

大模型专有云服务市场分析

(2026 年)

中国信息通信研究院云计算与数字化研究所

2026年7月

版权声明

本报告版权属于中国信息通信研究院，并受法律保护。
转载、摘编或利用其他方式使用本报告文字或者观点的，
应注明“来源：中国信息通信研究院”。违反上述声明者，
本院将追究其相关法律责任。

前 言

以专属隔离的 AI 基础设施，打造资源独享、弹性按需的专有云推理服务是大模型推理服务市场的重要组成部分，已成为政务、金融、能源等行业大模型推理服务的主流承载方式，持续驱动数字经济与实体经济深度融合、深化 AI 应用场景落地。

近年来，在“人工智能+”行动上升为国家核心战略的时代背景下，大模型服务已从“技术创新验证”全面迈入“产业规模落地”的关键阶段。大模型专有云服务聚焦企业级非公有云场景下的大模型推理服务，服务商为机构提供自建或托管专用的智算资源，支持基础大模型本地化部署，构建起安全隔离、弹性扩展、深度定制的大模型推理服务体系。随着应用场景的持续深化、企业级场景下词元（Token）调用量的指数级增长，大模型专有云服务在技术架构、服务模式、生态体系持续迭代升级，进一步巩固了其在我国 AI 产业落地中的重要市场地位。

在此背景下，中国信息通信研究院云计算与数字化研究所基于在云计算、智能算力、大模型产业领域的长期研究积累，组织开展大模型专有云服务市场专项调研，覆盖百余家行业龙头企业、头部云服务商、大模型厂商及行业用户，编写形成本报告。本报告从大模型专有云服务的发展背景、市场现状、竞争格局、发展趋势、落地场景等维度剖析我国大模型专有云服务市场发展现状，梳理产业发展核心逻辑与演进规律，旨在为国家相关政策制定、企业市场布

局与选型决策提供数据支撑与参考依据。

本报告中相关测算数据，除特别标注外，均仅代表我单位作为科研机构的学术研究成果，仅供学习参考，不代表政府官方数据口径。

目 录

一、战略、产业与合规共同驱动大模型专有云服务市场发展.....	1
（一）政策牵引：国家战略顶层赋能，引导产业持续高速发展.....	1
（二）产业驱动：模型应用规模爆发，催生本地部署刚性需求.....	3
（三）合规约束：监管体系持续健全，筑牢产业规范发展底座.....	4
二、大模型专有云服务迈入规模化、精细化、价值化发展新格局.....	6
（一）规模增速领跑行业，服务能力跨越提升.....	6
（二）重点行业因地制宜，专有云部署形态分化成型.....	8
（三）供给能力持续完善，场景需求加速释放.....	11
三、全链协同构筑起产业竞争的新壁垒.....	12
（一）产业格局：全链协同深度融合，构筑产业竞争内核.....	13
（二）交付逻辑：从资源供给到场景交付，多元供给体系加速成型.....	15
（三）市场梯队：头部厂商全域领跑，行业份额高度集中.....	17
（四）主体策略：厂商布局差异赋能，形成多元竞争格局.....	18
四、政务、金融、能源行业典型落地实践场景.....	20
（一）政务行业：集约化部署标杆引领，打造安全可靠底座.....	21
（二）金融行业：高价值场景集聚赋能，构建合规应用范式.....	23
（三）能源行业：分级部署适配生产需求，赋能能源安全与绿色转型.....	24
五、大模型专有云服务未来发展趋势.....	26
（一）规模结构：市场体量持续扩容，专有云占比稳步攀升.....	27
（二）服务模式：智能体驱动服务跃迁，词元经营走向精细化.....	28
（三）行业竞争：竞争维度全面升级，全栈和国产化是核心方向.....	29
（四）落地应用：场景渗透纵深推进，垂直定制激活增长动能.....	31

图 目 录

图 1 大模型专有云“算力规模→日均词元调用量”测算模型.....	7
图 2 政务行业大模型服务形态分布.....	9
图 3 金融行业大模型服务形态分布.....	10
图 4 能源行业大模型服务形态分布.....	11

一、战略、产业与合规共同驱动大模型专有云服务市场发展

我国大模型专有云服务市场快速崛起，是国家战略引领、产业发展倒逼、合规监管驱动三重因素叠加的必然结果。政策层面顶层设计持续加码，明确产业发展方向；产业层面数智化转型提速，催生本地化部署刚需；合规层面监管体系日趋完善，划定安全发展边界，三大动力协同发力，推动专有云成为大模型产业落地的主流形态。

（一）政策牵引：国家战略顶层赋能，引导产业持续高速发展

国际层面，美国等发达国家布局“云+AI”赛道，抢占数字科技竞争制高点。美国持续以国家级战略强化全球云智产业主导权，发布《赢得人工智能竞赛：美国人工智能行动计划》《国家算力战略》等系列顶层文件，从政策层面巩固亚马逊、微软、谷歌等企业在全球云计算与 AI 服务市场的主导地位，通过“星际之门”万亿级算力基建计划，构建跨洲际算力互联网络。欧盟以“技术自主可控、数据安全合规、绿色低碳发展”为核心主线，依托正式生效的《欧盟人工智能法案》和《数字市场法案》划定 AI 服务与云平台运营的合规边界，升级“欧洲云计划（Gaia-X）”2.0 版本，推动泛欧主权云网络与一体化算力基础设施建设，着力破解对外部云服务的依赖风险。

国内层面，AI 云产业发展进入新阶段，巩固人工智能产业自主发展能力。我国 AI 云产业已从技术跟随、规模扩张的起步期，全面迈入技术自主、生态完善、场景深耕的高质量发展新阶段。国家通过深入贯彻“十五五”规划纲要关于全面实施“人工智能+”行动的战略部署，发布《关于深入实施“人工智能+”行动的意见》等纲领性文件，从政策层面构建起 AI 云产业底层硬件适配、云平台研发以及大模型训练推理、智能体应用的全栈技术突破和全链条推进体系，为推动人工智能与千行百业深度融合、产业自主发展筑牢了政策底座。

与全球主要经济体以公有云为企业级 AI 服务绝对主导的发展路径不同，我国立足数字安全底线与实体经济转型需求，形成了“公有云推动算力普惠与民营企业发展、专有云深度服务大型央国企与强监管行业”的双轨并行、特色化发展格局。一方面，公有云凭借轻量化、低成本、开箱即用的服务模式，不仅快速响应海量中小微企业的智能化诉求，更成为众多大型企业实现技术普惠的重要底座，极大降低了 AI 应用门槛，在营销、客服、办公等通用场景中推动人工智能的普惠化落地。另一方面，针对政务、金融、能源等关系国计民生的关键领域，以及大型央国企对安全合规和高性能定制的刚性要求，专有云构建起物理隔离或逻辑强隔离的算力与模型服务环境，支撑大模型的私有化精调、安全部署与持续迭代，在筑牢安全防线的前提下实现复杂场景的深度智能化。这种“公有云+专有云”

