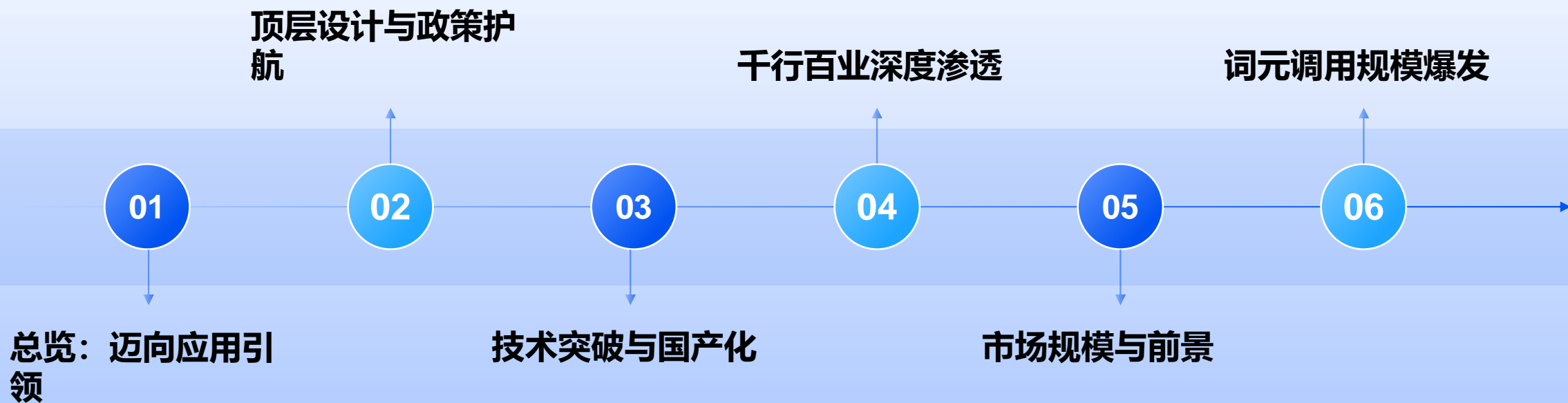


中国AI大模型产业黄金时代

规模化落地与全球引领 (2026)

北京研精毕智信息咨询有限公司

目录



总览：迈向应用引领

2026年，中国AI大模型产业迎来规模化落地的黄金时代，
凭借海量数据、庞大市场与政策支持，在应用活跃度上实现
对美国的超越，标志着产业重心从技术追赶转向应用引
领。

全球引领新纪元

中国AI产业实现从“追赶”到“引领”的关键转折

规模化落地元年

2026年是中国AI大模型产业全面进入规模化商业落地的关键年份，产业重心从技术验证转向价值创造。

01

日均调用量破140万亿

中国AI大模型日均词元调用量突破140万亿，占全球总调用量的60%以上，连续多月超越美国。

03

应用活跃度核心指标

词元（Token）调用量成为衡量AI应用活跃度的核心指标，反映了模型在实际场景中的使用频率和深度。

02

全球创新最快市场

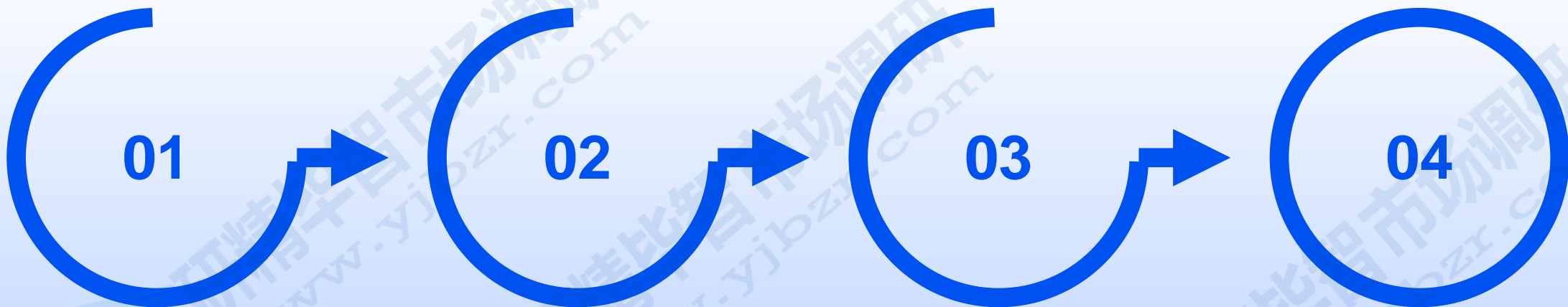
庞大的调用量和丰富的应用场景，使中国成为全球AI应用最活跃、创新迭代速度最快的市场。

