

# 2026年全球算力芯片行业：AI军备竞赛下的“芯”战场

Global computing power chip industry

世界のコンピューティングチップ業界

(精华版)

报告标签：AI芯片、国产替代、人工智能硬件

2026年1月

报告提供的任何内容（包括但不限于数据、文字、图表、图像等）均系头豹研究院独有的高度机密性文件（在报告中另行标明出处者除外）。未经头豹研究院事先书面许可，任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容，若有违反上述约定的行为发生，头豹研究院保留采取法律措施、追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用“头豹研究院”或“头豹”的商号、商标，头豹研究院无任何前述名称之外的其他分支机构，也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。



## 观点摘要

在AI大模型训练与推理需求爆发的推动下，全球算力芯片行业正经历深刻重构：竞争焦点从单一制程性能转向“专用架构+软件生态+能效效率”的系统级较量。美国持续收紧先进芯片及制造设备出口管制，叠加全球AI资本开支向头部云厂商高度集中，加速了算力供应链的地缘割裂，各国纷纷押注自主生态以突破CUDA垄断。

报告深入对比华为、寒武纪等厂商在硬件架构、编译器优化与场景适配上的战略差异，重点评估国产GPU在推理市场规模化落地与训练芯片生态破局的可行性。通过剖析商业模式变迁、生态壁垒高度及“非对称创新”机会，为产业决策与投资布局提供兼具前瞻性与实操性的战略参考。

### ✓ 全球算力爆发与国产替代加速

AIGC与数字经济驱动下，全球算力规模从2021年的615 EFLOPS跃升至2030年预计的20 ZFLOPS（即20,000,000 EFLOPS），智能算力占比达94.1%，AI加速芯片市场规模扩张至7,300亿美元，GPU凭借生态优势主导全球。同期，中国算力增长17倍至3,639.9 EFLOPS，支撑全球40%以上大模型研发，AI加速服务器市场规模达1,662.5亿美元；但受美国持续技术管制影响，非GPU加速服务器占比反超至57%，倒逼中国加速构建自主可控的智能算力体系。

### ✓ GPU：从图形处理器到AI算力核心

GPU是专为并行计算设计的处理器，凭借高ALU密度与高带宽显存，在AI和图形加速中显著优于CPU，其中全功能GPU因软件生态完善成为智能计算核心。自1962年图形学奠基以来，GPU历经图形普及到AI驱动的演进，由NVIDIA主导、AMD追赶，形成以高性能计算为核心的全球格局。其性能由核心数、频率、显存容量、带宽及多精度支持等指标综合决定，持续支撑AI训练、科学计算与渲染升级。

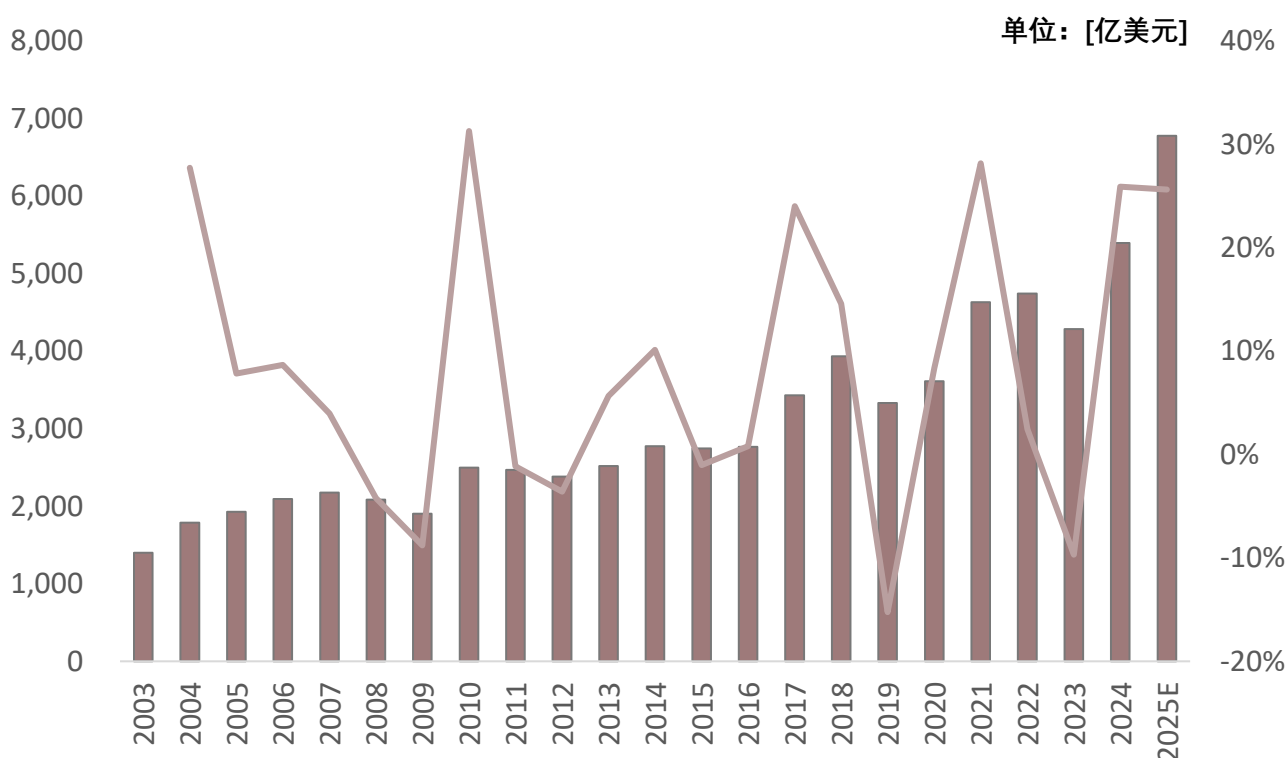
### ✓ 中国AI芯片：多路径突围构建自主生态

中国AI芯片市场呈现“华为与英伟达双雄并立、多元技术路线并进”格局：华为昇腾910B在整型算力上领先，寒武纪MLU590以1,280 GB/s显存带宽突出，昆仑芯专注推理场景优化，摩尔线程与沐曦则发力通用GPU（GPGPU）以兼容图形与计算负载。五家头部厂商正加速构建涵盖芯片、驱动、编译器与框架的全栈软硬件生态，力图在训练与推理双赛道实现非对称突破，支撑国家算力安全与大模型产业可持续发展。

