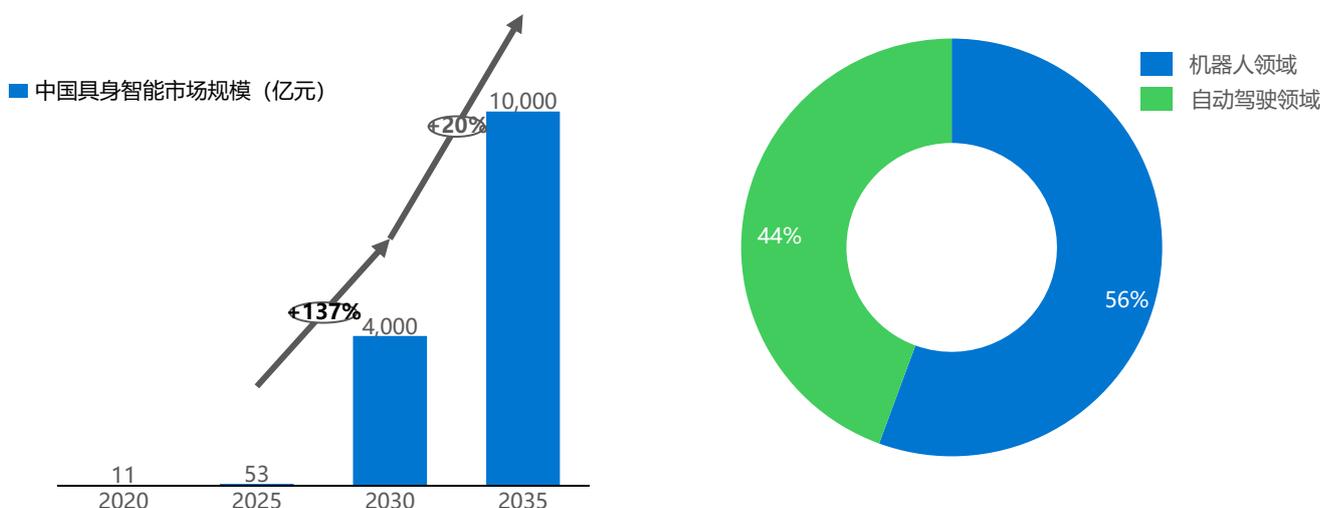


图示：中国AI原生App用户规模及增长率（2024.03–2025.03）

数据来源：QuestMobile，36氪研究院整理

## 大模型厂商多向探索，具身智能成为智能进化的新高地，研究与应用布局同步推进

- 大模型的技术演进已进入从通用能力构建向多方向探索的阶段。国内厂商在持续优化算法、模型与算力体系的同时，正围绕多模态融合、智能体演化、端侧部署等方向展开广泛实验。其中，具身智能因具备显著的技术牵引效应与产业成长潜力，逐渐成为行业共识性探索方向。
- 具身智能的核心在于将语言模型与多模态模型的认知能力嵌入机器人、车载系统和智能终端等物理载体，使模型具备感知环境、自主推理、任务执行与情境交互的能力。它不仅是大模型能力外化的最佳载体，也是验证大模型通用性与智能水平的理想载体。对企业而言，具身智能既是技术边界的实验平台，也是连接产业应用的新型增长点。
- 根据国务院发展研究中心《中国发展报告2025》，到2030年中国具身智能市场规模预计达到4,000亿元，2035年有望突破万亿元，增长潜能巨大。从应用结构来看，具身智能技术的市场应用主要集中在机器人和自动驾驶载具两大领域。其中，机器人领域占比约56%，自动驾驶领域占比约44%，构成当前具身智能产业发展的两大核心赛道。



图示：中国具身智能市场规模预测及应用市场分布

数据来源：国务院发展研究中心、智研咨询，36氪研究院整理

## 1.3 发展现状

- 具身智能平台作为连接认知智能与物理执行的中枢载体，承担着将语言理解、视觉感知、任务规划等认知能力转化为具体行动的核心使命。当前，中国具身智能技术体系正从算法突破走向系统集成与产业落地的阶段，在具身智能的系统化演进过程中，企业正基于自身技术积累形成差异化路径。
- 中国具身智能平台形成“智能认知、多机协同、行业融合”三大体系，构建起贯通“感知、认知、执行、应用”的技术演进格局。其中，智能认知体系以大模型为核心，聚焦语义理解、多模态融合与任务决策，成为推动具身智能由认知驱动向行动落地转化的关键阶段。以科大讯飞为代表，其“机器人超脑”平台依托“讯飞超脑2030”技术底座，率先实现语音、视觉、运动控制等底层能力统一集成，形成软硬一体、跨设备协同的智能生态。

体系类型	技术定位	代表企业 / 平台	核心特征	典型应用领域
智能认知	具身智能的大脑核心，聚焦多模态认知与任务规划	科大讯飞：机器人超脑平台 商汤科技：悟能平台	多模态融合、认知推理、软硬一体	服务机器人、教育陪伴、政务交互
多机协同	具身智能的执行基座，聚焦运动控制与群体协作	优必选：群脑网络 达闼科技：云端大脑系统	群体调度、云端协同、系统稳定	制造业、仓储物流、安防巡检
行业融合	具身智能的产业接口，推动模型与场景融合落地	华为云：CloudRobo平台 百度：Apollo车载平台	行业知识融合、仿真训练、生态协同	交通出行、工业制造、教育培训

图示：中国具身智能平台类型与代表案例对比

参考资料：36氪研究院根据公开资料整理



图示：科大讯飞“机器人超脑”平台核心能力体系

参考资料：36氪研究院根据公开资料整理

## 02

# 大模型生态新范式与企业竞争力

- CBDG四维生态模型成为新范式
- 中国大模型企业竞争力

## CBDG四维生态模型：解析中国大模型发展的新范式

- 与全球其他市场相比，中国大模型行业发展呈现出明显的复合性特征：它既植根于庞大的内需市场，又受到国家政策导向的引领与塑造，同时还顺应了软硬件一体化等技术发展趋势。为解释这一复杂系统中的内在逻辑与驱动力，CBDG四维生态模型成为分析新范式。该模型揭示，消费者（Consumer）、企业（Business）、设备（Device）与政府（Government）四大维度并非简单并列，而是通过深度耦合，形成了一个相互作用与赋能的有机整体。

### 消费者 (Consumer)

是应用创新的活力源泉与商业模式验证的试金石，其快速反馈驱动模型迭代

### 企业 (Business)

是技术价值实现的核心阵地，通过将大模型深度融入生产流程，催生效率革命与商业模式创新

### 设备 (Device)

作为大模型触达物理世界的重要载体，通过手机、汽车、机器人、AIoT设备等实现能力的无处不在，形成软硬一体与数据闭环

### 政府 (Government)

是引导者、规范者与关键用户的多重角色，通过顶层设计、政策扶持与公共领域采购，为产业发展指明方向、提供初期市场并构建有序环境。



图示：CBDG四维生态模型

- 消费者（C端）是需求牵引与创新源泉。**中国庞大的互联网用户群体不仅为模型迭代提供了海量数据燃料，其快速演进的数字化生活习惯和对新技术的积极拥抱态度，也为大模型的创新应用提供了宝贵的实践场。用户反馈的实时注入，加速推动大模型从技术可用向产品好用的跨越。
- 企业（B端）是价值创造与落地核心。**强大的企业级市场是技术实现商业闭环的主战场，各行各业迫切的降本增效与智能化转型需求，驱动大模型技术与行业Know-how深度融合，催生面向特定场景的深度解决方案，从而将技术潜力转化为切实的生产力提升，推动商业模式创新。