双轨并行的特色化格局，走出了一条兼顾效率与安全、普惠与深耕的中国式 AI 产业落地之路。

（二）产业驱动：模型应用规模爆发，催生本地部署刚性需求

我国云计算 IaaS 市场规模持续高速增长，专有智算资源部署成核心增长极。据相关研究机构统计数据，2025 年上半年中国公有云计算基础资源即服务（Infrastructure as a Service,IaaS）市场规模已突破 1200 亿元人民币，同比增长接近 20%；2025 年下半年市场规模达到 1085.96 亿人民币，同比增长 20.9%¹。据此测算，2025 年中国公有云 IaaS 市场规模约为 2300 - 2400 亿元人民币，智能算力服务需求成为增长的最主要因素。其中企业自建或定向托管专用的智算资源市场占比超过 28.8%。海量本地化、专属化、安全隔离的智算部署刚性需求持续赋能政务、金融、能源等强监管、高敏感行业，以及大型央国企的核心业务智能化升级。

大模型服务从训练转向推理，本地化推理服务成为主要增量。在云计算市场规模持续扩容的背后，以大模型推理服务为核心的智能算力服务需求，已成为拉动 AI 云市场高速增长的核心驱动因素。根据摩根士丹利预测，中国 AI 云市场 2024 至 2029 年复合年增长率（CAGR）预计高达 72%²。从产业服务结构来看，AI 产业发展重心

¹ IDC，《中国公有云服务市场跟踪报告》（China Public Cloud Services Market Tracker,2025H2），2026。

² 摩根士丹利(Morgan Stanley)，《中国 AI 路径：词元用量激增下的 AI 云变现》（China's AI Path: Monetizing Surging Token Use via AI Cloud），互联网行业研报，2026 年 3 月 16 日。

已从模型训练转向模型推理服务，推理算力服务成为智算服务的主要形态。IDC 预测，2025 年中国推理工作负载占比达 67%³。中国工程院邬贺铨院士提到：“2026 年全球推理服务占 AI 服务总负载 70%—80%，中国推理需求为训练的 8 倍⁴”。Gartner 预测至 2028 年推理与训练算力占比将达 3:1，中国甚至更高，可能达到 4:1 以上⁵。

词元调用量千倍增长，大模型专有云服务成为企业级 AI 落地核心载体。人工智能服务持续高速发展推动词元调用指数级增长，夯实推理服务落地基础。截至 2026 年 3 月，我国智能算力规模超过 1882EFLOPS⁶，人工智能基础设施底座持续夯实。与此同时，全国大模型词元调用需求呈几何级增长，国家数据局数据显示，全国日均词元调用量从 2024 年初 1000 亿攀升至 2026 年 3 月 140 万亿，两年增幅超千倍⁷，其中企业本地化部署词元调用占比持续提升，大模型专有云已然成为承接产业级 AI 应用落地的核心载体。

（三）合规约束：监管体系持续健全，筑牢产业规范发展底座

法律层面，我国数据与 AI 相关法律法规体系完善，划定专有云

³ 国际数据公司（IDC）与浪潮信息联合发布，《2025 年中国人工智能算力发展评估报告》，2025 年 2 月。

⁴ 邬贺铨，《智能体互联 云网算协同》主题演讲，2026 云网智联大会，2026 年 4 月（转引自《通信世界》报道，2026-04-16）。

⁵ 曾劲清（Gartner 高级研究总监），访谈转引自《大模型重点转向 业内预计 2027 年 AI 加速器出货量将超越 GPU》，新浪财经/《中国经营报》系转载，2025 年 6 月。原述：“预计到 2028 年，推理所占用的算力与训练所占用的算力比例将达到 3:1，在中国这一差距可能更大，至少达到 4:1 以上。”

⁶ 国务院新闻办公室，4 月 21 日新闻发布会上工业和信息化部副部长张云明表示，算力基础设施已成为驱动人工智能发展的关键底座。截至 3 月底，我国智能算力规模达 1882EFLOPS。。

⁷ 国家数据局，《刘烈宏出席国新办新闻发布会介绍第九届数字中国建设峰会有关情况并答记者问》，国家数据局官网，2026 年 3 月 24 日。

部署安全底线。随着我国数字安全法律体系持续完善，《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《关键信息基础设施安全保护条例》《生成式 AI 服务管理暂行办法》等法律法规颁布实施形成叠加效应，构建了覆盖数据全生命周期的监管体系，明确要求金融、政务、能源等关键信息基础设施运营者，以及处理大量个人信息和重要数据的行业主体，必须保障核心数据安全，实现“核心数据不出域、关键业务本地处理”。

监管层面，重点行业监管细则持续收紧，公有云模式难以适配合规要求，为专有云部署创造市场机遇。金融监管部门明确要求银行、证券、保险机构的核心业务数据与客户敏感信息必须实现本地存储、全程可追溯，严禁向第三方非可控主体泄露；政务领域出台专项政策，要求政务数据必须在政务专有云内闭环处理，严禁政务敏感数据上公有云；医疗行业明确患者诊疗数据、健康数据等敏感信息不得出境、不得跨域非授权传输，必须在院内或区域医疗专有云内完成处理。

国家战略的顶层引领、产业升级的内生驱动与合规监管的刚性约束三重力量深度交织、协同发力，共同构筑了我国大模型专有云服务市场发展的坚实基础与独特优势。在全球云智产业竞争日趋激烈的背景下，我国既通过公有云实现了人工智能技术的普惠普及，更依托专有云筑牢了国计民生关键领域的数字安全防线、支撑了实体经济的深度智能化转型。随着三者的持续协同深化，大模型专有

云服务将进一步巩固其作为企业级大模型落地核心载体的战略地位，成为推动我国人工智能产业自主发展、培育新质生产力的关键支撑力量。

二、大模型专有云服务迈入规模化、精细化、价值化发展新格局

在政策、产业、合规三重动力驱动下，大模型专有云市场迈入高速发展、结构优化、生态完善的高质量发展新阶段。整体市场规模跨越式增长，行业部署形态梯度分化，服务交付模式迭代升级，供需两端双向提质增效，产业呈现规模化、精细化、价值化发展新格局⁸。

（一）规模增速领跑行业，服务能力跨越提升

从产业规模来看，大模型专有云服务推理算力占据国内 AI 推理服务超三成份额，成为主要服务承载体。基于国家相关调研情况与行业统计数据分析，本报告建立了“资源规模→日均词元调用量”测算模型（下见图 1）。根据模型测算，截至 2026 年 2 月，中国企业自建智算云推理资源规模已超过 350EFlops，占推理资源总规模的 1/3 以上，能够支撑日均 Token 产能超 70 万亿，成为中国大模型推理服务的核心承载形态⁹。我国 AI 推理服务资源总规模已突破

⁸ 本章所引用的核心数据及相关分析结论，均基于团队于 2026 年 3 月开展的《企业专有云算力使用情况调查》定向问卷调查。本次调研共回收有效样本 120 余份。覆盖 30 余家省市级政务单位，20 余家头部银行、证券、保险等金融企业，15 家能源公司以及医疗、教育、工业制造、运营商等企事业单位。受限于抽样方式及样本规模，本章展示的调研数据仅反映特定受访群体在调研期间的反馈，可能与宏观全貌存在一定统计学偏差。相关数据与结论仅供行业交流与研究参考，不作为官方口径或商业投资决策依据。

