

小米 AI 战略演进：从“端侧优先”到“全生态 AI 化”

核心观点

小米当前对 AI 的投入转变，本质上是公司在完成一次与 2021 年造车相似、但层级更高的“第二次重仓下注”——从押注一个新终端品类，升级为押注下一代系统级能力。造车阶段，市场给小米加估值，更多是因为它多了一条大赛道和第二增长曲线。AI 阶段，如果市场真正买账，改变的就不仅是某个分部估值，而是整个集团的属性认知。如果市场认知到：小米不是硬件拼盘，而是一个能用 AI 重新组织终端生态价值的公司，那么传统 SOTP 并不完整，而是应该转向“整体平台重估”，AI 比造车更难短期定价，但长期想象空间更大，小米是稀缺的同时具备自研模型、系统权限、终端入口、制造场景的全栈型 AI 玩家。

摘要

小米当前对 AI 的投入转变，本质上是公司在完成一次与 2021 年造车相似、但层级更高的“第二次重仓下注”——从押注一个新终端品类，升级为押注下一代系统级能力。造车时期，小米是用首期 100 亿元、十年 100 亿美元去换取“人车家”生态中“车”这一重资产入口；而本轮 AI 布局中，小米已明确未来三年投入 600 亿元、2026 年单年投入 160 亿元、2026-2030 年研发合计超 2000 亿元，其目标不再只是增加一条业务线，而是争夺未来“人车家”所有终端的统一智能中枢、交互入口与系统定义权。

雷军作为创始人仍在一线且话语权极高，其核心意义不只是“提振士气”，而是显著降低了小米在战略跃迁期最容易出现三类风险：方向摇摆、组织内耗与资源错配。如果说造车验证了“雷军亲自下场，小米能把一个外界普遍认为极难的重资产新业务做成”，那么 AI 阶段验证的将是“能不能把小米从硬件生态公司推向 AI 平台公司”。二者最大的差别在于：造车更多改变收入结构，AI 则可能重写估值框架。

为什么要把小米 AI 与当年造车放在一起看？意义不只是做历史类比，而是要回答三个更本质的问题。第一，小米为什么总是在看起来“主业未必完全无忧”时做重投入战略转向；第二，雷军在这种转向中扮演的角色，到底是品牌型创始人还是实质性决策中枢；第三，本轮 AI 与当年造车相比，究竟只是重复过去的方法论，还是在战略层级上已经更进一步。

小米集团-W(1810.HK)

维持

买入

于伯韬

yubotao@csc.com.cn

SAC 编号:S1440520110001

SFC 编号:BRR519

发布日期：2026 年 06 月 07 日

当前股价：27.80 港元

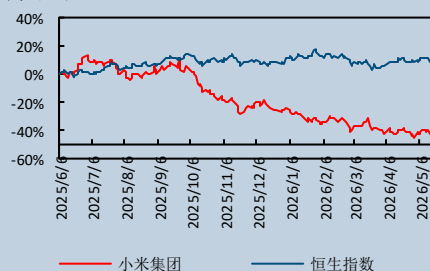
目标价格 6 个月：38.7 港元

主要数据

股票价格绝对/相对市场表现 (%)

1 个月	3 个月	12 个月
-10.26/-5.92	-13.12/-11.99	-47.99/-53.52
12 月最高/最低价 (港元)		60.15/27.80
总股本 (万股)		2,583,989.03
流通 H 股 (万股)		2,132,150.13
总市值 (亿港元)		7,183.49
流通市值 (亿港元)		5,927.38
近 3 月日均成交量 (万)		16590.58
主要股东		
Smart Mobile Holdings Limited		22.66%

股价表现



相关研究报告

- 26.05.28 【中信建投海外研究】小米集团-W(1810):26Q1 财报点评：股东回报持续加码，业绩短期承压下关注下半年产品周期+大模型进展小米集团发布 2026Q1 业绩，总收入 991 亿元 (yoy-10.9%，qoq-15)
- 26.05.08 【中信建投海外研究】小米集团-W(1810):26Q1 业绩前瞻：手机毛利率预计得到较好控制，新一代 su7 锁单量持续爬升

基座模型：MiMo 万亿级模型，进入开源模型第一梯队。小米基座大模型的演进路径，核心不是简单追求参数扩张，而是沿着“端侧轻量化起步—云边端矩阵化扩展—开源推理基座突破—Agent 与全生态底座化”的路线持续升级。2023 年是技术立项与 MiLM 验证期，2024 年是 MiLM2 产品化落地期，2025 年是 MiMo 开源基座跃迁期，2026 年是 MiMo-V2/V2.5 冲击全球第一梯队的能力释放期。2026 年：MiMo-V2 与 V2.5 将小米推入全球第一梯队。2026 年是小米大模型能力集中释放的一年。2026 年 3 月，小米正式发布面向 Agent 时代的旗舰基座模型 MiMo-V2-Pro，总参数超过 1 万亿，激活参数 42B，支持 100 万 token 上下文。同期还发布 MiMo-V2-Omni 与 MiMo-V2-TTS，小米移动端 Agent Xiaomi miclaw 亦进入小范围封测，开始验证模型从“对话能力”向“系统级执行能力”的升级路径。同样重要的是其经济性和 Token 使用效率：在能力相近情况下，相比 Claude Opus 4.6、Gemini 3.1 Pro、GPT-5.4，可减少约 40%-60% token 消耗。对企业级调用场景而言，是直接影响商业化成本结构的核心变量。