## ■ 精华摘要

全球集成电路市场规模从2003年1,399.6亿美元稳步增长，2017年后在AI、5G等驱动下加速上行，2024年达5,395.1亿美元，预计2025年将升至6,778.5亿美元，长期趋势向上

全球集成电路市场规模，2003-2025E



### ■ 全球集成电路产业：穿越周期波动，AI驱动再创历史新高

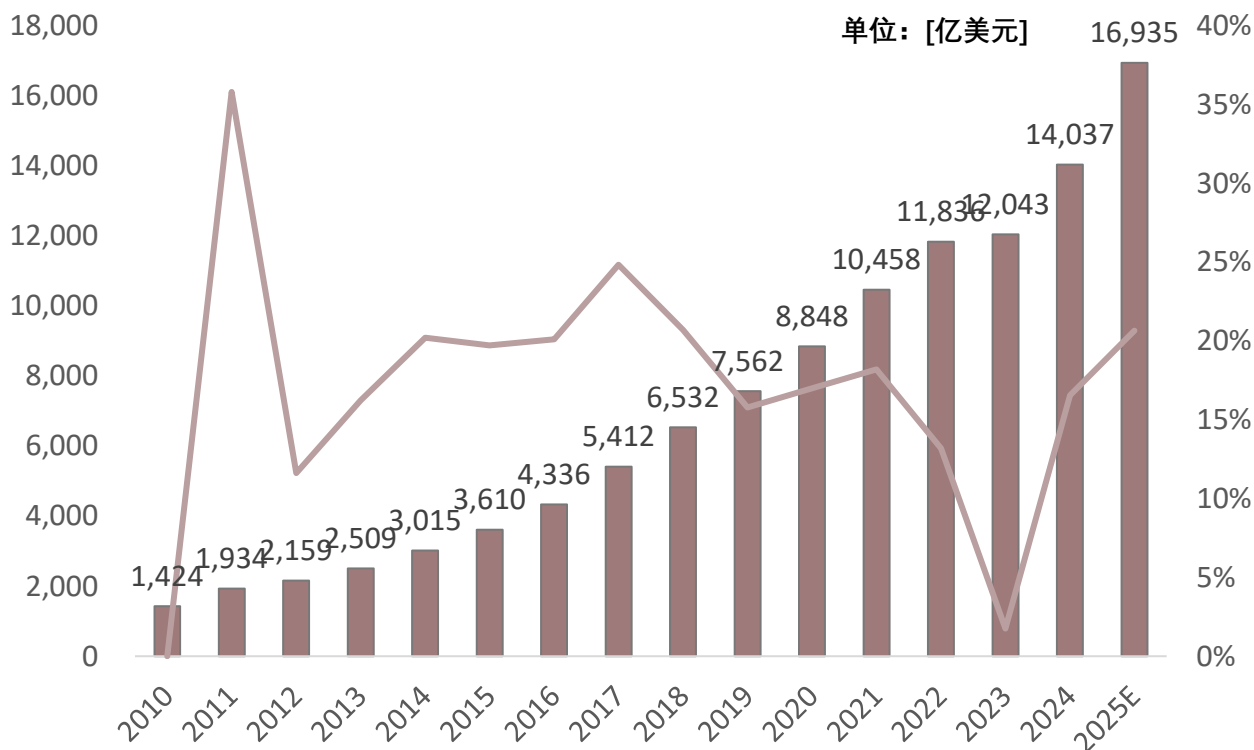
自2017年起，随着人工智能、5G通信、云计算、大数据中心、物联网及自动驾驶等新兴技术加速落地，集成电路产业进入新一轮高速增长周期。市场规模从2017年的3,431.9亿美元快速攀升，在2021年达到4,630.0亿美元的历史高点。尽管2019年受中美贸易摩擦和智能手机市场饱和影响回落至3,333.5亿美元，2023年又因全球通胀高企、库存去化及消费电子需求疲软而暂时回调至4,284.4亿美元，但行业基本面依然坚实。2024年市场强势复苏至5,395.1亿美元，同比增长逾25%，显示出AI芯片、高性能计算和先进制程需求的爆发式增长。展望2025年，WSTS预测全球集成电路市场规模有望进一步扩大至6,778.5亿美元，再创历史新高，凸显半导体作为数字经济核心基石的战略地位持续强化。

来源：WITS，头豹研究院

## ■ 精华摘要

中国集成电路市场规模从2010年的1,424.0亿元持续增长至2024年的14,037.0亿元，年均增速超17%，受益于国产替代、数字经济及AI等新兴需求驱动，预计2025年将达16,935.0亿元