⁹ 国际数据公司（IDC）《IDC×信通院联合重磅：十大真相颠覆认知，中国企业算力“短缺”背后，揭秘更

1065.3EFLOPS，其中大模型专有云推理服务占全国 AI 推理服务总规模的比重超 34.4%，对应智算规模约 366.46EFLOPS，较 2025 年初实现翻倍增长，已成为我国大模型推理服务的核心承载形态。



来源：中国信息通信研究院

图 1 大模型专有云“资源规模→日均词元调用量”测算模型

从词元调用来看，专有云承接企业级大模型推理需求，日均词元调用量占据产业核心份额。通过大模型专有云推理服务日均词元调用量 $=\left(\text{大模型专有云推理服务资源规模} \times \text{利用率} \times \text{MFU} \times 86400\right) / \left(2 \times \text{模型激活参数量}\right)$ 测算，结合头部云厂商非公有云业务披露数据、头部企业自建智算资源抽样监测数据验证推测，截至 2026 年 2 月，大模型专有云承接政务、金融、能源等关键行业的核心业务推理服务需求，日均词元调用量达到数十万亿。

从市场动能来看，呈现规模扩容与价值提升双轮驱动格局，多行业需求释放推动市场高质量增长。一方面，政务、金融、能源等核心行业的大模型专有云渗透率持续提升，教育、制造、医疗等新

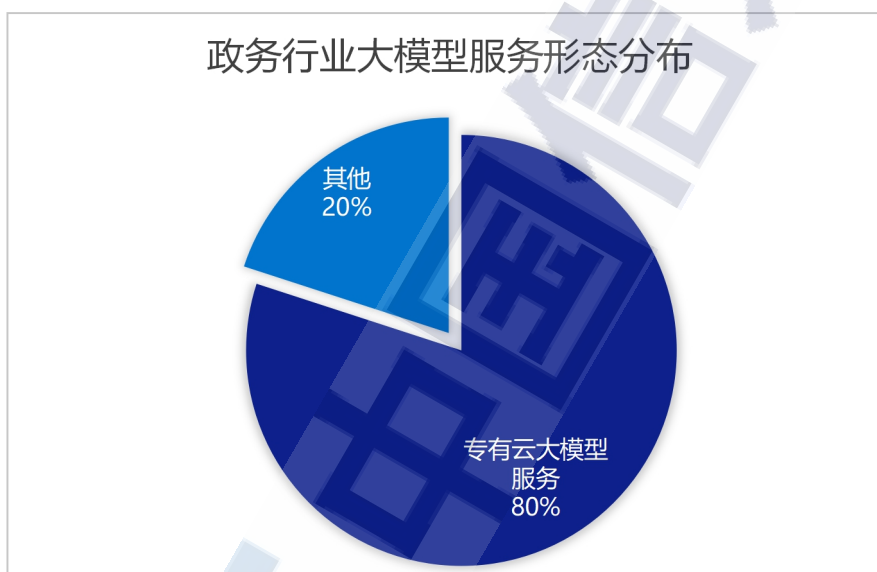
兴行业的部署需求快速释放，推动市场规模持续扩容；另一方面，大模型应用从知识问答、日常办公、公文处理等浅层场景，向金融风控、城市治理等核心业务场景扩展渗透，单个词元对应的商业价值密度持续提升，推动市场从“规模扩张”向“价值增长”转型。IDC 中国数据显示，2025 年千问大模型 Qwen、DeepSeek 等开源大模型掀起企业私有部署浪潮，在金融、医疗、能源等行业，82% 企业倾向私有部署大模型以满足数据安全要求，为市场增长提供了持续的内生动力。

（二）重点行业因地制宜，专有云部署形态分化成型

受行业监管规则、业务安全要求以及智算资源投入规模等核心因素驱动，各行业基于自身分工体系，在大模型专有云推理服务的部署形态上呈现出显著差异，共同构成了梯度协同的布局格局。政务行业由省属大数据公司或政务平台公司主导专有云建设，底层架构大多采用云服务商成熟解决方案。金融行业以自身私有云建设经验为依托，由科技子公司或内部 IT 团队开展专有云服务建设及运维。能源制造行业主要依靠企业数字化部门与云服务商深度协作，推动专有云部署。民生行业则在医疗、教育等领域的解决方案供应商与运营商的共同拓展下，大模型专有云的渗透率持续提升。

政务行业以托管专用部署模式为主，集约化建设模式成熟，实现安全、效率、成本协同统一。本次针对全国二十余家省市级政务单位开展的专项调研结果显示，超 80% 的省市级政务单位通过省属

大数据公司或政务平台公司托管专用资源部署大模型服务，依托云服务商专有云架构的强隔离特性与云化服务能力，同时兼顾了政务数据的安全合规要求与服务运维效率。从部署特征来看，政务行业形成了“省级统筹建设、地市分级接入”的专用部署模式，依托全省统一的政务专有云平台，实现大模型能力的全省复用与集约化运营，大幅降低了重复建设成本。

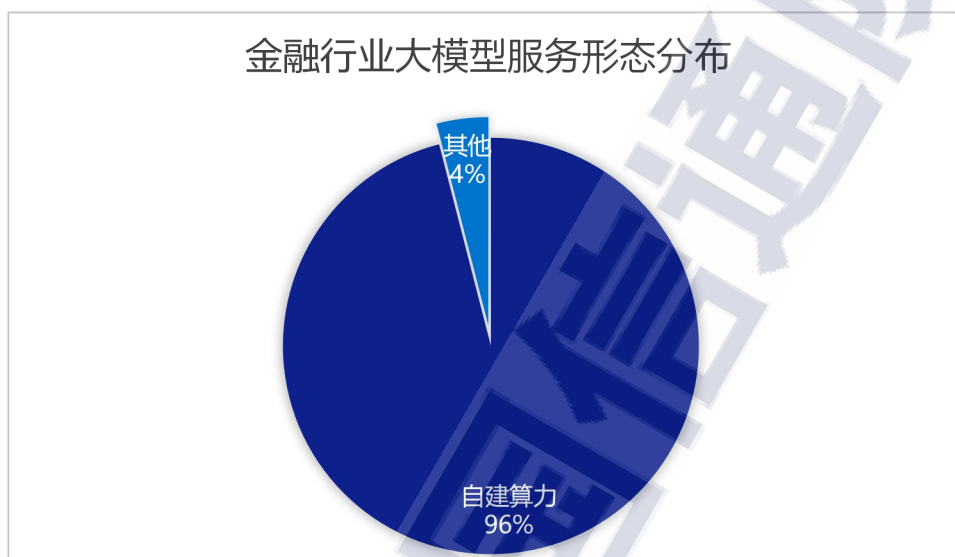


来源：中国信息通信研究院

图 2 政务行业大模型服务形态分布

金融行业适配不同规模机构的严监管合规需求，形成云服务商专属托管和自建专有资源池的混合部署格局。当前国有六大银行均已完成企业级自建智算资源布局，依托自建完全自主的物理隔离特性，充分满足金融核心业务场景的强监管与高安全要求，调研数据显示，金融行业自建智算资源规模占比达 96%（按智算规模口径，头部机构主导）。从细分业态来看，国有大型商业银行、全国性股