- **设备（D端）是能力载体与生态接口。**中国强大的智能设备制造与普及能力，为大模型提供了无缝融入物理世界的广泛入口。从AI PC、AI手机到AI终端，从智能网联汽车、到人形机器人，硬件载体使得大模型能力无处不在，并在此过程中形成感知、决策、反馈的数据闭环，持续优化系统智能。
- **政府（G端）是战略引导与秩序保障。**政府通过顶层设计、产业政策、标准制定与先行先试的公共采购，扮演着多重角色。它不仅是产业发展的方向指引者和初期市场的重要提供者，同时也是安全、可信、可控发展红线的划定者，为行业的长期健康发展构建了规范有序的环境。
- **这四大维度之间构成了一个高效运转的飞轮效应。**G端的政策引导与场景开放，催化了B端的技术攻关与方案落地；B端的技术赋能提升了C端的应用体验与D端的终端智能；C端的使用反馈与D端采集的数据，反过来又优化了B端的模型与方案，并为G端的精准治理与科学决策提供了新工具。这一自我增强的闭环，共同构成了中国大模型行业具有本土特色的核心驱动力。
- **理解CBDG模型，是准确把握中国大模型市场增长逻辑的关键。**该模型揭示了驱动中国大模型行业发展的是多因素变量，而非单一的技术突破。G端的顶层设计通过产业政策与公共采购，清晰指明了技术发展的优先领域与关键赛道，企业因此敢于进行面向实体经济的长期、重投入的领域深耕，而非仅仅追逐短期的流量红利。这使得中国的大模型产业能够迅速在智能制造、智慧政务、生物医药等国家战略亟需的领域形成有效供给能力。同时，纯粹的云端模型服务在面临增长瓶颈与同质化竞争时，D端（设备）的深度整合成为破局关键。通过软硬结合的方式，让大模型更深度地嵌入生产流程与生活场景，能够有效突破当前困局。通过CBDG这一模型，可以看到，未来的行业竞争制高点，将不再仅由少数几个通用大模型的性能决定，而是取决于谁能在CBDG构建的复杂网络中，更高效地整合资源、捕捉需求并创造闭环价值。这意味着，拥有深厚行业知识（B端）、强大硬件生态（D端）并能深刻理解政策导向（G端）以服务海量用户（C端）的企业，将获得更大的生态主导权与价值分配权。

## 基于CBDG四维生态模型，构建中国大模型企业竞争力指标评价体系

- 基于CBDG四维生态模型分析可知，中国大模型企业的竞争格局已发生深刻转变，企业竞争力已经从技术或产品的单点对决，演变为全方位的体系化竞争，主要体现在生态构建、技术研发、行业赋能、商业变现与创新拓展等五大能力维度的全面较量。这五大能力共同界定了企业在当前及未来市场格局中的战略地位与发展潜力。



图示：衡量大模型企业竞争力的五大能力维度

- 基于上述五大能力，本报告构建中国大模型企业竞争力指标评价体系，用以全面、科学地评估中国大模型企业综合竞争力。该体系一方面为需求方选型提供客观、系统的价值指南；另一方面为供给方提供清晰的战略对标与自我诊断工具。指标体系的搭建遵循“价值创造-能力支撑-生态演进”的内在逻辑：价值定位（Why）是大模型企业的立身之本，决定其为CBDG四维生态（用户、企业、设备、政府）创造价值的广度与深度；能力支柱（What）是大模型企业实现价值、构建竞争壁垒的核心能力；支撑要素（How）是大模型企业维持并放大核心能力的基础运营与资源保障。三者环环相扣，共同定义大模型的市场地位与发展潜力。指标能力的权重配比设置紧扣中国大模型企业的体系化竞争特征，采用“双核心+双落地+一潜力”结构（生态构建、技术研发各占25%；行业赋能、商业变现各占20%；创新拓展占10%），形成重点突出、兼顾当前竞争力与未来潜力的权重分配体系。