中国集成电路市场规模，2010-2025E



### ■ 中国集成电路产业：万亿底座之上，国产化与AI双轮驱动迈向1.7万亿新高地

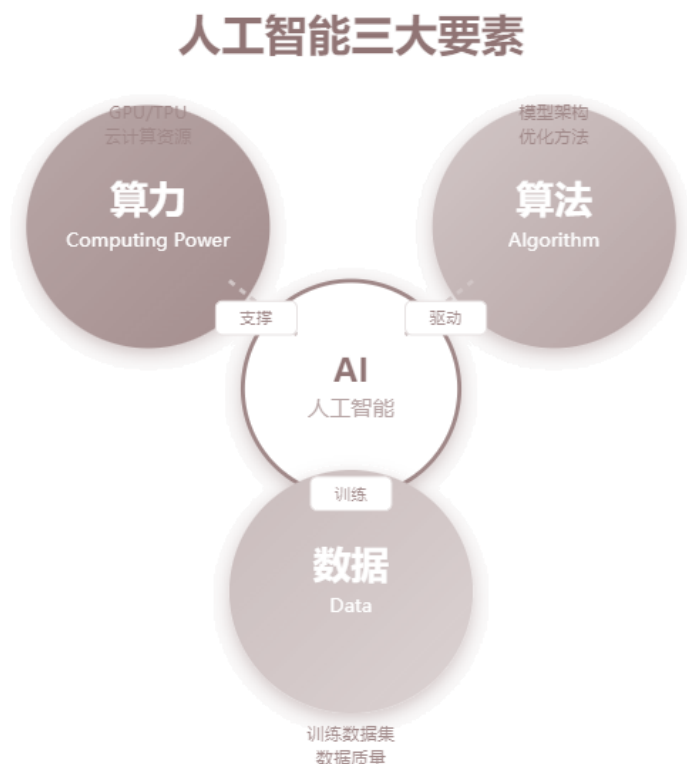
2017年以来，在“中国制造2025”“新基建”及“信创”等国家战略推动下，叠加人工智能、5G通信、新能源汽车、数据中心、工业控制等新兴应用场景爆发，中国集成电路市场进入新一轮高速增长周期。市场规模从2017年的5,411.8亿元一路攀升，2020年突破8,800亿元，2021年迈过万亿元大关，达10,458.3亿元。尽管2022-2023年受全球半导体下行周期、库存调整及地缘政治等因素影响，2023年仅微增至12,043.0亿元，增速明显放缓，但自2024年起行业迎来强劲复苏，全年市场规模达14,037.0亿元，同比增长约16.6%。2025年，预计市场规模将进一步扩大至16,935.0亿元，持续巩固中国作为全球最大集成电路消费市场和关键制造基地的地位，同时也凸显国产化率提升与技术突破的长期战略价值。

来源：CSIA，摩尔线程，头豹研究院

## ■ 精华摘要

在人工智能体系中，算力作为核心动能，依托GPU/TPU等专用芯片与云边协同架构，支撑大模型从千亿到万亿参数的训练与推理，已成为驱动生成式AI突破和数字经济发展的关键底座

### 人工智能三大要素



#### ■ 算力筑基：AI大模型时代下的数字基础设施新底座

在人工智能体系中，算力、算法与数据构成三大不可或缺的核心支柱，三者相互依存、协同演进，共同推动人工智能特别是生成式AI（AIGC）的快速发展。算力是动能，为模型训练与推理提供物理基础和性能保障。近年来，以GPU、TPU、FPGA及ASIC为代表的专用计算芯片持续迭代，单芯片算力已迈入千TFLOPS量级；同时，云计算、边缘计算与雾计算构建起多层次、弹性化的算力供给体系，支撑从超大规模数据中心训练到终端低延迟推理的全场景需求。随着大模型参数量从亿级跃升至万亿级，对AI服务器、高速光模块、交换机及IDC基础设施的需求显著增长，算力产业链正成为数字经济的关键底座。

来源：《“算力数据算法”一体化发展趋势》，科普中国，头豹研究院

## ■ 精华摘要

2022年以来，美国对华AI芯片管制从全面禁售转向2026年对H200等“逐案审查”，在技术封锁与商业压力间寻求平衡，凸显中美算力博弈长期化

### 近年美国对中国的芯片政策

时间	事件	核心内容	中方反应/影响
2022年8月	美国《芯片与科学法案》签署	拨款520亿美元扶持本土半导体制造，遏制中国在先进制程领域的崛起	中国批评美方“泛化国家安全”，加速国产替代
2022年10月	首轮对华高端芯片出口管制	限制A100/H100等高性能AI芯片及制造设备对华出口，聚焦军事用途风险	中国称美方“恶意打压”，华为等企业启动自研突围
2022年12月	扩大实体清单	将36家中国芯片及国防相关企业列入黑名单，切断其获取美技术渠道	多家企业被列入，国产EDA、设备攻关提速
2023年10月	升级管制范围	将L40、A800等“降规版”芯片纳入管制，回应华为Mate 60搭载自研7nm芯片	华为突破引发美方警觉，管制从“顶尖”扩展至“中高端”
2024年12月-2025年1月	拜登政府“最后加码”	要求全球供应链配合管制，限制第三国向中转售先进芯片；设“友好国家”例外	中国抗议“长臂管辖”，推动RISC-V、Chiplet等非美路径
2025年1月	DeepSeek大模型出圈	中国AI公司DeepSeek发布高性能开源模型，登顶App Store榜单	彰显中国在算法与应用层追赶能力，刺激美方焦虑
2025年5月	特朗普政府部分松绑	取消部分国家芯片进口上限，担忧过度管制反促他国转向中国方案	美国内部出现分歧：安全派 vs. 商业派
2025年4月	H20芯片被临时叫停	英伟达专为中国设计的H20因“可能用于超算”被要求申请许可	出口政策不确定性加剧，企业合规成本上升
2025年7月	H20恢复对华销售	美方承诺简化H20许可流程，黄仁勋年内三访中国	显示商业利益与地缘政治间的动态平衡
2026年1月13日	H200等芯片管制微调	将H200、MI325X等符合性能阈值（TPP<21,000，带宽<6,500 GB/s）的芯片出口许可由“推定拒绝”改为“逐案审查”	政策小幅松动，但保留军事用途“推定拒绝”红线
2026年1月12日	《远程访问安全法案》通过	堵住“芯片不出境、算力远程用”的漏洞，限制境外用户远程调用美国AI芯片	针对“云服务绕道”，强化全链条管控