小米 2025—2026 年 AI 应用落地的核心，是把 Agent 从手机端的任务执行器，扩展为贯穿手机、汽车、家庭与机器人四类物理终端的统一调度中枢。这一进程的关键变化，在于 AI 不再停留在“问答式入口”，而是开始获得系统权限、跨端记忆、多模态感知、设备编排与物理执行能力，从而推动小米“人车家全生态”由互联网络升级为可被统一驱动的行动网络。

从落地节奏看，小米已经形成较清晰的三层推进路径：第一层是手机 Agent 化，以 Xiaomi miclaw 为代表，验证系统级执行闭环；第二层是车和家的空间 Agent 化，分别通过 XLA/OneVL 认知智驾与 Miloco 全屋智能方案，把 AI 从屏幕带入驾驶空间和居住空间；第三层则是具身执行端，借助 Xiaomi-Robotics-0、TacRefineNet 与工厂实习机器人，将 AI 最终延伸至真实工业动作执行。

从小米的角度看，AI 最现实的回报路径，未必是订阅，而是先体现在产品力与组织效率。从商业模式上，小米现阶段并不急于把 AI 单独变现，当前更关注模型、应用与生态能力成熟，Token Plan、订阅制、高端产品溢价、API 服务等均在探索中，但尚未把其设定为核心 KPI。

AI 比造车更难短期定价，但长期想象空间更大。造车的好处是可见度更高，销量、ASP、毛利率都能跟踪；AI 的难点在于大量价值短期难进入报表，所以更容易出现“看得到战略，看不到业绩”的定价错配。但也正因为如此，一旦 AI 真正打通系统级入口、订阅和硬件溢价，其整体重估幅度可能高于单一汽车业务。小米已经成为少数同时具备**自研模型、系统权限、终端入口、制造场景**的全栈型 AI 玩家。

尽管短期传统主业在大周期下承压，存储的价格仍在快速上涨，成本周期拐点可能在 2027 年以后，但受益于前期存储备货节奏、公司主动调整产品结构、以及 4 月起部分产品涨价拉动，手机毛利率预计得到较好控制，此外汽车新品周期+产能爬坡有望形成催化。另一方面，公司仍然积极投入 AI，在更长远的时间将是最核心的竞争力之一。

预计小米集团 2026/2027 年收入 4846/6077 亿元(yoy+6%/+25%)，调后净利润 286/393 亿元(yoy-27%/+37%)，我们给予公司核心主业 26 年 20x PE，汽车 26 年 2.5x PS，目标价 38.7 港币，给予“买入”评级。

目录

一、小米 AI 战略演进：从“端侧优先”到“全生态 AI 化”	1
1.1 从“谨慎布局”到“全方位进军”，小米 AI 投入发生了什么转变？	2
1.2 AI 组织架构变化&创始人作为核心领导的意义	4
1.2.1 雷军仍在一线且话语权高，意味着什么？	4
1.2.2 MiMo Core 团队：从“端侧 AI 能力补位”转向“集团基座模型中枢”	5
1.3 投入对比：非云业务视角下的投入强度加码	7
二、基座模型：MiMo 万亿级模型，进入开源模型第一梯队	8
三、AI 应用落地方向：人车家全生态的 Agent 化进程	13
3.1 手机端先行：miclaw 初步试水	13
3.2 汽车：智驾与座舱是 AI 反哺业务最直接的利润场景	15
3.3 家庭空间 Agent 化，AI 将“可连接”推进到“可理解、可主动服务”	15
3.4 机器人与制造：AI 外溢到物理世界，是长期想象力所在	16
投资建议与盈利预测	18
风险分析	21

图目录

图 1：小米 AI 投入的转变	2
图 2：小米集团聚焦 AI 战略，研发规模持续扩大	3
图 3：雷军在造车与 AI 战略中的角色对比	5
图 4：小米大模型团队组织架构	6
图 5：小米历年研发费用	7
图 6：小米历年 capex 投入	7
图 7：Artificial Analysis Intelligence Index	9
图 8：Artificial Analysis Coding Index	9
图 9：Artificial Analysis Agentic Index	9
图 10：MiMo-V2.5-Pro&MiMo-V2-Pro benchmark	11
图 11：Price of leading AI models-blended price	12
图 12：Intelligence vs. Cost to Run Artificial Analysis Intelligence Index	12
图 13：小米围绕人车家全生态建设应用 AI 大模型与 Agent 技术的技术架构	13
图 14：小米最新一代机器人（2026 投资者日）	16
图 15：Xiaomi-Robotics-0 模型与真机后训练流程示意图	17