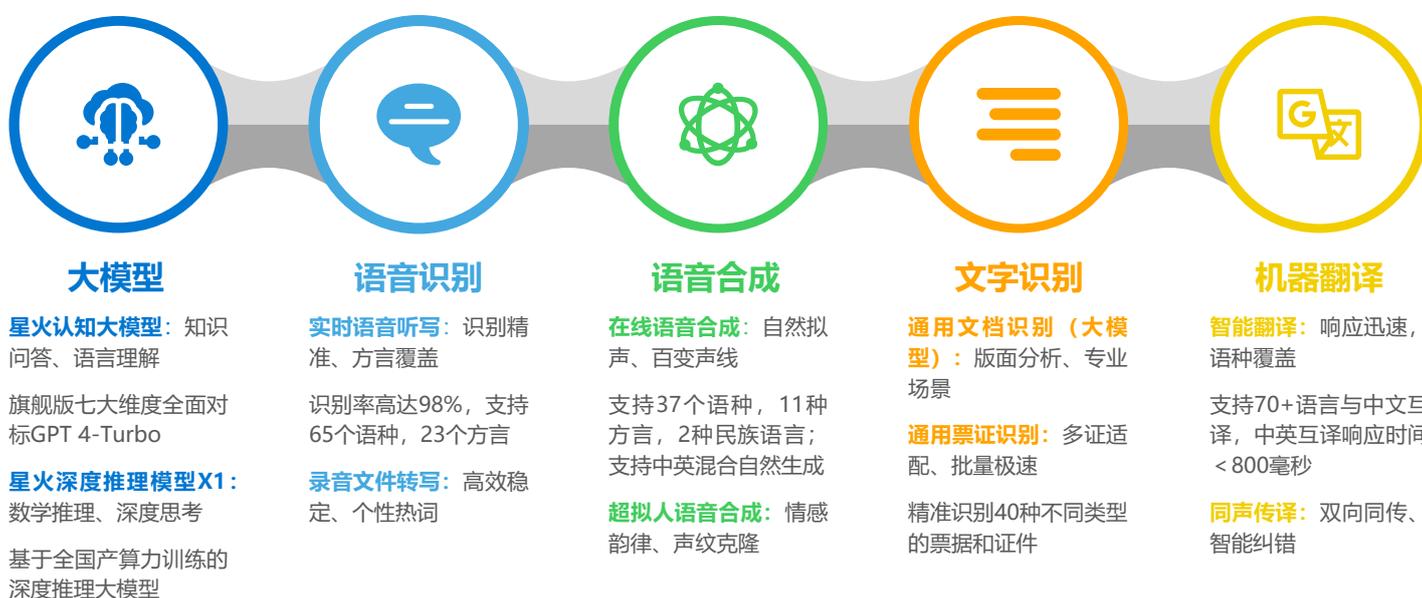
## 中国大模型企业竞争力指标评价体系

一级指标	二级指标	序号	三级指标	赋分
生态构建能力 (25%)	开发者生态	1	注册开发者数量 (万人)	5分
		2	API月度总调用次数均值 (亿次)	6分
	客户生态	3	付费企业客户总数 (家)	5分
		4	面向消费者的App月活跃用户数 (MAU, 万人)	5分
	合作伙伴生态	5	年度新增战略技术合作伙伴数量 (家)	2分
		6	年度新增行业解决方案合作伙伴 (ISV/SI) 数量 (家)	2分
技术研发能力 (25%)	通用核心性能	7	国内权威基准 (如C-Eval/CMMLU) 平均排名 (名)	3分
		8	国际综合基准 (如MMLU) 平均得分 (分)	3分
		9	代码生成基准 (如HumanEval) 通过率 (%)	3分
	关键功能特性	10	模型支持的最长上下文窗口 (K Tokens)	3分
		11	支持的多模态任务数量 (种)	3分
		12	智能体 (Agent) 框架支持的可调用工具数量 (个)	3分
	技术底蕴与迭代	13	研发人员占总员工比例 (%)	2分
		14	年度大版本迭代次数 (次)	2分
	技术开放性	15	开源模型年度下载总量 (万次)	3分
行业赋能能力 (20%)	解决方案深度	16	已发布的行业解决方案总数 (个)	8分
		17	解决方案是否深度集成行业知识库、工作流和业务逻辑	2分
	应用效果验证	18	典型客户案例中官方宣传的效率提升平均值 (%)	5分
	客户满意度	19	用户评分 (5分制)	5分
商业变现能力 (20%)	市场占有率	20	API调用付费金额 (亿元)	8分
	收入与盈利	21	年度经常性收入 (ARR, 亿元)	8分
		22	企业级客户平均合同金额 (ACV, 万元)	4分
创新拓展能力 (10%)	软硬一体整合	23	开放平台累计覆盖终端设备数 (亿个)	4分
	前沿领域探索	24	年度研发费用总额 (亿元)	3分
	产业影响力	25	获得的专利数量 (项)	3分

注：时间节点统一为截至2025年6月（如三级指标1、3等）；2. 年度统一为2024年度（如三级指标5、6等）

## 生态构建：行业竞争已经进入到高活性、高粘性、具备自演进能力的生态体系阶段

- 生态构建能力是衡量大模型企业能否从产品技术成功迈向商业生态化的关键能力。当前，行业竞争已超越简单的开发者数量积累，进入到构建高活性、高粘性、具备自演进能力的生态体系阶段。
- 在开发者生态方面，领先企业正通过提供精细化工具链、优化计费策略与强化调试支持，系统提升开发效率与应用创新活力。与此相对应，除了注册开发者规模，API月度调用总量与高质量原生应用产出率成为能够真实反映生态活跃度与价值转化能力的关键指标。以科大讯飞为例，其依托讯飞开放平台的全栈能力，满足全场景需求，已汇聚150万+大模型开发者，在行业内保持市占率领先，形成了明显的生态规模与协同优势。



### 场景产品创新，让办公轻松高效

讯飞绘文

讯飞智文

讯飞文书

讯飞智检

讯飞智作

讯飞翻译

讯飞绘镜

星火陪练

星火快答

图示：讯飞开放平台主要产品能力与场景创新

- 在客户生态构建方面，行业形成了多种路线打法，依托自身核心资源形成差异化的竞争逻辑。一类是以流量为核心的C端规模路线，如字节跳动（豆包）、阿里（夸克），通过高频场景切入和生态协同赋能快速构建用户与数据壁垒，构建品牌认知。这种路线的核心在于通过规模效应沉淀场景数据，反哺模型优化，并为后续商业变现铺垫基础。另一类是主攻B/G端的价值深耕路线，以科大讯飞为典型，通过合规底座、场景定制与生态联动构建高壁垒商业闭环。该路线以付费客户质量为核心，通过长期服务绑定形成可持续收益。2025年1-9月，公开渠道统计到的科大讯飞大模型相关项目中标数量和金额均位居第一，形成行业领先优势\*。
- 生态构建方面，2024年以来，与头部独立软件开发商（ISV）及系统集成商（SI）建立战略合作，已成为快速切入垂直领域、共建行业解决方案的主流路径。从合作模式来看，部分企业与ISV共建联合实验室，部分企业探索按结果付费等灵活合作机制。

### 以流量为核心的C端规模路线



通过具有竞争力的C端应用（如豆包、夸克）快速获取用户、沉淀数据与构建品牌认知

### 主攻B/G端的价值深耕路线



深耕B端与G端市场（如科大讯飞），以付费客户数量与客单价为核心，构建可持续的商业化深度，形成差异化竞争壁垒

图示：客户生态构建的两种路线

### 开发者生态

从开发者的数量积累转向质量与活性的提升，通过精细化工具链、灵活计费等服务驱动开发效率与创新

### 客户生态

采用C端和B/G端不同路径来构建客户生态。C端通过爆款应用快速获取用户、沉淀数据与品牌认知；B/G端以付费客户规模与客单价为核心，打造可持续商业化深度

### 合作伙伴生态

以与头部ISV/SI战略合作为核心路径切入垂直领域，合作模式从简单分销升级为技术和场景相结合的深度协同

图示：生态构建能力的主要方面

\*数据来源：智能超参数，36氪研究院整理

## 技术研发：基于基准测试排名，侧重特定场景下的技术深度、实用性与迭代速度

- 技术研发能力构成了大模型企业最核心的竞争壁垒。当前行业的竞争焦点已不再局限于通用基准测试的排名，而是转向了在特定场景下的技术深度、实用性与迭代速度的综合较量。
- 通用核心性能依然是行业准入的基本门槛，其在国内外权威基准测试中的表现是衡量模型基础能力的客观标尺。此外，真正的技术竞争力更体现在关键功能特性的突破上：模型支持的上下文长度持续突破，处理长文档和复杂代码库已成为标配；多模态能力正从静态的图文生成向动态的视频理解与交互演进；而智能体框架的工具调用能力与任务执行的可靠性，则直接决定了模型能否从辅助工具升级为真正的生产力。例如，科大讯飞以通用模型为根基，在语音与多语种领域突破显著。讯飞星火认知大模型V3.5实现文本生成、逻辑推理等七大核心能力全面提升，升级后的星火X1深度推理大模型更在数学、知识问答等通用任务上对标国际顶尖水平，在模型参数更小的情况下展现出强劲通用性能，充分满足行业准入的高阶要求。

### 1 深度推理

数理逻辑（逻辑推理、数学解题）在国内外任务测试集中表现强势；指令跟随能精准理解意图，挖掘深层需求，强化输出控制；代码生成高效可靠；多轮交互稳定性强，精准捕捉关联上下文

### 2 多模生成

多模理解完成图片素材识别理解与准确描述；视觉问答围绕上传的图片素材，响应用户问题；多模生成根据用户描述，生成合成音频和视频；虚拟人视频整合AI虚拟人，生成与描述匹配的视频

### 3 代码生成

智能生成单行或函数级代码建议，自动生成代码；对类、函数及其作用逐一详解，逐行代码注释；代码纠错，一键修改；为选中代码生成单测用例，多主流单元测试框架，智能生成单元测试数据

### 4 内容创作

根据用户要求，创作具有吸引力与情感共鸣的商业文案；根据描述，创作多种形式的营销方案；使用英文完成多风格多任务的长写作任务；根据新闻要素，创作严谨详实的新闻通稿

### 5 数学能力

可处理方程求解，涵盖多种多元方程；几何问题，涵盖平面几何与立体几何；微积分，处理导数、积分等问题；概率统计，包含随机变量、概率分布、假设检验等内容相关需求

### 6 语言理解

机器翻译，包括英语、中文、法语、德语、西班牙语等常用语种；快速理解文章核心观点，生成文本摘要；检查语法错误并提供正确的语法建议；情感分析，分析文本中的情感色彩

### 7 知识问答

提供饮食、运动、旅游等日常生活常识；沟通技巧、时间管理、团队协作等工作技能指导；基础健康保健及疾病预防、诊断、治疗相关医学知识；输出历史事件、名人故事等历史人文类文案