来源：NDTV, CGGT, 头豹研究院

## ■ 精华摘要

计算加速芯片产业链涵盖上游EDA/IP/设备材料、中游IDM/Fabless芯片设计与制造（如英伟达、华为海思、寒武纪）、下游AI服务器及终端应用（如人工智能、自动驾驶、机器人、消费电子）

计算加速芯片产业链



来源：各企业官网，摩尔线程，沐曦股份，头豹研究院

## ■ 精华摘要

2025年国家密集出台政策，从芯片攻关、算力基建、标准体系到税收优惠，全链条支持计算加速芯片研发与应用，加速构建自主可控的智能算力生态

中国算力芯片相关政策，2025年



### ■ 政策加码AI芯：2025年国家力推计算加速芯片筑牢新质生产力底座

2025年，国家层面密集出台多项政策，系统性强化对计算加速芯片及智能算力基础设施的支持。国务院在《关于深入实施“人工智能+”行动的意见》和《2025年政府工作报告》中明确提出，要支持人工智能芯片攻坚创新与使能软件生态培育，加快超大规模智算集群技术突破，并将AI芯片作为发展智能终端（如AI手机、智能机器人、智能网联汽车）和智能制造的核心支撑，凸显计算加速芯片在新质生产力构建中的战略地位。

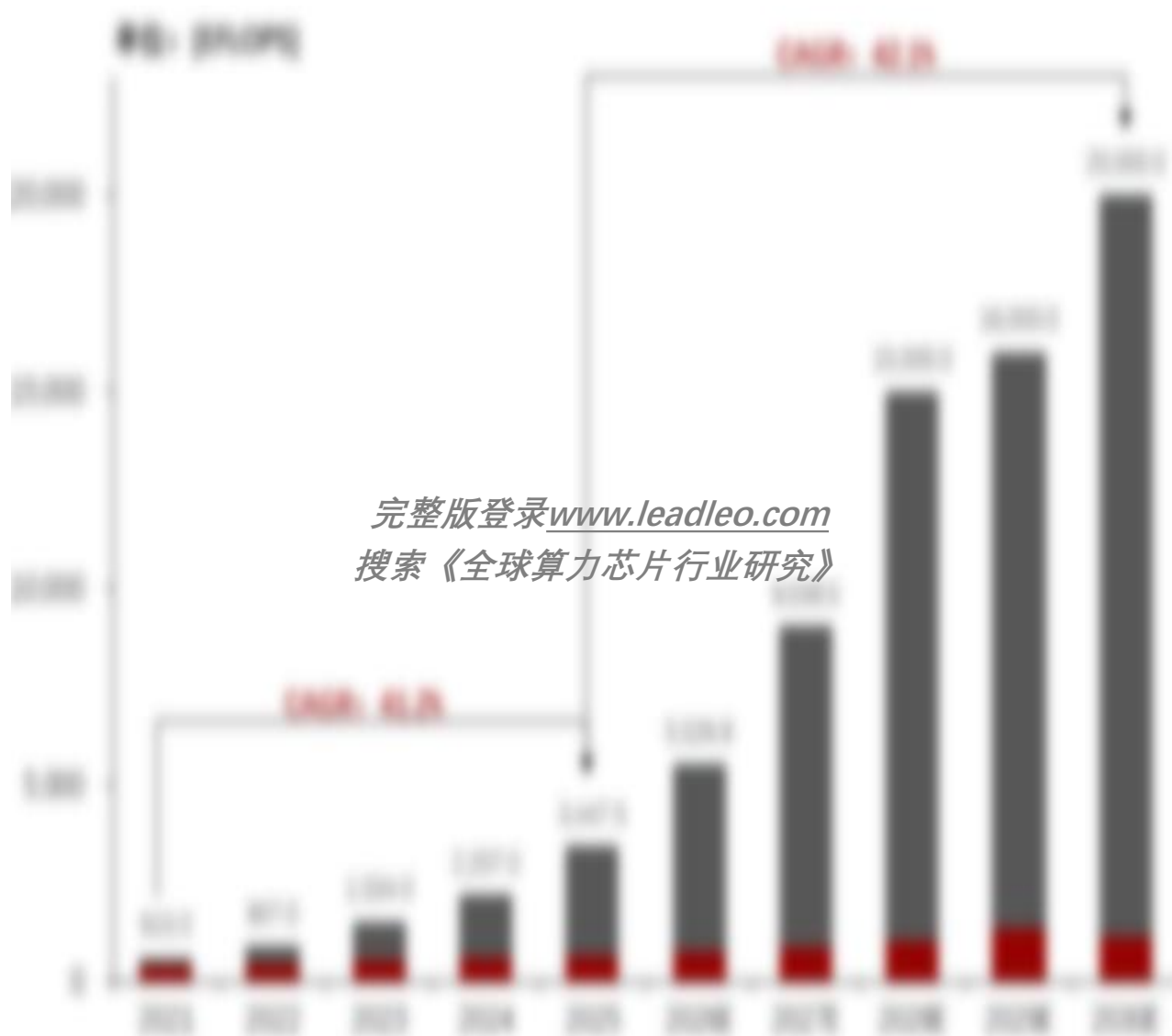
来源：中国政府网，摩尔线程，头豹研究院

## ■ 精华摘要

在AIGC浪潮、AI芯片大规模部署及全球算力基建加速推动下，2021至2030年，全球算力总规模（FP32等效）从615EFLOPS跃升至20 ZFLOPS，其中智能算力占比从37.7%升至94.1%