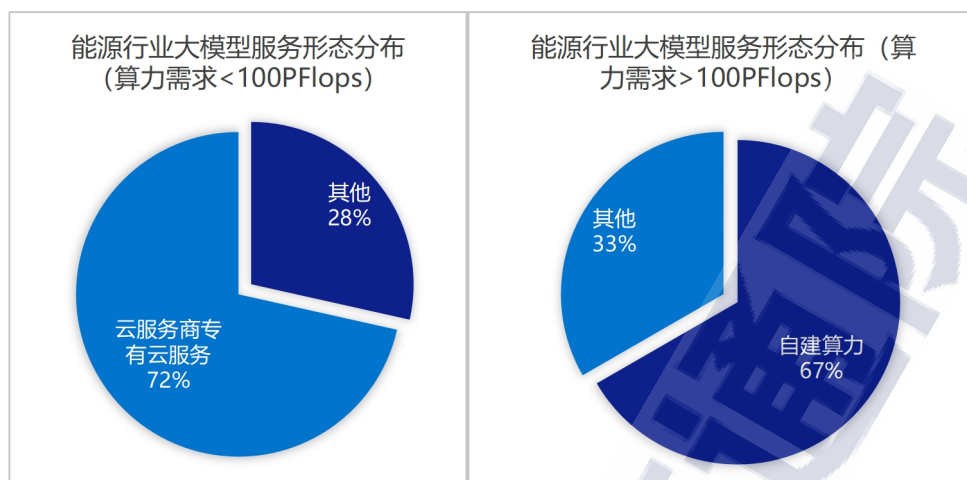
股份制银行以自建为主，头部券商、保险机构逐步形成“核心业务自建、非核心业务托管”的混合部署模式，城商行、农商行等中小金融机构则更多采用行业托管专用模式，降低部署门槛与运维成本。



来源：中国信息通信研究院

图 3 金融行业大模型服务形态分布

能源行业大模型专有云服务以采购云服务商专有智算服务为主，自建专属资源池为辅，分级适配核心与通用场景，平衡安全与效率。行业主体普遍形成了分级化、场景化的服务灵活部署模式：针对单主体累计超过 100PFlops 的智算资源需求、核心生产业务场景，普遍优先采用自建部署模式，保障生产数据安全性与业务连续性；针对 100PFlops 以下的小规模智算资源需求、非核心办公与管理场景，以采购云服务商专有云智算服务为行业主流选择，兼顾成本与效率。



来源：中国信息通信研究院

图 4 能源行业大模型服务形态分布

医疗、教育、智能制造等民生行业专有云渗透率持续提升，持续拓宽市场覆盖边界。医疗行业围绕电子病历分析、医学影像辅助诊断、医院智慧管理等场景，加快区域医疗专有云建设。教育行业聚焦智慧教学、科研计算、校园智慧治理等需求，推动教育专有云的规模化落地。智能制造领域则依托工业互联网平台，面向质量检测、设备预测性维护、工艺参数优化和生产调度协同等关键场景，加快建设制造专有云。这些民生行业的专有云加速渗透，正成为大模型专有云服务市场的重要增量来源。

（三）供给能力持续完善，场景需求加速释放

从供给侧看，云服务提供商正积极构建智算资源、模型、方案、生态一体化体系，以全面夯实专有云服务支撑能力。智算资源底座层面，国产智算芯片、服务器、智算中心建设持续突破，为专有云服务提供了坚实的智算资源支撑；模型能力层面，开源模型领跑专

有云市场，形成了以千问大模型 Qwen、DeepSeek、文心一言、科大讯飞、智谱、Kimi 等为核心的模型供给体系；场景方案层面，服务商已沉淀出覆盖政务、金融、能源、制造等数十个行业的标准化解决方案，实现了大模型能力与行业业务场景的深度融合；生态服务层面，已形成云服务商、大模型厂商、芯片厂商、ISV 服务商、行业客户协同发展的产业生态，大幅提升了专有云服务的落地效率。

从需求侧看，全域场景持续渗透，实体经济行业成为核心增量，供需双向赋能形成良性循环。行业场景化需求加速深度释放，形成了“核心行业引领、全行业覆盖”的需求格局。政务、金融、能源三大核心行业是专有云市场的核心支柱；制造行业增速最快，2025 年市场规模同比增速超 150%，成为市场核心增长极；医疗、教育、交通、物流等行业的专有云部署需求持续释放，推动市场需求结构不断优化。从场景深度来看，大模型应用已从智能客服、公文起草、内容生成等浅层通用场景，向政务城市治理、金融信贷风控、工业生产调度、医疗辅助诊断等核心业务场景深度渗透，场景复杂度持续提升，对专有云服务的定制化、专业化要求不断提高，形成了需求升级与供给能力提升的正向循环。

三、全链协同构筑起产业竞争的新壁垒

我国大模型专有云服务市场已进入体系化竞争的全新阶段，市场竞争从单一的智算价格、模型参数比拼，升级为全链路能力、全生态布局的综合竞争。产业链各环节协同持续深化，市场集中度持

续提升，头部厂商凭借全栈能力形成显著竞争优势，不同市场主体基于自身资源禀赋形成了差异化的竞争策略，整体呈现出“头部集中、梯队分化、多元赛道协同发展”的竞争格局。

（一）产业格局：全链协同深度融合，构筑产业竞争内核

大模型专有云服务产业链覆盖基础设施供给层（词元生产）、模型能力与服务供给层（词元流通）、行业应用需求层（词元应用）三大环节，产业链各环节边界持续模糊，全链路协同、芯云模体一体化成为产业发展的核心竞争逻辑，具备全栈自主能力的厂商在市场竞争中占据显著优势。

基础设施层是产业硬件底座，当前呈现国际引领、国产加速迭代的发展格局。基础设施供给层主要包括智算芯片、服务器、数据中心、网络设备等硬件设施供给商，以及操作系统、云平台、数据库、虚拟化软件等基础软件厂商。该环节决定了专有云服务的底层智算服务性能、稳定性与成本控制能力，国内市场已形成“国际厂商主导高端市场、国产厂商加速渗透”的格局，华为昇腾、平头哥真武、海光 DCU 等国产芯片已在政务、特定行业场景取得应用突破，浪潮、新华三、华为等厂商在服务器市场占据领先地位，万国数据、世纪互联、三大运营商等企业在数据中心领域形成了规模化布局。

模型与服务层是产业价值核心，四类市场主体依托差异化优势错位发展。服务与模型能力供给层是市场竞争的核心赛道，主要包括云服务商、大模型厂商、运营商、第三方专有云服务商四类主体。