图示：讯飞星火深度推理X1大模型，综合性能显著提升，整体对标OpenAI o3

- 在大模型综合性能显著提升的基础上，科大讯飞持续优化语音效果，多语种技术实现技术突破，获得国家科学技术进步奖一等奖。面向万物智联时代的高噪复杂场景，科大讯飞的多模态降噪技术展现出领先实力，在CHIME-8大赛中实现五连冠，成功攻克动态多人会议等场景下的语音重叠、远场混响等难题，让高难度语音识别成为现实。此外，讯飞星火还支持130+语种问答，并于近期发布东盟多语言大模型，为世界提供了多语种智能新选择。
- 在技术底蕴与持续迭代方面，高比例的研发投入与稳定的大版本更新频率是维持技术领先地位的关键保障。例如，据浙商证券数据，字节跳动2024年在AI领域的资本开支达800亿元，接近百度、阿里、腾讯三家总和（约1,000亿元），2025年更计划将资本开支提升至1,600亿元，其中900亿元专项用于AI算力采购，支撑模型快速迭代。科大讯飞则始终保持常态化的高比例研发投入，据年报数据，2024年研发费用占营收比重19.62%，其中研发费用的16.2%专项用于大模型研发，支撑星火大模型实现季度级功能迭代。
- 行业在技术开放性上呈现出战略分化，闭源路线致力于追求极致的性能与商业控制力，开源路径则着眼于构建广泛的开发者生态与行业影响力。由于国产模型持续开源，与国际顶级闭源模型性能差距几近抹平，加之其尺寸、类别丰富，能让企业按需定制，据Frost&Sullivan预测，未来超过80%的企业将采用开源大模型，预示着开源模型将在行业应用中占据主导性增长。

### 通用核心性能



作为行业准入基本门槛，以国内外权威基准（如C-Eval/CMMLU、MMLU）表现及代码生成基准（如HumanEval）通过率为客观衡量标尺

### 关键功能特性



核心体现在上下文长度持续突破、多模态能力、智能体工具调用与任务执行可靠性提升，可通过上下文窗口、多模态任务数、可调用工具数衡量

### 技术底蕴与迭代



高比例研发投入、稳定大版本更新频率是维持领先的关键，研发人员占比、年度大版本迭代次数为重要参考维度

### 技术开放性

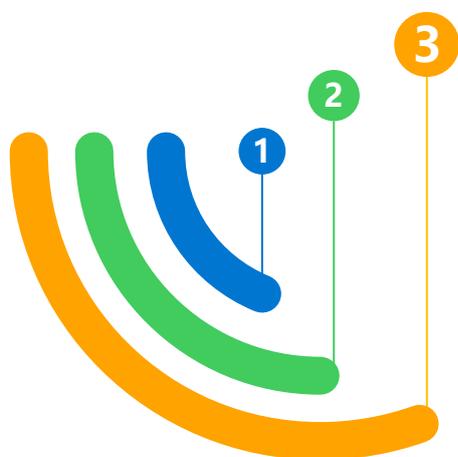


行业呈战略分化，闭源追求性能与商业控制，开源追求生态与影响力，随着国产开源模型性能逼近闭源，未来或主导行业应用增长

图示：技术研发能力的主要方面

## 行业赋能：从技术工具的单点应用向与行业知识、业务流程深度重构的根本性转变

- 行业赋能能力直接衡量大模型技术从实验室走向产业实践的深度与实效，是检验其商业价值的核心维度。当前，这一过程正经历从技术工具的单点应用向与行业知识、业务流程深度重构的根本性转变。竞争力的核心首先体现在解决方案的深度与广度上，具体而言，企业已发布的行业解决方案总数反映了其技术应用的覆盖范围，而解决方案是否深度集成行业知识库、工作流和业务逻辑，则决定了其解决实际问题的精准度与有效性。在金融风控、医疗辅助诊断、工业质检等高价值场景中，这种深度整合能力尤为关键。其次，应用效果的量化验证成为最具说服力的竞争要素，典型客户案例中通过效率提升百分比、成本降低率等具体数据，直观展现技术带来的实际价值。最后，持续的用户满意度不仅是解决方案成熟度的体现，也是推动产品持续更新迭代、建立客户忠诚度的重要保障。在金融、政务、教育、工业等垂直行业领域，能够提供深度融合的端到端解决方案、并展现出显著降本增效成果的企业，如科大讯飞在科研与求职等垂直场景中建立的实践标杆，正在构筑起差异化竞争力。



### 解决方案深度

- 不仅需要行业解决方案总数来覆盖更多领域，更要深度融合行业专属知识库、业务工作流与核心逻辑，以“广度+深度”来构筑行业壁垒

### 应用效果验证

- 以典型客户案例中的具体数据为核心验证标准，通过效率、成本、质量等维度的可量化成果来锚定商业价值，证明对业务的实际增益，避免技术空转

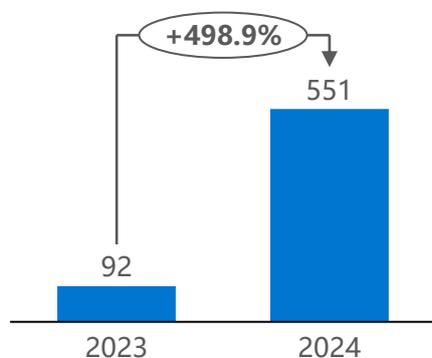
### 用户满意度

- 体现解决方案成熟度，同时通过用户满意度的反馈来反哺产品迭代，形成从结果反馈到长期价值构建的循环，

图示：行业赋能能力的主要方面

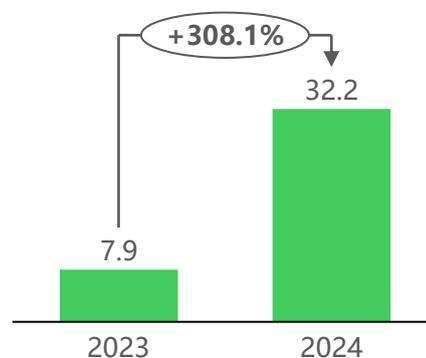
## 商业变现：从商业模式探索期进入初步的规模化变现阶段

- 商业变现能力是将技术投资转化为可持续发展的动力源泉。2024-2025年，中国大模型行业整体从商业模式探索期迈入初步的规模化变现阶段。市场占有率通过API调用付费金额这一硬指标得以清晰体现，反映了模型服务在开发者群体中的接受度与商业价值。单纯的API调用模式存在两大挑战：其一，算力成本居高不下，中小厂商的月均机器成本远超收入规模；其二，头部企业价格战进一步压缩利润空间，中小厂商陷入用户越多亏损越大的困境。因此，企业将商业模式转向了高附加值的解决方案与模型授权，这直接体现在年度经常性收入（ARR）和企业级客户平均合同金额（ACV）提升上。
- 此外，政企市场成为变现主阵地。2024年全年，可统计的有931家央国企开展大模型招标，在整体市场上的数量占比为61.25%，其中有551个招标项目披露金额约32.2亿元，占整体市场的49.8%\*。这类项目普遍具有金额大、周期长的特点，且对企业的品牌公信力、服务能力和安全合规要求极高。多种商业变现模式交织，当前市场分化已十分明显。部分头部企业如科大讯飞在政企市场斩获超14亿元的重大订单，在商业化落地层面确立了领先优势，而众多中小厂商则仍在盈亏平衡线上探索。