2021至2030年，全球算力总规模（统一按FP32等效口径）从615 EFLOPS迅猛增长至20,000 EFLOPS（即20 ZFLOPS），近十年扩张超32倍。其中，智能算力成为绝对主导力量，占比由2021年的37.7%跃升至2030年的94.1%；基础算力稳步提升但增速趋缓，而超算算力保持低速增长，整体结构呈现“智能算力一极引领、通算超算协同支撑”的格局。

全球算力规模（FP32），2021-2030年预测



完整版登录[www.leadleo.com](http://www.leadleo.com)  
搜索《全球算力芯片行业研究》

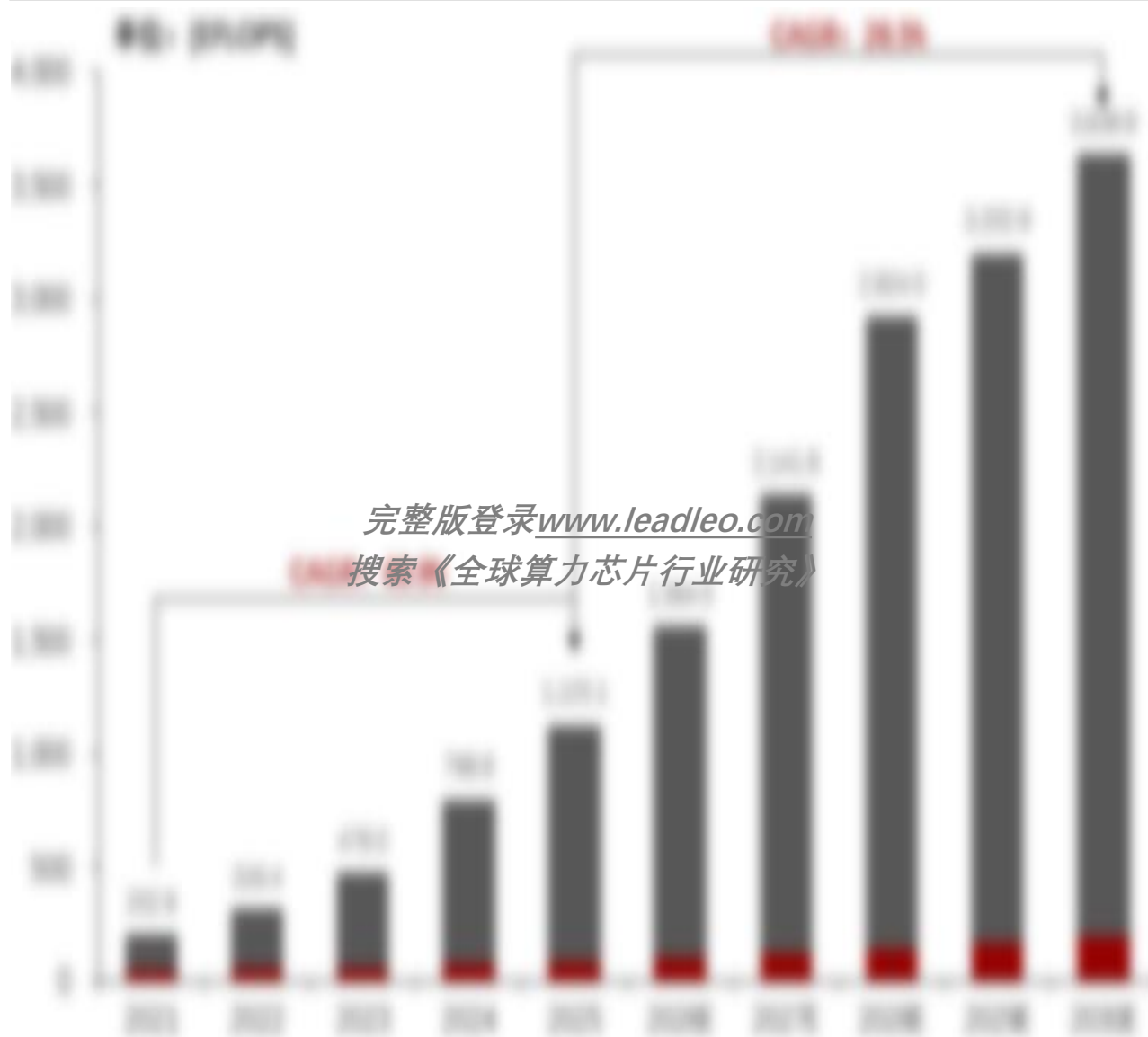
来源：中国信通院，深圳市人工智能产业协会，头豹研究院

## ■ 精华摘要

2021至2030年，中国算力总规模（FP32）从202.9 EFLOPS增至3,639.9 EFLOPS，增长约17倍，智能算力占比升至94.6%，在用机架超4,100万架，目前支撑全球40%以上大模型研发，稳居世界算力第一梯队

2021至2030年，中国算力总规模（FP32统一口径）从202.9 EFLOPS增至3,639.9 EFLOPS，增长约17倍，年均复合增速超36%。智能算力成为核心驱动力，占比由76.5%提升至94.6%，在用标准机架数从230.6万架扩至4,138.3万架；同期，中国占全球算力比重从33.0%调整至18.2%，主要因全球AI竞赛全面提速，多国大规模投入智算基建，推动全球算力总量跃升至更高量级，持续发挥引领作用。