04



产业重心转移

产业发展焦点从底层技术竞赛转向上层应用创新



技术验证期结束

全球AI产业已走过单纯追求参数规模的“参数竞赛期”和技术初步应用的“应用落地期”。

商业落地期开启

产业焦点全面转向商业化落地，核心目标是将技术转化为实际生产力，实现降本增效。

竞争焦点转移

行业竞争从“参数大小”转向“推理效率、成本控制、场景适配”，技术创新更注重系统性优化。

差异化优势形成

中国厂商通过“算法-算力-数据”协同优化，实现“性能不降、成本大降”，为规模化应用奠定基础。

顶层设计与政策护航

国家与地方层面形成“国家统筹、地方联动、伦理兜底”的政策体系，为产业爆发提供“真金白银”的支持与“精准高效”的服务，同时筑牢安全防线。

国家级战略支持

顶层设计为AI大模型产业发展划定“红线”与“底线”

01 顶层规划密集出台

《生成式人工智能服务管理暂行办法》《新一代人工智能发展规划（2026-2030年）》等文件，构建了全周期产业支持框架。

02 明确产业发展目标

将AI大模型定位为“新质生产力核心引擎”，提出到2030年技术与应用水平稳居世界前列的发展目标。

03 统一产业计量标准

国家数据局确立“词元（Token）”为AI大模型核心计量单位，发布《人工智能词元计量规范》，为产业标准化奠定基础。

地方特色政策

各地政府因地制宜，形成各具特色的政策支持体系

01

北京：创新策源地

设立100亿元AI产业基金，提供最高30%的算力补贴，开放政务、医疗、教育等重点场景。

02

上海：产业高地

建设张江人工智能岛等产业载体，对顶尖AI人才给予最高500万元安家补贴。

03

深圳：应用示范

依托数字经济优势，推出“AI大模型应用示范项目”，对入选项目给予最高2000万元奖励。

04

杭州：开源生态

聚焦开源生态建设，支持阿里、DeepSeek等企业构建开源社区，对开源模型给予年度运营补贴。

伦理与安全治理

完善的治理体系为产业长期健康发展奠定制度保障

伦理审查制度化

印发《人工智能科技伦理审查与服务办法》，明确审查重点，要求企业建立伦理审查委员会。

01

备案监管全流程

国家网信办建立“生成式AI服务备案管理平台”，已有860余家企业完成备案，形成“备案+监管”机制。

02

安全管理条例

《人工智能安全管理条例》草案公开征求意见，明确数据安全、模型评估、应急处置等要求。

03

防范潜在风险

政策体系旨在防范偏见歧视、算法压榨、数据泄露等风险，推动AI“安全、可控、可信”发展。

04

技术突破与国产化

中国AI厂商走出差异化技术路线，在效率革命与国产化替代两方面取得重大突破，通过算法优化与工程创新，显著降低推理成本，提升应用效率。

效率革命性突破

通过系列工程创新，实现推理成本大幅下降

1

算法优化核心技术

广泛采用模型蒸馏、量化压缩、动态路由 (MoE架构) 等技术, 显著提升推理效率。

2

算力调度智能化

通过分布式训练框架、异构算力融合等技术, 大幅提高算力利用率, 降低训练与推理成本。

3

数据治理精细化

构建高质量中文数据集, 优化数据清洗、标注、对齐流程, 提升模型训练效率与准确性。

4

推理成本降至1/5

综合技术优化使中国大模型推理成本降至海外同类产品的1/5-1/10, 为规模化应用扫清障碍。



关键技术新突破

2026年技术突破集中在架构、推理与可解释性



1

MoE分布式架构普及

阿里通义千问4.0采用第三代MoE架构，知识共享效率提升60%，推理速度提升3倍。

2