图示：全年市场公开披露的大模型中标项目数量（个）

数据来源：智能超参数，36氪研究院整理



图示：全年市场公开披露的大模型中标项目金额（亿元）

数据来源：智能超参数，36氪研究院整理

\*数据来源：智能超参数，36氪研究院整理

## 创新拓展：关乎在下一代产业变革中的生态位争夺，抢占未来生态制高点

- 创新拓展能力直接关乎企业在下一代产业变革中的生态位争夺，体现了其面向未来的战略布局与前瞻视野。当前，行业的共识性方向聚焦于软硬一体整合与具身智能等前沿领域。在软硬一体方面，大模型正加速与各类终端设备深度融合，其开放平台累计覆盖的终端设备数量成为衡量生态延展广度与未来影响力的关键先行指标。在前沿探索层面，具身智能作为公认的万亿级蓝海市场，吸引了科技巨头与科研机构战略性投入，其中机器人/具身大模型作为实现智能体自主决策的“中枢大脑”，已成为兵家必争之地。这一探索的强度与决心，往往通过企业的年度研发投入总额等指标得以体现。与此同时，产业影响力的构建变得愈发重要，具体表现为在核心技术领域的专利积累，以及参与乃至主导行业关键标准制定的能力。科大讯飞在该领域布局成效显著：已赋能超过500家机器人企业，机器人语音交互市占率领先，且在人形机器人行业Top20公司、商用机器人领域的语音交互应用中占有率均居前列；发布智能语音背包以模块化AI接入打通开发应用链路，已落地中石化易捷易享节、讯飞学习机营销等场景；同时获授人形机器人标准化工作组副组长单位，作为起草单位参与制定并发布《人形机器人语音交互技术规范》，全面展现前瞻布局与产业影响力。

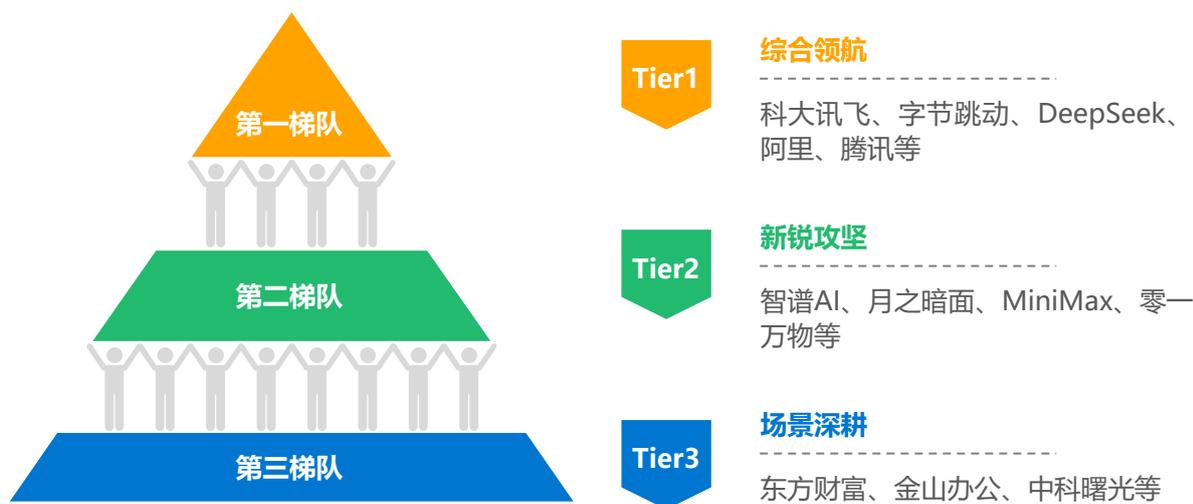


图示：科大讯飞在创新拓展维度的主要能力

## 科大讯飞凭借在生态、技术与政企市场等多维度能力优势，在大模型市场中位居第一梯队

- 当前，中国大模型领域已进入体系化竞争阶段，市场呈现出多元且复杂的竞争态势。市场参与者主要可划分为以下五类主体：以全栈布局构建护城河的互联网/科技巨头、以技术锐度著称的专注于模型的新贵、以行业纵深创造价值的垂直领域深耕者、在转型中寻求突破的AI 1.0转型者，以及作为创新源头的科研机构。其中，部分企业通过多维度能力建设取得跨越式发展，如科大讯飞凭借其百万量级的开发者生态、在语音与多语种技术领域的持续突破、政企市场的规模化落地能力以及在具身智能等前沿领域的创新布局，构建了全面的综合竞争优势，位居行业玩家第一梯队。

市场主体	主要特点	典型企业
互联网/科技巨头	打造从芯片、框架、模型到应用的全栈式生态。有自有场景和流量，以云服务为出口，技术栈全面	阿里、百度、华为、科大讯飞、腾讯、字节
模型技术新贵	专注于大模型技术本身的突破，追求在通用能力或特定技术点上达到世界领先水平	DeepSeek、零一科技、Minimax、月之暗面、智谱AI
垂直领域深耕者	将大模型能力与特定行业知识和数据相结合，解决行业核心问题。以解决方案为导向，深厚的行业积累、数据和客户资源是其护城河	东方财富、金山办公、医渡科技
AI 1.0转型者	原本在计算机视觉、语音识别等时代领先的AI公司，全力向大模型赛道转型	旷视科技、商汤科技、云从科技
科研机构	它们虽非企业，但通过发布开源模型，是技术和人才的重要源泉	北京大学、北京智源研究院、中国科学院



图示：中国大模型市场竞争梯队

注：中国大模型市场的典型企业由36氪研究院梳理，只列出部分企业为代表，未覆盖全产业

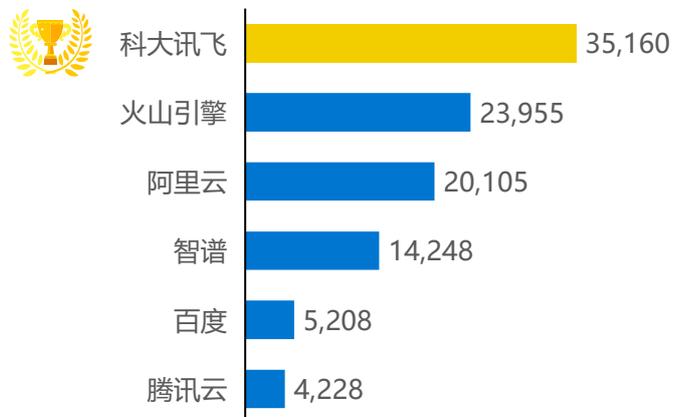
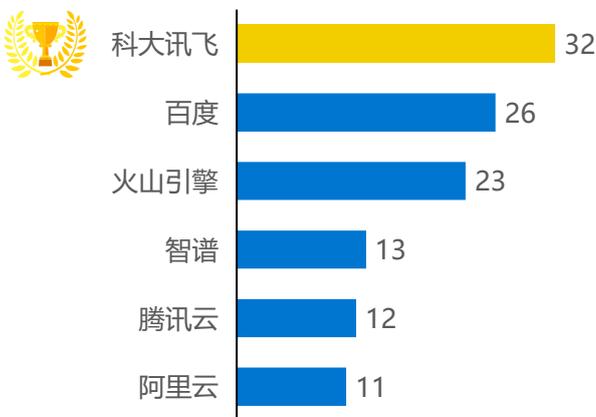
## 03