中国算力规模（FP32），2021-2030年预测



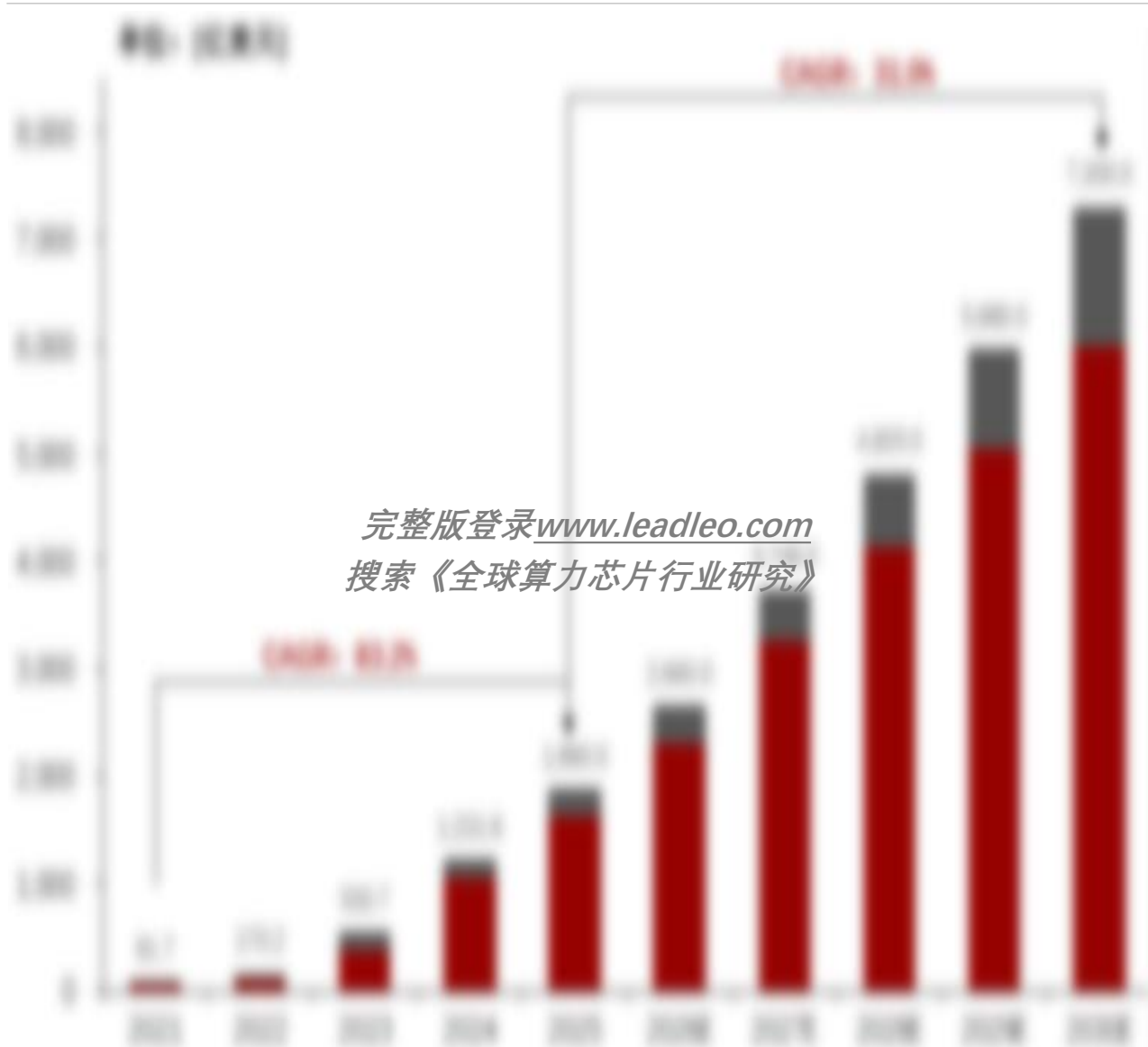
来源：中国信通院，浪潮信息，工信部，全国数据资源调查报告（2024年），头豹研究院

## ■ 精华摘要

2021至2030年，全球AI加速芯片规模从91.7亿增至7,300亿美元，GPU凭借生态与性能优势主导市场（占比73%→82%），ASIC/FPGA在推理场景协同补充，共同支撑大模型驱动的算力爆发

这一格局主要源于大模型训练对通用、高吞吐、强生态的GPU架构的高度依赖。以NVIDIA H100/B100为代表的GPU凭借CUDA生态、大规模并行能力和持续迭代的硬件优势，成为全球云厂商和AI实验室的首选；同时，“东数西算”、国家级智算中心建设及AIGC产业化浪潮推动数据中心资本开支激增，进一步巩固GPU在AI基础设施中的核心地位，而ASIC/FPGA则在特定推理场景中发挥能效优势，形成“GPU主导、多元协同”的算力生态。