神经符号融合成熟

百度文心一言6.0采用“双脑架构”，复杂数学推理准确率提升18%，代码生成通过率提升25%。

3

可解释性能力提升

DeepSeek-V3推出“透明推理层”，实现推理轨迹追溯，推动模型从“黑箱”走向“透明”。

4

满足关键领域要求

技术突破有效满足了金融、医疗等关键领域对模型准确性、稳定性和合规性的严苛要求。

国产化替代加速

核心软硬件自主可控进程加快，市场占比显著提升

01

AI芯片性能跃升

华为昇腾910B、寒武纪思元370、壁仞BR100等国产芯片性能达到国际先进水平，打破算力垄断。

02

国产芯片市场占比

2026年国产AI芯片市场占比达35%，预计2028年将超过50%，实现对政务、金融等领域的批量替代。

03

框架生态主导地位

百度飞桨、华为MindSpore等国产框架占据国内市场主导，开发者数量超500万。

04

高质量中文数据集

国内厂商构建的海量中文数据集，数据质量与安全性显著提升，减少对海外数据的依赖。



千行百业深度渗透

AI大模型已从“可选工具”变为“必备能力”，在C端、B端和垂直行业全面渗透，成为推动数字经济与实体经济深度融合的核心引擎。

C端市场多元化

C端应用成为词元调用量最大的场景，渗透率高

1

智能对话成日常

豆包、元宝等智能对话应用成为用户信息查询、问题解答的日常工具，日均活跃用户超3亿。

2

内容创作降门槛

AI写作、AI绘画等应用降低创作门槛，覆盖自媒体、设计师等群体，激发个人创造力。

3

AI陪伴新需求

虚拟恋人、AI导师等应用满足用户情感与学习需求，用户日均使用时长超1.5小时。

4

C端调用占比45%

C端应用词元调用占比达45%，是词元调用量的核心来源，字节、阿里、腾讯占据主导。



B端市场降本增效

B端应用成为产业增长的“压舱石”，增速持续

1

办公自动化提效

AI实现会议纪要、邮件撰写等功能，企业办公效率提升40%，人力成本下降25%。

2

智能客服升级

智能客服处理70%以上的常规咨询，响应时间从分钟级缩短至秒级，客户满意度提升30%

3

研发设计加速

AI辅助代码生成、工业设计等，研发周期缩短30%-50%，研发成本下降20%-30%。

4

生产调度优化

AI优化生产排产、质量检测等流程，生产效率提升15%-20%，不良率下降10%-15%

垂直行业定制化

行业大模型成为垂直领域应用的关键，增速超通用模型

01

政务领域提效

某省会城市采用政务大模型后，行政审批效率提升60%，群众办事时长缩短70%。

02

金融风控升级

某国有银行通过金融大模型实现信贷风险识别准确率提升22%，不良贷款率下降0.8个百分点。

03

医疗诊断辅助

某三甲医院采用医疗大模型后，常见病诊断准确率提升15%，病历书写时间缩短80%。

04

工业智能制造

某汽车厂商采用工业大模型后，生产效率提升18%，制造成本下降12%。

05

垂直应用占比15%

垂直行业应用词元调用占比达15%，预计2028年将达30%，成为产业增长新引擎。

市场规模与前景

中国AI大模型市场已迈入高速增长的黄金期，预计到2030年，市场规模将达到3250亿元，成为全球AI产业增长的核心引擎。

市场高速增长轨迹

2023-2026年市场规模持续攀升，结构不断优化



2030年市场展望

产业进入稳定成熟期，成为数字经济核心基础设施



01

市场规模达3250亿

预计到2030年，中国AI大模型市场规模将增至3250亿元，进入稳定成熟期。

02