该环节的核心价值在于打通底层智算与上层应用，通过智算调度、模型适配、场景化开发、全生命周期运维，为下游客户提供一站式专有云服务，是产业链的价值中枢。其中，头部云服务商凭借“智算基础设施+专有云平台+自研大模型”的全栈能力，占据市场主导地位；大模型厂商以模型能力为核心，轻量化切入专有云市场；运营商依托网络资源、政务渠道与本地化服务能力，在下沉市场形成差异化优势；第三方服务商则聚焦垂直行业定制化服务，与头部厂商形成生态互补。

行业需求层主导市场迭代方向，客户综合需求升级驱动服务商能力完善。行业应用需求层主要包括政务、金融、能源、制造、医疗、教育等行业客户，以及互联网、中小企业等其他市场主体。该环节决定了市场的需求结构与发展方向，其中政务、金融行业是市场的核心需求方；能源、制造等实体经济行业是市场增长最快的需求领域；医疗、教育等民生行业需求持续释放，推动市场需求结构持续优化。下游客户的需求从基础的智算资源，向全栈式、场景化、可落地的行业解决方案升级，倒逼中游服务商持续提升全链路服务能力。

整体来看，产业链迈入全链路整合协同阶段，全栈优化能力成为厂商核心竞争壁垒。头部厂商通过纵向整合，实现了从芯片适配、智算资源调度、模型研发到场景落地的全链条自主可控，全栈协同能力已成为市场竞争的核心逻辑。通过实现“基础设施+云平台+大

模型+行业解决方案”的全链路打通，为客户提供安全合规、高效稳定、高性价比的专有云服务，才能在市场竞争中构建长期壁垒。

（二）交付逻辑：从资源供给到场景交付，多元供给体系加速成型

当前阶段，大模型专有云服务的服务模式正经历深刻变革，行业已逐步告别传统 IDC “机柜租赁”、云计算“资源分时租赁”的粗放模式，全面从“资源供给”向“场景交付”转型，从“硬件售卖”向“全生命周期服务运营”升级，形成了多元化、精细化、场景化的服务交付体系。

一是资源交付模式升级，从标准化供给转向场景化定制，提升资源适配与利用效率。传统私有云部署以固定规格的服务器、智算集群租赁为核心，交付重心在于硬件资源的可用性；当前专有云服务已实现针对行业场景的深度定制，针对政务公文处理、金融长文本合规审查、工业视觉质检等不同场景的智算需求，定制化适配 GPU/NPU 异构智算集群，优化单机架功率、网络带宽、存储架构，实现智算资源与业务场景的精准匹配，大幅提升资源利用效率。

二是平台服务边界延伸，实现从单一部署到全生命周期运维、自主迭代的一体化赋能。头部服务商已不再局限于为客户完成大模型的本地化部署，而是提供覆盖模型微调、场景适配、推理优化、运维监控、安全合规管控的全流程平台化服务。通过构建专有云大模型服务平台，为客户提供低代码/无代码的模型微调工具、场景化

应用开发框架、全链路可观测性监控系统，让客户无需专业的 AI 技术团队，即可完成大模型的行业化适配与应用落地，大幅降低了大模型专有部署的技术门槛。

三是成果交付逻辑重构，转向任务式完成履约计费，实现服务商与客户业务目标深度绑定。行业服务重心从“智算资源租赁时长”向“业务任务完成效果”转移，计费模式从传统的按卡时、按智算规格计费，向按词元消耗量、按任务完成度、按业务价值贡献计费升级。例如，针对政务 12345 热线场景，按有效咨询处理量计费；针对金融风控场景，按风险识别准确率与不良率降低效果计费；针对工业质检场景，按缺陷检出率与生产效率提升效果计费，真正实现了服务价值与客户业务目标的深度绑定，推动专有云服务从成本中心向价值中心转型。

四是开启词元资产化运营范式，依托闭环环境沉淀数字资产，放大长期服务价值。行业已初步开启从“词元调用”到“词元经营”的范式升级，头部服务商通过专有云环境，帮助企业系统性沉淀两类核心资产：一是 **Context** 资产，即企业专属的场景图谱，包含业务规则、客户偏好、历史经验；二是 **Memory** 资产，即决策轨迹记忆，包含成功路径、试错经验、用户反馈。让每一次词元调用都转化为企业可积累的业务资产，形成 AI 能力的正向复利飞轮，进一步强化了专有云服务的长期价值。这一范式的系统化成熟与精细化运营，仍有赖于专有云服务能力的持续深化。

（三）市场梯队：头部厂商全域领跑，行业份额高度集中

当前大模型专有云服务市场呈现出显著的头部集中效应，市场份额向具备全栈能力的头部厂商快速聚集。

一是领航者梯队，具备全栈自主能力与全域生态优势，牢牢占据市场主导地位。该梯队厂商的核心特征是具备完整的全栈技术能力，实现了从底层智算基础设施、专有云操作系统、大模型研发到行业解决方案的全链路自主研发，拥有全国化一体化的算力布局、完善的生态体系与海量的头部客户案例。从市场表现来看，该梯队厂商覆盖了政务、金融、能源、工业制造等专有云大模型应用的全部核心行业，是国内专有云推理场景中覆盖行业最广、适配部署形态最全、政企二次开发活跃度最高的服务商体系。

二是拓局者梯队，依托单点优势深耕细分赛道，持续向全栈服务模式升级突围。该梯队厂商的核心特征是在细分赛道形成了差异化的核心优势，或在模型能力或在特定行业或在区域市场构建了自身的竞争壁垒，是市场的重要参与者。该梯队厂商虽未形成全市场的主导能力，但在细分赛道实现了对第一梯队的追赶。其中，云厂商依托自身互联网生态与大模型技术积累，在互联网、政务、金融等行业形成了稳定的客户群体；大模型厂商凭借开源模型的生态优势，在中小企业私有部署市场实现快速破局，在开源模型专有部署领域占据重要市场份额。该梯队厂商在细分赛道持续巩固差异化优势，逐步向全栈市场拓展。

三是竞争者梯队，聚焦细分市场，依托本地化服务与生态合作实现生态补位。该梯队厂商的核心特征是聚焦单一区域市场或特定垂直行业，通过灵活的定制化服务、本地化响应能力、高性价比的产品，在细分市场占据一席之地。这类厂商虽在技术深度、布局广度、规模效应上与前两大梯队存在显著差距，但凭借对区域市场、垂直行业的深度理解，能够为客户提供高度定制化的专有云服务。面对头部厂商下沉市场带来的竞争压力，该梯队厂商需持续挖掘细分领域潜力，强化与头部厂商的生态合作，巩固自身生存空间。

（四）主体策略：厂商布局差异赋能，形成多元竞争格局

不同市场主体基于自身资源禀赋、技术积累与市场定位，采取了差异化的竞争策略，形成了“全栈生态主导、垂直赛道深耕、区域市场聚焦”的多元竞争赛道，各类主体优势互补、协同发展，共同构建了大模型专有云服务市场的完整生态。