全球AI加速芯片规模，2021-2030年预测



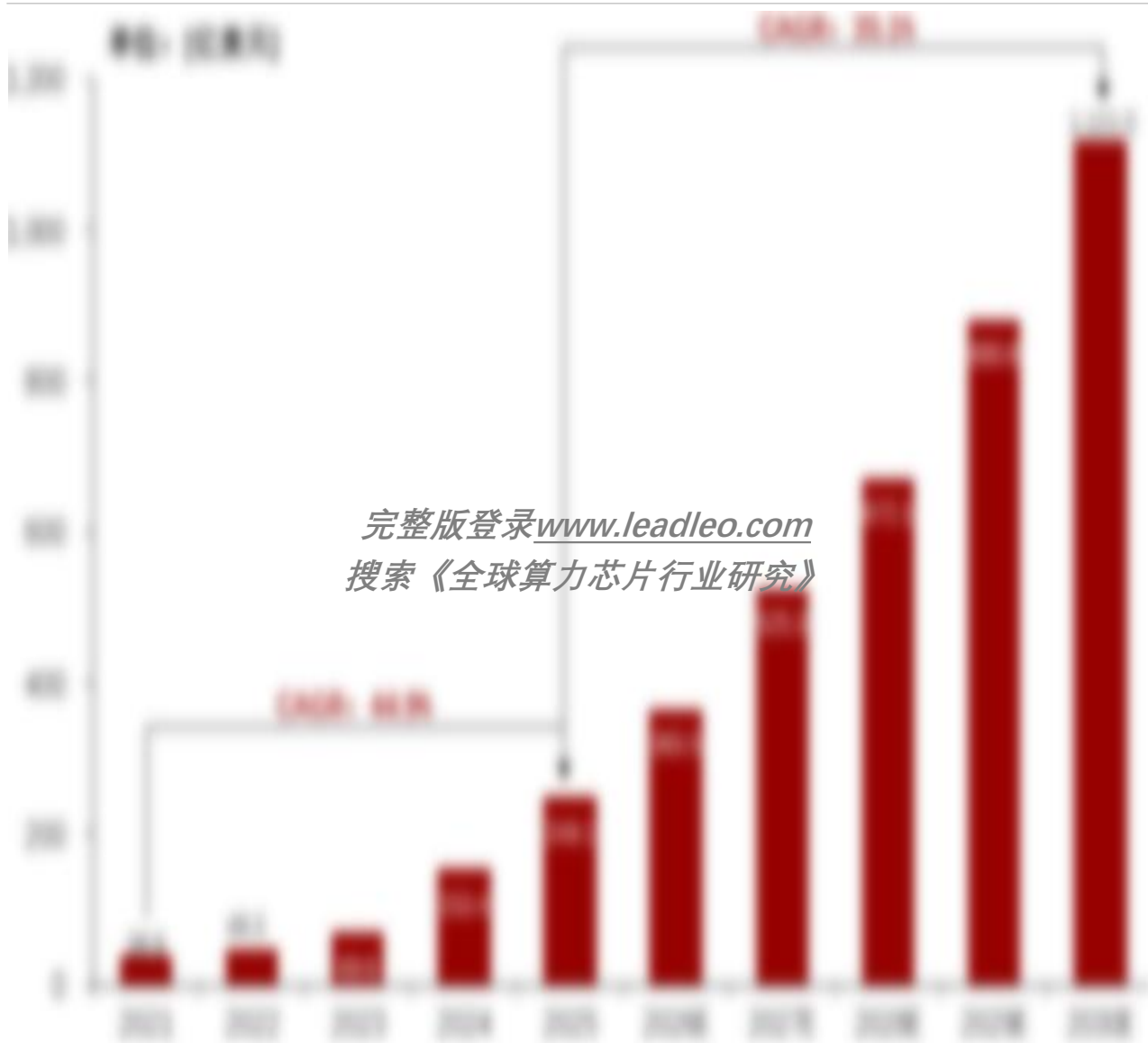
来源：公开资料搜集，头豹研究院

## ■ 精华摘要

2021至2025年，中国AI加速芯片市场规模从38.9亿增至249.1亿美元，在大模型、智算基建与国产替代驱动下爆发式增长，预计2030年将超1,121亿美元，加速构建自主可控的全球智能计算核心力量

展望未来，在“东数西算”工程持续推进、各地智算中心密集建设、行业大模型广泛应用以及政策对自主可控芯片的强力扶持下，中国AI加速芯片市场有望保持强劲增长势头。预计2026年规模将达365.5亿美元，并于2030年突破1,121亿美元，2025-2030年复合增长率约35%，不仅支撑起国内超大规模AI训练与推理需求，也将进一步巩固中国在全球智能计算生态中的战略地位。