行业格局基本定型

头部5家企业（CR5）市场份额预计将超过60%，市场竞争从规模扩张转向价值提升。

03

成为新质生产力载体

AI大模型将成为数字经济核心基础设施，像电力、互联网一样融入生产生活，对GDP贡献超5%。



词元调用规模爆发

词元调用量作为核心“硬指标”，其爆发式增长直接反映了中国AI大模型在经济社会中的深度渗透与广泛应用，标志着产业从“实验室”走向“生产线”。

三年千倍增长奇迹

中国AI大模型词元调用量实现指数级跃升

01

2024年初：千亿级

日均词元调用量仅1000亿，主要来自C端对话应用，用户规模有限，调用行为碎片化。

02

2024年底：十万亿级

日均词元调用量达10万亿，同比增长100倍，B端场景放量，厂商策略转向“实用优先”。

03

2025年底：百万亿级

日均词元调用量跃升至100万亿，同比增长10倍，智能体框架开源引爆市场，单次调用词元量突破10万。

04

2026年：全球第一

日均词元调用量突破140万亿，占全球60%以上，连续多月超越美国，成为全球最大调用市场。



全球第一领跑优势

中国厂商在全球词元调用榜单中占据主导地位

01

调用量超美2.4倍

2026年4月，中国周词元调用量达7.942万亿，日均超1.13万亿，为美国的2.4倍。

02

垂直行业成主力

B端与垂直行业调用量占比达55%，标志着产业赋能成为词元增长的核心动力。

03

头部厂商主导全球

阿里通义千问、字节豆包、DeepSeek三家厂商词元调用量合计占全球40%。

04

单品调用量登顶

腾讯Hy3preview以2.66万亿Token登顶全球单品榜首，DeepSeek-V4-Flash以2.06万亿Token紧随其后。



应用深度广度彰显

海量词元调用背后是AI对生产生活的全面重塑

1

每人每天10万词元

相当于中国14亿人口每人每天占用10万个词元，AI已成为全民级生产工具。

2

等效46.7亿本书

140万亿词元若以每本书30万字计算，可等效为46.7亿本书，彰显内容处理能力。

3

填满1000个国家图书馆

打印成书可填满1000个国家图书馆，直观体现应用规模的宏大。

4

产业提质增效核心

在山钢集团，AI将5个工作日的报表统计压缩至实时完成；在政务领域，智能审批系统将办事时长从数天缩短至数小时。

分析师声明

负责本研究报告的分析师在本报告中所采用的数据均来自合规渠道，报告的观点、逻辑和论据均为分析师本人研究成果，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。

公司声明

本报告的著作权归北京精毕智信息咨询有限公司(简称为“研精毕智”)所有。本报告是研精毕智研究与统计成果，所载的观点、结论和建议仅代表行业基本状况，仅为市场及客户提供基本参考。

本报告调研方法主要是桌面研究、行业访谈等，结合公司内部逻辑算法，通过定量和定性分析分析，客观阐述行业的现状，科学预测行业未来的发展趋势。

我们力求报告内容客观、公正，但受到调研方法及调查资料收集范围的局限，本报告所述的观点、数据并不一定完全准确。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式篡改、复制和发布。如引用、转载需注明出处，且不得对本报告进行有悖原意的引用和修改。

本研究报告仅供北京研精毕智信息咨询有限公司客户和经本公司授权机构的客户使用，未经授权私自刊载的机构以及其阅读和使用者应慎重使用报告，本公司不承担由此所产生的相关风险和责任。