头部全栈云厂商构建端到端技术壁垒，凭借综合优势持续领跑市场。以阿里云为代表的头部云厂商，核心竞争策略是构建“自研芯片+专有云平台+原生大模型+行业解决方案”的全栈技术体系，通过端到端软硬一体协同优化，为客户提供安全合规、高效稳定、高性价比的专有云服务。这类厂商通过全栈优化，实现了专有云部署下词元产出效能的显著提升，根据调查问卷数据及测试结，“全栈云服务+企业自研模型”的组合部署形态，对比传统的“自选模型+

自选云服务厂商”混合部署形态，词元产出效能提升超过 20%¹⁰。同时，这类厂商通过开放平台接口横向整合第三方技术，打造一站式智算解决方案，构建了完善的产业生态，通过生态协同强化市场壁垒，牢牢占据市场主导地位。

运营商云厂商依托算网融合与渠道优势，深耕政务及下沉市场实现差异化竞争。以天翼云、移动云、联通云为代表的运营商云厂商，核心竞争策略是依托自身骨干传输网络、全国本地化服务网点、政务渠道资源优势，聚焦政务市场、县域下沉市场，打造“云网融合”的专有云服务。这类厂商通过“算网大脑”“息壤”等智算调度平台，纳管全网智算资源，实现云网边端一体化的智算资源调度，为政务、中小城市客户提供本地化、高安全的专有云服务。同时，运营商依托自身国资背景，深度适配政务、国企客户的国产、合规需求，在政务专有云、国企私有云市场实现了快速扩张，市场份额持续提升。

专业大模型厂商凭借技术与开源生态优势，轻量化切入市场、快速拓展细分赛道。以千问大模型 Qwen、DeepSeek、智谱 AI 为代表的大模型厂商，核心竞争策略是依托开源模型的生态优势、模型性能的技术优势，以轻量化部署模式切入专有云市场，通过与云厂商、硬件厂商的生态合作，实现规模化落地。其中，千问大模型 Qwen 凭借断层式的开源生态优势，截至 2026 年 1 月初，系列模型累计下

¹⁰ 词元产出效能：相同场景下单位算力日均 Token 调用量。

载量已突破 7 亿次，累计开放 300+ 开源模型，衍生的二次模型超过 17 万个，形成了全球最大的开源模型生态集群，为其专有云部署奠定了坚实的开发者基础；DeepSeek、智谱 AI 等厂商则通过模型性能优化、垂直行业微调，在金融、医疗等行业形成了差异化的模型优势，快速拓展私有部署市场。

垂直行业服务商深耕行业经验，错位竞争，成为产业生态的重要补充力量。以行业 ISV、区域型专有云服务商为代表的第三方厂商，核心竞争策略是避开与头部厂商的正面竞争，聚焦特定垂直行业或区域市场，通过深度的行业积累，为客户提供高度定制化的专有云部署、行业化模型微调、场景化应用开发全流程服务。这类厂商通常与头部云厂商、大模型厂商达成生态合作，基于头部厂商的智算底座与模型能力，做行业化的二次开发与本地化服务，在中小金融机构、地方制造企业、县域政务等细分市场形成了错位竞争优势，成为市场生态的重要补充。

四、政务、金融、能源行业典型落地实践场景

大模型专有云服务的价值最终通过行业场景落地实现，政务、金融作为强合规、高敏感行业，是专有云部署最成熟、渗透率最高的两大核心赛道，也是专有云服务模式创新、技术迭代的核心试验场。两大行业基于自身的监管要求、业务特征，形成了差异化的部署模式、场景结构与发展特征，成为引领全行业专有云服务发展的标杆。

（一）政务行业：集约化部署标杆引领，打造安全可靠底座

政务行业为专有云最成熟赛道，高安全合规属性锁定专有云主流部署形态。政务数据涉及国家秘密、公共安全、民生保障等核心敏感信息，具有极强的安全合规要求，“数据不出域、业务全闭环、自主可控”是政务大模型部署的首要原则，专有云凭借强隔离、高安全、云化运维、集约化运营的核心优势，成为政务大模型落地的主流模式。

从部署渗透率看，政务行业专有云渗透率领跑全行业，省级统筹集约化部署模式降本增效、合规可控。当前，政务专有云已成为数字政府建设的核心基座，渗透率显著领先于金融、能源等强合规行业。在建设模式上，各省普遍采用“全省一朵云”的统管架构，以省级平台为枢纽实现计算、存储、安全能力的统一纳管与弹性供给，地市及区县以逻辑租户形式按需接入，避免各单位重复建设机房与平台，显著节约财政资金。集约化运营使得安全策略统一配置、威胁情报全域共享、系统版本集中升级，极大缩小了基层的运维盲区与防护短板。同时，该模式遵循“数据不出域、本地化留存”原则，通过专有云物理或逻辑隔离环境确保全链路数据闭环，辅以分级分权审计，实现全域操作可追溯、可问责，在降低建设与运维成本的同时，牢牢守住了政务数据安全合规的红线。

从市场规模看，政务专有云市场规模高速增长，国产开源模型成为行业主流选型。政务行业大模型专有云服务市场规模 2025 年实现翻倍增长，省级政务专有云平台实现全覆盖，地市级政务专有云部署率超 70%，区县政务单位接入率持续提升。从模型选型来看，政务行业专有云部署高度集中于千问大模型 Qwen 等国产开源大模型，其中千问大模型 Qwen 在政务领域的部署渗透率已超过 40%，成为当前政务大模型落地的主要方案之一。

从落地场景看，政务专有云形成三大核心场景，全面赋能政务服务与城市治理现代化。政务行业大模型专有云落地已形成智能问答与咨询、政务服务热线智能处理、公文起草与城市治理三大高频核心场景，覆盖政务服务、办公协同、城市治理全流程。智能问答与咨询场景，是政务行业落地最成熟、调用量最高的场景，为企业和群众提供政策咨询、办事流程指引、申请材料预审等服务。单轮政务咨询问答平均词元消耗量达 600-1000，复杂政策解读场景单轮词元消耗量可达 2000 以上，是政务行业词元消耗的第一大来源。政务服务热线智能处理场景，是政务大模型落地的核心刚需场景，通过专有云部署的大模型，实现热线来电的智能分类、诉求自动派单、办理结果智能回访、投诉内容智能分析，大幅提升热线办理效率与群众满意度。公文起草与城市治理场景，是政务大模型向核心办公场景渗透的关键方向。通过大模型服务实现政策文稿起草、跨部门公文流转审核、会议纪要生成、舆情信息分析、城市运行态势研判

等功能，大幅提升政务办公效率。同时，基于城市运行海量数据，通过专有云大模型实现交通调度、应急管理、生态环保等城市治理场景的智能决策，推动城市治理能力现代化。

（二）金融行业：高价值场景集聚赋能，构建合规应用范式

金融行业是专有云高价值核心赛道，强监管属性推动专有云闭环部署模式普及。金融行业受最严格的数据安全监管约束，客户信息、交易数据、信贷数据等核心业务数据具有极高的敏感性，监管明确要求金融核心数据必须实现本地存储、全程可追溯、不出域管控，是专有云服务市场的核心价值高地。