中国AI加速芯片规模，2021-2030年预测



完整版登录 [www.leadleo.com](http://www.leadleo.com)  
搜索《全球算力芯片行业研究》

来源：公开资料搜集，头豹研究院

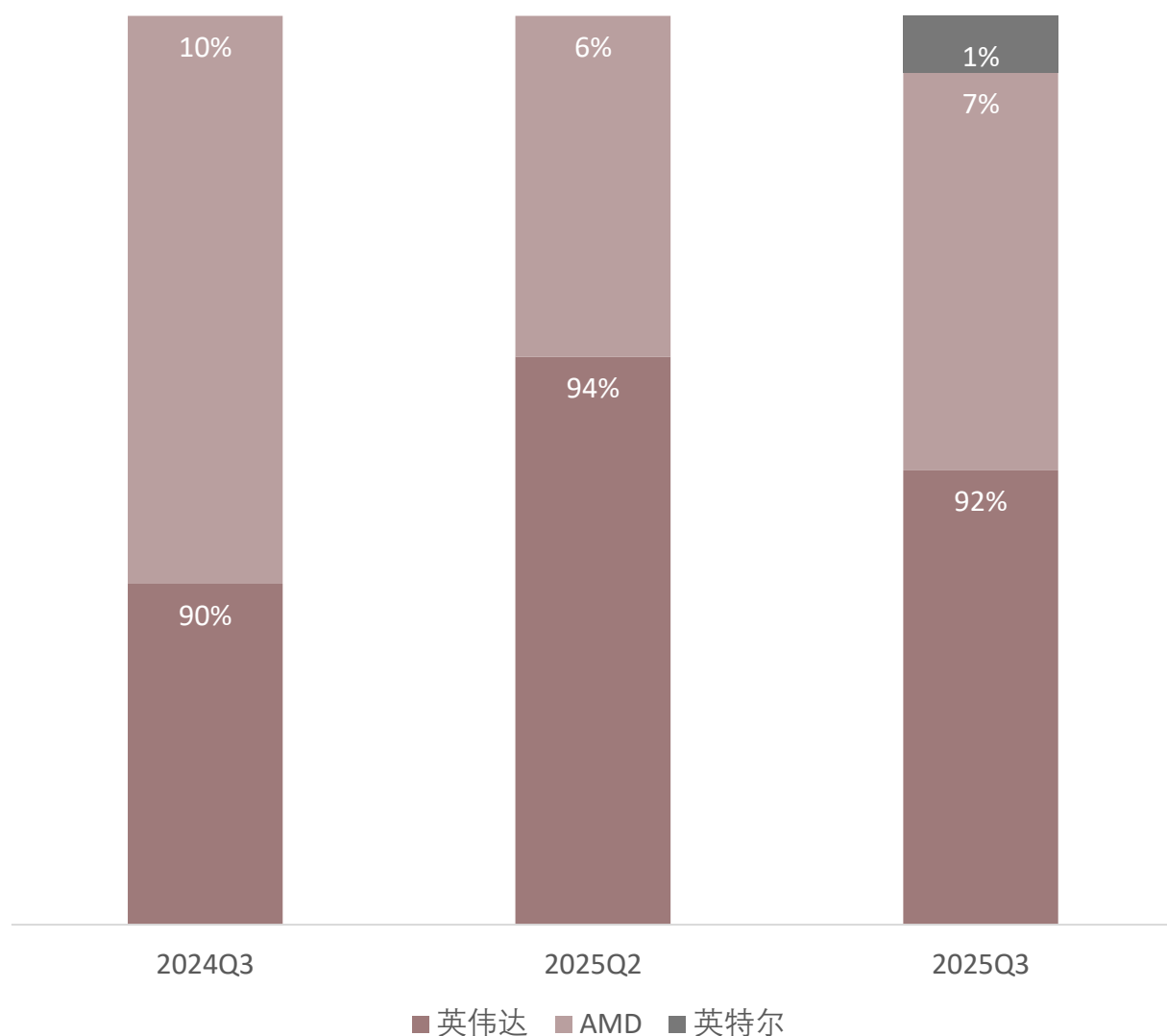
## ■ 精华摘要

2025年全球数据中心GPU市场高度分化。英伟达主导数据中心与独立显卡，AMD在PC GPU市场反超至61%，中国AI加速芯片领域则由华为与英伟达双雄并立

2025年，全球GPU市场呈现显著分化：在数据中心GPU市场中，英伟达预计占据79%份额，AMD占7%，其他厂商（包括英特尔核显及国产芯片）合计14%。

但在独立显卡市场（仅含插槽式高性能显卡），英伟达依然强势主导，2025年份额维持在92-94%，AMD约6-7%，英特尔尚未形成有效竞争。

全球独立显卡市场份额，2024Q3-2025Q3



来源：JPR, 伯恩斯坦, 头豹研究院

未完待续

下篇正在进行中

若您期待尽快看到下篇报告或对下篇报告的内容有独到见解，头豹欢迎您加入到此篇报告的研究中。相关咨询，欢迎联系头豹研究院工业研究团队

邮箱：

[sharlin.chen@leadleo.com](mailto:sharlin.chen@leadleo.com)

18129990784

## 完整版研究报告阅读渠道：

- 登录[www.leadleo.com](http://www.leadleo.com)，搜索《2026年全球算力芯片行业：AI军备竞赛下的“芯”战场》

## 了解其他集成电路系列课题，登陆头豹研究院官网搜索查阅：

- 中国集成电路产业白皮书
- 中国集成电路产业政策分析概览
- 2024年集成电路封测行业概览
- 2023年集成电路半导体设备制造行业概览
- 2021年卡脖子系列—全球光刻机行业概览

## 方法论

- ◆ 头豹研究院布局中国市场，深入研究19大行业，持续跟踪532个垂直行业的市场变化，已沉淀超过100万行业研究价值数据元素，完成超过1万个独立的研究咨询项目。
- ◆ 研究院依托中国活跃的经济环境，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- ◆ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

## 法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何投资建议。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。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。本报告所指的公司或投资标的的价值、价格及投资收入可升可跌。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本文所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本文所载资料、意见及推测不一致的报告和文章。头豹不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

# 头豹业务合作

## 数据库/会员账号

可阅读全部原创报告和百万数据，提供数据库API接口服务

## 定制报告

行企研究多模态搜索引擎及数据库，募投可研、尽调、IRPR等研究咨询

## 定制白皮书

对产业及细分行业进行现状梳理和趋势洞察，输出全局观深度研究报告

## 招股书引用

研究覆盖国民经济19+核心产业，内容可授权使用至上市文件、年报

## 市场地位确认

对客户竞争优势进行评估和调研确认，助力企业品牌影响力传播

## 行研训练营

依托完善行业研究体系，帮助学生掌握行业研究能力，丰富简历履历

## 报告作者



陈夏琳

首席分析师

sharlin.chen@leadleo.com



马天奇

行业分析师

Kareem.ma@leadleo.com

## 业务咨询

- 客服电话：400-072-5588
- 官方网站：[www.leadleo.com](http://www.leadleo.com)



商务咨询与深度合作

### 深圳办公室

广东省深圳市南山区粤海街道华润置地大厦E座4105室

邮编：518057

### 上海办公室

上海市静安区南京西1717号会德丰国际广场 2701室

邮编：200040

### 南京办公室

江苏省南京市栖霞区经济开发区兴智科技园B栋401

邮编：210046

主办方：高顿教育 FROST & SULLIVAN 沙利文 阿里云 | 协办单位：头豹 LeadLeo

# 第四届全国大学生 预见未来行业研究大赛

现已正式开启报名通道

报名时间：2026年4月1日 - 4月20日



## • 比赛收获 •

丰富简历

名企实习offer

100% 官方参赛证书

搭建专属工作流+行研skills

大咖导师带教

行研技能提升

万元现金奖金