# 大模型企业典型案例分析

- 科大讯飞：G/B端深耕与抢滩AI终端构筑壁垒
- 字节跳动：数据飞轮与C端流量驱动的生态闭环
- 阿里：夸克的用户粘性策略与差异化突围

## 科大讯飞：深耕G/B端，巩固政企市场领先地位，并加强C端和抢滩AI终端，构筑竞争壁垒

- 科大讯飞提出“优选G端、做深B端、加强C端”的战略思路，以坚实的B端（企业）和G端（政府）市场为商业化主阵地，同时体系化布局C端应用与AI终端，构建软硬一体、覆盖多场景的协同生态。
- 政企市场的优势领跑。**大模型2C侧用户付费意愿较低，盈利模式单一，可持续的商业闭环尚未形成；相比之下，2B/2G领域凭借明确的刚性需求、强支付能力以及政策红利，成为大模型规模化变现的主战场。科大讯飞得益于强大的品牌公信力、全国产化的算力底座以及为客户提供软硬一体深度解决方案的能力，在大模型2B和2G商业化落地赛道，展现出显著领先优势，深度覆盖央国企客户群体。科大讯飞独创的“通专结合”模式能够充分融合一个通用大模型的强大底座技术能力，同时灵活调用几亿、几十亿乃至几百亿参数规模的大模型，组合完成复杂任务，为央国企及各行业提供高性价比的定制化解决方案。IDC数据显示，2024年中国央国企大模型解决方案市场规模已达31.8亿元，其中，科大讯飞市场份额居首位。《中国大模型中标项目监测与洞察报告(2025年Q3)》显示，科大讯飞的中标项目数量和中标金额在2025年第三季度依然实现双第一，领先地位持续巩固。



图示：Top6通用大模型厂商中标数量（个）排行榜（2025年第三季度）  
数据来源：智能超参数，36氪研究院整理

图示：Top6通用大模型厂商中标金额（亿元）排行榜（2025年第三季度）  
数据来源：智能超参数，36氪研究院整理

- C端应用的多元化探索与重点突破。** 在消费端，科大讯飞采取“广覆盖+深聚焦”的双轨策略，通过11个核心赛道27款应用的广泛布局，成功孵化出讯飞星火（MAU超1200万）、讯飞AI学、讯飞晓医三款百万级月活的标杆应用，形成了具备持续增长能力的C端产品矩阵。其中，讯飞星火APP端月活跃用户达1274.95万，在中国APP端AI应用排行榜中位列第7，跻身聊天类机器人赛道前五。2025年4月，讯飞星火X1全新升级发布，作为唯一的基于全国产算力训练的深度推理大模型，在数学、代码、逻辑推理等通用任务效果显著提升，整体效果对标OpenAI o1和DeepSeek R1。
- 抢占AI终端的生态蓝海。** 科大讯飞积极布局“大模型+机器人”等前沿领域，赋能机器人企业超500家，抢滩具身智能市场；同时作为AI硬件生态的领军者，通过办公本、学习机等硬件形成数据闭环和生态粘性。公司以智能语音技术和星火大模型等核心技术为基础，构建了涵盖教育、办公、翻译和音频等多个场景的硬件矩阵，持续创新功能，精准贴合用户需求。
- 科大讯飞在B/G端商业化上遥遥领先，同时通过星火大模型和硬件生态，在CBDG四个维度均占据了牢固的生态位。



图示：科大讯飞在CBDG四大维度均占据了牢固的生态位

## 字节跳动：全栈AI布局，场景赛道扫描，基于数据飞轮与C端流量驱动，打造生态闭环

- 作为行业后来者，字节跳动的AI战略框架日渐清晰：纵向，构建从底层硬件到上层应用的全栈能力，通过技术协同形成成本优化与体验升级的飞轮效应；横向，以C端超级App流量与B端企业服务为双轮驱动，搭建开放生态。其核心逻辑是并非追求单点技术极致，而是依托庞大的流量与产品矩阵广铺赛道，以数量和规模覆盖博取胜率。
- **全栈AI布局，多层次突破。** 算力层面，2025年计划在AI基础设施投入超120亿美元，积极建设海内外算力中心；基于自研硬件优化计算集群，性能较上一代最高提升三倍以上，为技术落地提供底层支撑。模型层面，最新开源的Seed-OSS-36B模型支持超长文本处理，创新引入推理效率优化机制，用户可灵活调整计算资源投入，在国际权威测试中取得优异成绩。多模态技术成为近期核心突破点：Waver1.0以统一架构实现文本、图像到视频/图像的生成，无需切换工具，重构内容创作流程；OmniHuman-1.5则实现“静态图像动态化”，仅凭一张人物照片和音频，就能让图像人物呈现说话、做手势等带情绪的动态效果，已在影视制作、教育培训等领域展现应用潜力。
- **场景赛道扫描，广铺多领域。** 从产品矩阵来看，字节的AI产品体系由豆包领衔，实现多场景覆盖，涵盖内容创作、社交互动和商业服务等多个热门领域。截至2025年4月，字节在双端推出超20款应用，布局10个赛道。这种策略旨在覆盖AI所有可能的应用方向，确保在任何一条赛道崛起时都不缺席。