从部署特征来看，金融行业形成大小机构差异化梯度部署格局，平衡合规要求与成本投入。国有大型商业银行、全国性股份制银行、头部券商、保险集团，均倾向开展企业级私有云智算资源布局，依托完全自主的物理隔离私有云环境，部署大模型服务，满足核心业务的强监管要求；城商行、农商行、中小保险机构等中小金融机构，受限于资金投入、技术能力，更多采用金融行业专有云模式，通过区域金融云、行业专属云实现大模型部署，兼顾合规要求与成本控制。

从竞争格局来看，金融专有云市场分层竞争格局清晰，头部厂商主导高端市场、多元主体协同补位。国有大行、股份制银行的企业级私有云市场，主要由阿里云、华为云等具备全栈能力的头部厂

商主导，其中千问大模型 Qwen 在金融行业的市场份额持续领先，在国有大行、股份制银行的私有云部署中实现了规模化落地；城商行、农商行等中小金融机构市场，由运营商云、区域金融云服务商与头部厂商协同布局；证券、保险等细分业态，形成了垂直行业解决方案厂商与头部云厂商互补发展的格局。

场景适配、效能优化、合规安全构成金融专有云三大核心竞争壁垒。金融行业专有云服务的核心竞争优势，集中于“金融场景深度适配能力”“全栈优化的词元效能”与“金融级安全合规能力”三大维度。头部厂商通过金融专属语料对大模型进行深度微调，使其精准适配金融风控、合规审查、财富管理等专业场景的需求，同时通过“私有云基础设施+大模型”的全栈协同优化，大幅提升单位智算词元产出效能，降低金融机构的智算投入成本，同时有效满足金融行业的强监管、高安全要求。

（三）能源行业：分级部署适配生产需求，赋能能源安全与绿色转型

能源行业作为国家关键基础设施与实体经济的重要代表，正加速大模型专有云服务的规模化落地。行业主体普遍形成“分级适配、场景驱动”的部署范式，成为驱动专有云服务市场高速增长的关键引擎。

组织结构上，能源行业形成了“集团统筹规划、分级分类部署、云边端协同”的特色部署模式。为了兼顾生产安全、运行效率与成

本控制，行业主体普遍遵循“核心业务自建、非核心业务托管”的分级原则。**从架构特征来看**，能源行业专有云普遍采用“集团总部云+区域分公司云+厂站边缘节点”的三级云边协同架构。总部云承载全局调度、大数据分析、模型训练等非实时任务，厂站边缘节点承载实时生产控制、设备巡检等低时延任务，专有云作为核心枢纽实现算力与数据的分级调度与协同。同时，能源行业作为国产要求的重点领域，对全栈国产适配要求极高，基于国产芯片、国产操作系统、国产大模型的专有云解决方案已成为新建项目的主流选择。

市场格局方面，能源行业专有云市场呈现“头部云厂商主导、运营商协同、行业独立服务商（ISV）深度参与”的竞争格局。头部全栈云厂商凭借全栈技术能力、国产适配能力与大型项目交付经验，与国家电网、南方电网、中石油、中石化等能源央企达成深度战略合作，占据了能源专有云市场的主要份额。运营商依托云网融合优势与边缘节点覆盖能力，重点布局地市级能源企业与厂站边缘侧场景。能源行业 ISV 则凭借深厚的行业经验，聚焦细分领域的定制化开发与本地化服务，与头部云厂商形成生态互补。

当前，能源行业大模型专有云服务已从初期的办公辅助场景，全面渗透至生产调度、设备运维、安全管控、碳资产管理等核心业务环节，形成智能生产调度、设备预测性维护、智能巡检与安全管控、能源交易与碳资产管理等高频高价值落地场景，有效支撑了能源行业的安全稳定运行与绿色低碳转型。**一是智能生产调度场景，**

依托专有云的高可靠、低时延智算服务支撑，融合电网负荷、气象数据、能源价格、设备状态等多维度数据，通过大模型实现电力系统日前调度、实时调度、新能源消纳优化、油气田生产动态调整等功能。**二是设备预测性维护场景**，专有云部署的设备健康管理大模型，可实时采集汽轮机、发电机、风电叶片、光伏组件、输油管道等设备的运行数据，通过多模态分析实现故障提前预警、剩余寿命预测与维护方案智能推荐。**三是智能巡检与安全管控场景**，依托专有云与边缘节点的协同架构，大模型可融合视频监控、传感器、无人机巡检等多模态数据，实现变电站无人巡检、油气管道泄漏检测、煤矿井下安全监测、危化品存储管控等功能，自动识别设备异常、安全隐患与违规操作，有效解决了能源行业点多面广、人工巡检难度大、安全风险高的痛点。**四是能源交易与碳资产管理场景**，该场景依托专有云的安全隔离特性，整合能源生产、消费、交易、碳排放等全链条数据，通过大模型实现电力现货交易智能报价、碳足迹全流程核算、碳资产优化管理等功能。

五、大模型专有云服务未来发展趋势

随着大模型、智能体等技术的突破性进展、多智能体应用的规模化落地，以及企业对数据主权与业务价值的双重重视，大模型专有云服务市场正步入由“技术驱动+场景变革+价值重构”共同推动的新阶段。未来，市场将在规模结构、服务模式、竞争升级、行业落地等维度呈现出清晰的演进趋势，产业将从当前的粗放式规模扩

张期，全面迈向高质量、集约化、生态化的全新发展格局。

（一）规模结构：市场体量持续扩容，专有云占比稳步攀升

未来 3—5 年，我国大模型专有云服务市场将保持 80%以上的高增速，市场规模持续跨越式扩容，在整个大模型服务市场中的占比将持续攀升，成为 AI 产业发展的核心增长极。

从总量看，本地化推理需求持续释放，推动专有云推理服务规模与市场占比持续攀升。AI 推理需求的爆发式增长将持续拉动专有云市场扩容，推理服务将成为新增 AI 服务需求的主要部分。随着大模型应用在千行百业的规模化落地，企业级本地化推理需求将持续释放，推动专有云推理服务规模持续增长。预计到 2026 年底，我国大模型专有云推理算力规模将突破 450EFLOPS，在全国 AI 推理算力中的占比将达到约 35%。

从结构看，市场供需持续优化，行业、场景、技术等全面升级，自研技术体系成新增量赛道。市场需求结构将持续升级，高价值场景占比持续提升。政务、金融、能源等核心行业的大模型应用将从浅层通用场景向核心业务场景深度渗透，单客户的智算投入与词元消耗量将持续提升；制造、医疗、交通等实体经济行业的专有云部署需求将加速释放，成为市场增长的核心动力，推动市场需求结构从“政务、金融、能源多核驱动”向“全行业协同增长”转型。同时，随着国产技术体系研发进程加速，基于国产芯片、国产操作系

统、国产大模型的全栈国产化专有云部署占比将持续提升，成为市场重要的增长方向。

（二）服务模式：智能体驱动服务跃迁，词元经营走向精细化

大模型专有云服务模式将加快完成从“资源租赁”向“任务式交付”、从“硬件资源”向“全生命周期服务运营”的根本性转型，未来，这一转型将在智能体技术的驱动下实现第二次跃迁，词元资产化运营也将从当前的范式探索期迈向体系化成熟期，服务模式的深化演进将持续重塑行业价值逻辑。