### 纵向：全栈AI布局

构建从底层硬件到上层应用的全栈能力，实现算力、模型、多模态三层突破

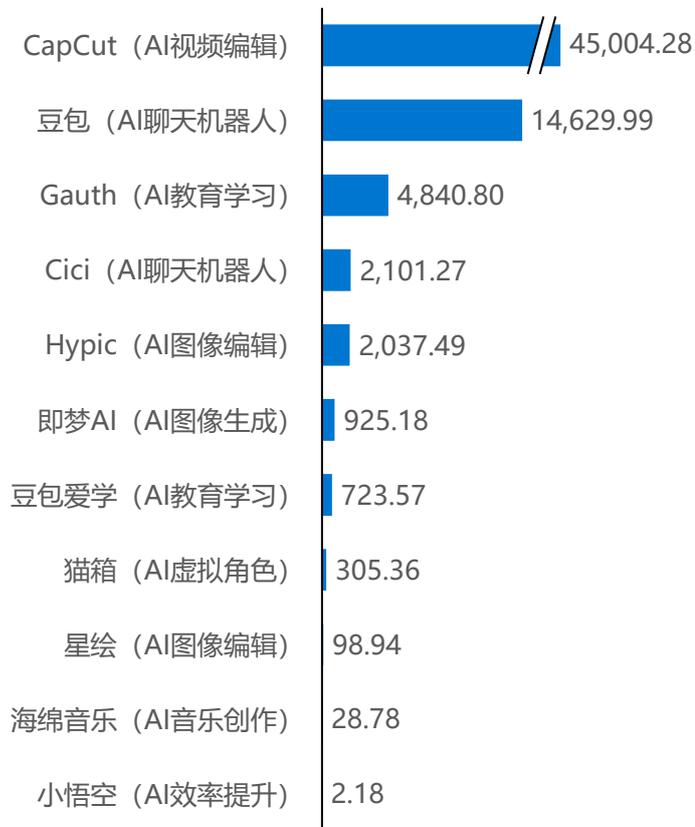
### 横向：赛道扫描全覆盖

AI产品体系由豆包领衔，实现多场景覆盖，确保在任何一条赛道崛起时都不缺席

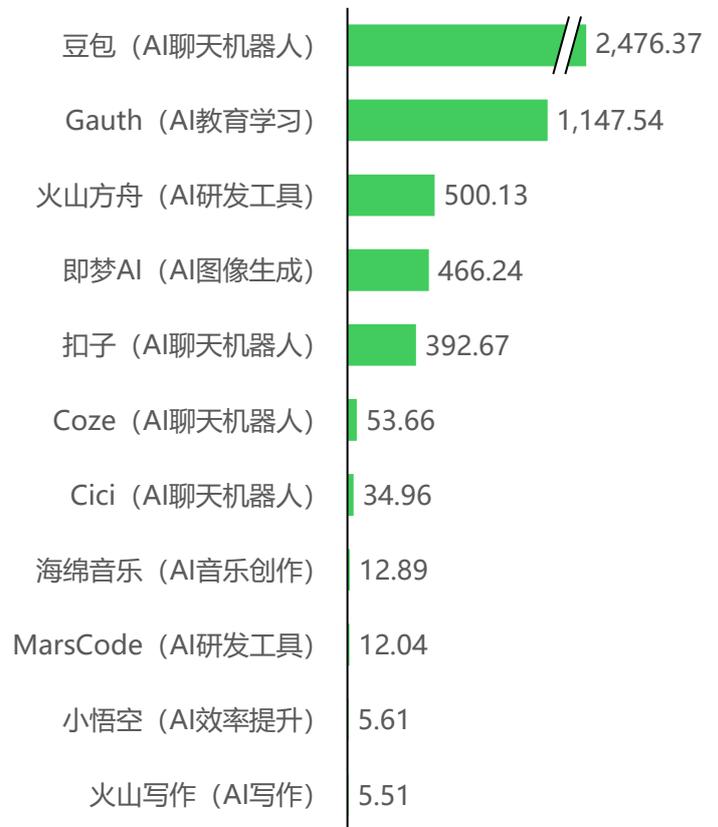
图示：字节的核心战略

- 数据飞轮，生态反哺。** 字节核心优势在于将抖音、今日头条等超级App的庞大用户流量与数据，转化为AI应用迭代的燃料，形成“产品-数据-模型-更好产品”的闭环。通过精准推荐和个性化定制，提供优质便捷的AI体验，成功打造了多款千万级活跃用户应用产品，验证了其流量驱动模式的可行性。
- 敏捷开发，速度制胜。** 作为活跃的行业领导者之一，字节强大的工程化能力支撑了应用的快速开发和上线，以速度应对市场的不确定性，展现了卓越的技术研发能力和用户需求把握能力。同时，字节也在积极拓展海外市场，将AI技术与产品推向全球。Sensor Tower数据显示，2025年Q3，豆包的海外姊妹版Cici在印度尼西亚、马来西亚、菲律宾、墨西哥和英国等市场，一直位列谷歌应用商店下载量前20名的免费应用。

#### AI APP产品



#### AI WEB产品



图示：字节跳动布局的AI应用产品、所属赛道和MAU（万）

数据来源：Xsignal, 36氪研究院整理

\*时间：2025年4月

## 阿里：将夸克作为竞争尖兵，向集成式AI工具集转型，提高用户粘性，寻求差异化突围

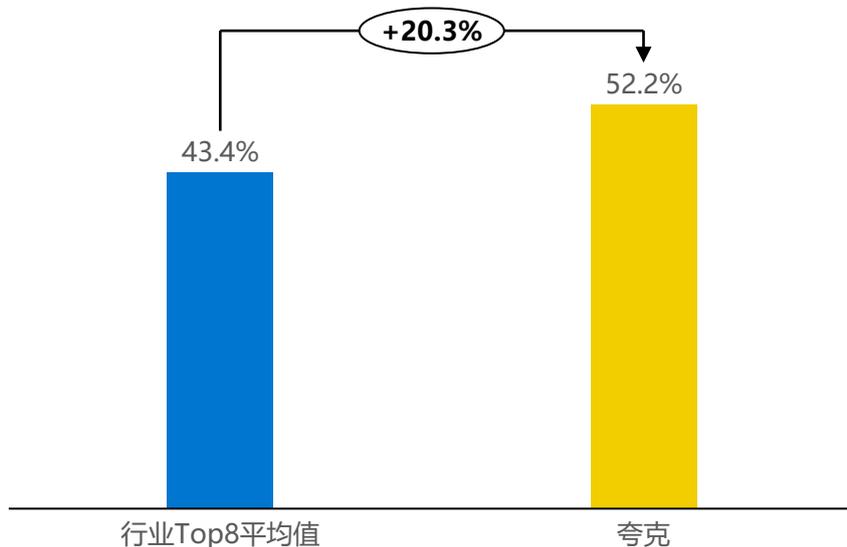
- 阿里规避与DeepSeek在通用大模型上的同质化正面竞争，而是将夸克作为尖兵，通过极致化的用户体验和功能聚合，打造高粘性用户体系，寻求差异化突围。
- 从搜索到“AI工具集”的转型。**在产品定位转型上，阿里顺应通用大模型技术普及与用户需求升级趋势，对夸克进行系统性重塑：摆脱传统搜索引擎作为信息检索工具的单一定位，以“AI超级框”为核心载体整合全链路能力，集成创作、学术、健康等高频场景服务，逐步升级为覆盖用户多元需求的全能型AI助手。
- 2025年9月，夸克进一步落地和升级该战略，上线全新AI创作平台“造点”，不仅集成Midjourney V7与通义万相Wan2.5双模型（双模型协同覆盖图像生成、创意设计、商业视觉等多元需求），还同步推出7天免费体验、会员价格减半的运营策略，大幅降低AIGC服务的使用门槛。从战略深意来看，“造点”的免费模式表面是抢占AIGC的用户心智，实则是阿里生态布局的关键一步。通过前端免费快速积累用户规模与场景化数据，如创作偏好、行业需求等，为后端会员增值服务、企业级定制方案、场景化广告等多元变现路径奠定基础，形成“用户+数据+商业”的闭环逻辑。

阶段	核心产品	主要特点	商业模式
初期	夸克浏览器	轻量、简洁、无广告	免费使用，基础服务
成长期	夸克App (AI超级框)	整合AI对话、超级搜索	部分功能付费，增值服务
扩张期	夸克扫描王、夸克学习	垂直领域深度工具	会员制，特定功能付费
生态期	“造点” AI创作平台	综合创意表达，双模型集成	免费体验+会员制，价格腰斩

图示：夸克AI产品矩阵的演变历程

数据来源：36氪研究院通过公开资料整理

- **用户粘性构建壁垒。** 阿里的竞争优势并非依赖单一的技术领先，而是通过“功能集成+场景渗透+生态协同”的组合策略，打造高用户依赖体系。据Xsignal统计数据，2025年2月，夸克AI的次日留存率达52.2%、用户人均单日使用时长33分钟，高于行业平均水平。这一数据优势的背后，是夸克通过“AI超级框”实现一次进入、多需求满足的体验设计，以及与阿里生态内电商、本地生活等服务的协同联动，大幅提升用户使用惯性与替换成本。
- **场景深耕替代技术空战。** 在场景落地层面，夸克以深耕刚需场景替代技术空战为路径，聚焦学习、办公、生活三大刚需场景，将自研模型能力转化为具体服务价值。在学习场景，其自研大模型支持精准搜题、知识点拆解与错题智能整理；在办公场景，整合AI写作、文档翻译、数据可视化工具，助力职场用户文档生成效率提升，支持多格式导出与跨平台协同。在生活场景，夸克围绕旅行规划、健康咨询、本地服务等核心需求提供结构化解决方案，同时整合场景资源，覆盖影视检索、听书听剧、网盘管理等高频功能，最终实现生活类需求“一个夸克全满足”的服务闭环。