智能体即服务（Agent-as-a-Service, AaaS）将成为新兴业态，推动服务模式从“模型调用”向“自主执行”升级。随着智能体应用从概念验证走向规模化落地，结合了智能体能力与专有云安全属性的“智能体助手”，将成为专有云服务市场增速最快的领域。专有云的物理隔离特性，既能满足企业对智能体任务执行、多工具协同的需求，又能保障核心业务数据、智能体交互记忆不出域，与智能体部署形成天然适配。未来，专有云服务商将提供覆盖智能体开发、部署、运营、治理的全生命周期 AaaS 服务，让企业无需复杂的技术投入，即可在专有云环境中快速部署适配自身业务的智能体，实现业务流程的端到端自主执行，进一步释放专有云服务的价值潜力。

词元资产化运营走向精细化成熟，从初步沉淀迈向系统化分层

经营体系。当前阶段，头部服务商已初步探索通过专有云环境帮助企业沉淀 **Context** 资产与 **Memory** 资产。未来，这一模式将从探索期迈向成熟期，词元将不再仅仅是计量计费的基本单位，而是真正成为专有云服务的核心价值标尺和企业业务资产沉淀的核心载体。服务商将通过专有云环境，帮助企业实现每一次词元调用的价值沉淀，构建企业专属的业务场景图谱与决策记忆体系，让 **AI** 模型越用越贴合企业业务需求，形成正向复利飞轮。同时，词元的分层运营体系将逐步完善，服务商将根据词元的应用场景、价值密度进行分层定价，针对核心业务场景的高价值词元、通用办公场景的基础词元，制定差异化的服务体系与计费模式，实现精细化运营。

任务式交付将成为主流服务模式，实现从“卖资源”向“卖结果”的转型。未来，专有云服务商将不再局限于提供智算资源与模型部署服务，而是直接面向客户的业务目标，提供端到端的任务式交付服务。客户无需关注底层智算调度、模型优化、运维管理等技术细节，只需明确业务任务目标与验收标准，由服务商完成全流程的服务交付，按任务完成效果进行结算。例如，面向政务 12345 热线场景，服务商直接交付“智能客服座席替代率、客户满意度提升”的业务结果；面向金融营销场景，服务商直接交付“客户转化率提升、营销成本降低”的业务成果，服务价值与客户业务目标深度绑定。

（三）行业竞争：竞争维度全面升级，全栈和国产化是

核心方向

未来，大模型专有云服务市场的竞争将持续升级，从单点能力比拼，全面转向全栈生态能力的综合竞争，市场集中度将进一步提升，头部厂商的领先优势将持续扩大，行业壁垒持续抬高。

竞争壁垒从“全链路整合”向“全链路智能化协同”跃升，具备自进化能力的全栈体系将成为新门槛。当前市场竞争已从单一智算、模型参数比拼逐步升级为全栈能力的综合竞争，而未来，市场竞争的焦点将进一步向更高维度跃迁。全栈协同优化带来的词元效能提升、安全合规保障、场景化适配能力、全生命周期服务能力，将成为厂商的核心竞争力。能否实现基于实际业务数据的持续智能化优化、能否将每一次客户交互转化为模型迭代的养料、能否构建“越用越强”的自进化全栈体系，将成为区分头部厂商与追赶者的核心分水岭。缺乏这些能力的中小厂商，将逐步被市场边缘化，或融入头部厂商的生态体系，市场马太效应将持续凸显。

各梯队发展路径加速分化，头部厂商领先优势持续扩大。第一梯队领航者厂商凭借全栈能力、规模效应与生态优势，将持续扩大市场份额，头部厂商将通过投资并购、生态合作等方式，整合区域市场、垂直行业的优质资源，进一步完善全栈能力与市场布局；第二梯队拓局者厂商将在细分赛道持续深耕，通过差异化能力巩固市场地位，部分具备技术优势的厂商将向第一梯队发起冲击；第三梯队竞争者厂商的市场空间将需通过与头部厂商的生态合作、垂直行

业的深度绑定、垂直行业的专业积累，持续挖掘细分领域潜力，寻求差异化生存与发展空间。

国内厂商技术体系加速成熟，全栈自研能力成为行业竞争门槛与增长赛道。随着国家对关键信息基础设施安全能力的要求持续提升，政务、金融、能源等关键行业的专有云部署，将面向全栈自研技术体系方向转型。具备“自研芯片+自研云平台+自研大模型”能力的厂商，将在市场竞争中占据显著优势。头部厂商将持续加大相关技术研发投入，完善全栈解决方案，推动算力、模型的深度适配与性能优化，抢占市场先机。

（四）落地应用：场景渗透纵深推进，垂直定制激活增长动能

未来，大模型专有云服务的行业落地将呈现“深度渗透、广度拓展”的双重特征，从政务、金融、能源核心行业向千行百业全面拓展，从浅层通用场景向行业核心业务场景深度渗透，垂直行业定制化解决方案将成为市场增长的核心驱动力。

应用场景从浅层通用服务，全面渗透至各行业核心生产、治理业务链路。当前大模型专有云应用仍以智能问答、公文起草、内容生成等浅层通用场景为主，未来将全面向行业核心业务场景深度渗透。政务领域将向城市治理、应急指挥、政策精准落地等核心场景延伸；金融领域将向信贷全流程风控、反洗钱监测、智能投顾等核心业务场景拓展；能源领域聚焦能源安全保供、生产效率提升、绿

色低碳转型等核心目标，布局新能源发电功率精准预测、电网智能调度与故障快速处置、综合能源系统多能协同优化等领域；制造领域将向生产调度、设备预测性维护、工艺优化等核心生产场景落地；医疗领域将向辅助诊断、新药研发、精准医疗等核心医疗场景渗透。核心业务场景的深度渗透，将大幅提升单词元的商业价值密度，推动市场价值的持续增长。

应用行业持续扩容，实体经济成为核心增量，市场向中小微及基层市场下沉。政务、金融、能源行业仍将是专有云服务市场的核心支柱，而制造、交通、物流、医疗、教育、农业等实体经济与民生行业，将成为市场增长的核心引擎。随着大模型技术的持续成熟与部署成本的持续降低，专有云服务将从大型国企、头部机构，向中小微企业、县域市场、基层单位全面下沉，实现千行百业的全覆盖。

垂直行业定制化将成为厂商竞争的核心方向，行业 Know-How 成为关键竞争力。通用大模型与标准化专有云方案，已难以满足不同行业的差异化业务需求，未来市场竞争的核心，将转向垂直行业的深度定制化能力。头部厂商将持续深耕重点行业，积累行业专属语料与业务 Know-How，打造适配不同行业的垂类大模型与行业专有云解决方案，实现大模型能力与行业业务流程的深度融合。具备深度行业理解、完善行业解决方案的厂商，将在细分行业市场构建起难以替代的竞争壁垒，实现差异化突围。

中国信息通信研究院 云计算与数字化研究所

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-62302074

传真：010-62302074

网址：www.caict.ac.cn