图示：AI应用的次日留存率（月度活跃用户）

数据来源：Xsignal, 36氪研究院整理

\*注：次日留存率（月度活跃用户）的含义是：统计周期内，每日活跃用户数在第2日仍启动该AI产品的用户数占比的平均值

\*时间：2025年2月

## 04

# 大模型行业未来发展趋势

- 技术趋势
- 产业趋势
- 治理趋势

## 大模型推动人工智能范式从多模态认知向具身化迁移，并架设起通向未来超级智能的桥梁

- 随着通用能力持续增强，大模型将成为具备自主规划与跨场景执行能力的智能体基础。未来，中国大模型技术将沿着以下三条路径演进：
- 多模态融合深化，奠定智能认知的世界模型基础。**大模型将进一步打破模态边界，实现从文本、语音、图像、视频到三维空间的统一建模与语义对齐，构建能够刻画环境状态、推演物理规律、预测动态变化的世界模型体系。未来多模态不再停留于信息融合，而将演化为对真实世界的结构化、可计算化的认知特征，成为具身智能理解世界的基础。
- 智能体化演进，构建主动学习与协同决策的新范式。**大模型将从被动响应的认知工具进化为能自主规划、执行与反思的智能体（Agent），具备目标驱动的任务分解、资源调度与持续学习能力。智能体将成为AI生态的核心运行单元，推动人工智能由单体模型向多Agent协同系统演化。
- 具身智能突破，开启人工智能与物理世界融合的新阶段。**大模型将与传感、运动控制及机器人系统深度融合，赋予机器感知、决策与行动能力，实现认知、决策、行动一体化智能闭环。以大模型为核心的具身智能系统，将支撑人形机器人、服务机器人及其他智能载体在现实场景中完成复杂任务，通过与环境的交互持续修正认知、优化行为，实现智能体的自我进化与外部共生。



图示：中国大模型技术演进三条路径

参考资料：36氪研究院根据公开资料整理

## 大模型行业竞争由技术突破转向生态协同，价值重心加速向应用与服务层聚集

- 随着通用模型的能力差距趋于收敛，生态主导权与场景掌控力将成为产业竞争的核心。未来，中国大模型行业发展将呈现以下三大趋势：
- **价值链重构，推动行业从技术领先走向生态协同。**随着通用模型趋于标准化，企业竞争重心由技术突破转向生态整合与价值共创。产业链上下游将通过数据、算力、模型与应用的深度联动，构建以API服务、模型授权、模型即服务（MaaS）与行业模型定制为核心的新型价值体系。
- **生态体系演进，构建多层协同的产业网络。**头部企业巩固通用模型与基础框架优势，新兴主体强化行业模型、轻量化部署与私域定制，产业分化格局加速形成。随着生态协同深化，模型能力将通过API、Agent框架等方式嵌入行业系统，实现从模型赋能向系统融合的跃升。
- **应用纵深拓展，驱动大模型向垂直行业纵深渗透。**在制造、金融、教育、医疗、能源等行业推动智能化转型；同时，在消费端，AI终端、智能交互和内容生成将成为用户入口，带动智能服务加速普及。



### 价值链重构

- 通用模型标准化加快，产业价值链转向下游应用与服务
- 企业竞争重心由算法突破转向生态整合，通过开放接口、模型授权和生态共建实现价值共创



### 生态体系演进

- 产业生态由单一平台迈向“基础—模型—应用—服务”多层结构
- 通用模型厂商强化平台优势，行业主体深耕垂直领域，生态协同与分化并进



### 应用落地深化

- 制造、金融、医疗、教育、能源等行业加速融合大模型能力
- AI终端与智能交互成为新入口，带动智能服务普及与产业价值延伸

图示：中国大模型产业重构三大趋势

参考资料：36氪研究院根据公开资料整理

## 大模型生态竞合加速，中国将贡献以安全、可信与责任为核心的全球治理新范式

- 当大模型的应用规模化不断深入，有效的治理已成为支撑产业可持续发展的关键支柱。未来，中国在大模型治理领域将重点沿以下三大路径展开：
- **安全可信，构筑大模型治理的核心能力。**大模型在关键领域的渗透，使安全、透明与可控成为治理的底线要求。产业需加快建设贯穿数据、训练与部署全链条的治理工具与责任化体系，将安全与伦理从外部约束内化为模型的核心能力，实现由合规驱动向内生可信的转变。
- **中国路径，探索发展与安全的动态平衡。**中国正探索一条兼顾技术创新与可控安全的治理道路。以《关于深入实施“人工智能+”行动的意见》（国发〔2025〕11号）为纲领，确立“自主可控”与“安全可信”双重战略目标，未来治理将体现“发展-规范-安全”的三元统一，实现活力与秩序的共赢。
- **国际竞合，塑造规则制定的新竞争格局。**全球AI竞争已从技术制高点转向规则与标准的制定权。在此背景下，中国所倡导的“政策引导、企业实践、社会协同”三位一体的治理模式，不仅是本土实践的总结，更旨在为全球治理提供可鉴的方案，并塑造于我国有利的国际话语体系。



图示：中国特色大模型治理体系的三维发展路径

参考资料：36氪研究院根据公开资料整理

36氪研究院根据行业发展、资本热度、政策导向等定期输出高质量研究报告，研究方向覆盖人工智能、5G、区块链、医疗、金融、物流、文娱、消费、汽车、教育等多个领域，帮助政府、企业、投资机构等快速了解行业动态，把握发展机遇和明确发展方向。同时，研究院致力于为全国各级政府、企业、VC/PE机构、政府引导基金、孵化器/产业园区等提供专业定制化咨询服务，服务内容包括行业研究、产业规划、用户研究、股权投资研究、指数研究、投资配置、基金/企业尽调、战略规划、园区规划等。



## 01 传播资源

36氪已成为中国流量与影响力较大的互联网新商业媒体，覆盖全球超1.5亿读者，累计发布超过50,000篇文章，拥有行业内较显著的流量和传播优势。



## 02 核心数据

36氪深耕创投市场超9年，深度服务过数千家客户，累积超80万家企业和投资机构资源，搭建了完善的一级市场数据库。



## 03 团队优势

研究院现有数十位成员，主要来自国内外知名咨询机构或研究机构，拥有丰富的研究及项目经验。



## 04 研究领域

36氪研究院主要关注领域包括人工智能、5G、区块链、医疗、金融、物流、文娱、消费、汽车、教育等。



## 05 品牌影响

36氪研究院发布的常规性研究报告，受业内专业人士一致好评。在政府合作层面，研究院已为国务院、国家发改委、中央网信办、工信部、基金业协会、北京发改委、南京发改委、青岛高新区政府、湖南湘江新区管委会、成都新经济委、杭州西湖区政府等提供过咨询服务；在企业/投资机构合作方面，研究院已与大众中国、苏宁易购、携程、京东、网易、转转、字节跳动、海尔资本、洪泰基金、首钢基金等建立了深度合作关系。

## 分析师声明

作者具有专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

## 免责声明

36氪不会因为接收人接受本报告而将其视为客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在法律许可的情况下，36氪及其所属关联机构可能会持有报告中提到的公司的股权，也可能为这些公司提供或者争取提供筹资或财务顾问等相关服务。

本报告的信息来源于已公开的资料，36氪对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映36氪于发布本报告当日的判断，本报告所指的公司或投资标的的价格、价值及投资收入可升可跌。过往表现不应作为日后的表现依据。在不同时期，36氪可发出与本报告所载资料、意见及推测不一致的报告。36氪不保证本报告所含信息保持在最新状态。同时，36氪对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

# 36Kr

让一部分人先看到未来



研究咨询邮箱: [research@36kr.com](mailto:research@36kr.com)

研究院网站: <https://36kr.com/academe.html>