表目录

表 1：小米造车与 AI 战略的核心对照框架	2
表 2：雷军在造车与 AI 两大战略中的核心作用	4
表 3：小米大模型组织重构与分工	6
表 4：小米基座大模型阶段性演进（2023-2026）	8
表 5：2025-2026 年小米 MiMo 系列与行业头部水平对比	10

表 6: MiMo-V2.5-Pro 较 MiMo-V2-Pro 的核心提升.....	12
表 7: miclaw 基本能力.....	14
表 8: 假设收到购票成功信息，miclaw 可能会做的步骤.....	14
表 9: Miloco 相对传统智能家居规则引擎的能力跃迁	16
表 10: 造车与 AI 对小米估值框架的影响差异	18

一、小米 AI 战略演进：从“端侧优先”到“全生态 AI 化”

小米当前对 AI 的投入转变，本质上是公司在完成一次与 2021 年造车相似、但层级更高的“第二次重仓下注”——从押注一个新终端品类，升级为押注下一代系统级能力。造车时期，小米是用首期 100 亿元、十年 100 亿美元去换取“人车家”生态中“车”这一重资产入口；而本轮 AI 时期，小米已明确未来三年投入 600 亿元、2026 年单年投入 160 亿元、2026-2030 年研发合计超 2000 亿元，其目标不再只是增加一条业务线，而是争夺未来“人车家”所有终端的统一智能中枢、交互入口与系统定义权。

雷军作为创始人仍在一线且话语权极高，其核心意义不只是“提振士气”，而是显著降低了小米在战略跃迁期最容易出现三类风险：方向摇摆、组织内耗与资源错配。造车阶段，雷军以“人生最后一次重大创业项目”“押上全部声誉”的方式，亲自担任汽车业务 CEO，形成了极强的组织穿透力与资源聚焦力；而 AI 阶段，雷军虽然不再复制造车时那种“事无巨细的前线指挥员”角色，但仍以战略拍板者、关键人才引入者、资源配置终审者身份深度介入，典型体现包括其在 2023 年 8 月将 AI 上升为集团科技战略核心、2024 年底亲自参与引进罗福莉、2025-2026 年连续将 AI 预算和技术路线公开升级。

如果说造车验证了“雷军亲自下场，小米能把一个外界普遍认为极难的重资产新业务做成”，那么 AI 阶段验证的将是“能不能把小米从硬件生态公司推向 AI 平台公司”。二者最大的差别在于：造车更多改变收入结构，AI 则可能重写估值框架。汽车业务到 2025 年已实现收入 1061 亿元、全年经营利润转正、收入占集团比重约 23%，证明第一轮重仓已基本跨过商业可行性门槛；AI 则仍处于“投入显著前置、盈利后置”的阶段，其短期更可能体现为硬件 ASP 抬升、组织效率改善、系统体验增强，中长期才可能进入 Token、订阅、Agent、Physical AI 的直接兑现期。

现在小米的 All in AI 与 2021 年造车宣布时的逻辑高度相似——都是在主业仍有基础盘时，提前对下一轮核心场景下注；但又明显更进一步，因为 AI 与芯片、OS、机器人、汽车、手机、IoT 之间的耦合度更高，成功后带来的不是单点增长，而是系统性溢价重估。

回顾小米造车和 AI：2021 年 3 月 30 日，小米宣布正式进军智能电动汽车业务，首期投入 100 亿元，未来十年投资 100 亿美元，并由雷军亲自兼任智能电动汽车业务 CEO。当时这是一场典型的“跨行业重资产下注”，当年市场关注的核心是：小米为什么要从高周转消费电子跨入低容错制造业同时又是汽车这样极度内卷的红海赛道，雷军为什么愿意把自己重新推回创业第一线。到了 2023 年 8 月 14 日，雷军在年度演讲中宣布小米科技战略升级，提出“深耕底层技术、长期持续投入、软硬深度融合、AI 全面赋能”。从这个节点开始，AI 已不再是“功能增强”或“语音助手演进”，而是被正式提升为集团级战略语言。而真正的转折则发生在 2024 年下半年至 2025 年：小米从此前更偏端侧、轻量化的 MiLM 体系，开始明显转向通用基座模型与 Agent 路线，并在 2024 年底亲自引入罗福莉，随后模型演进速度显著加快。

为什么要把小米 AI 与当年造车放在一起看？意义不只是做历史类比，而是要回答三个更本质的问题。第一，小米为什么总是在看起来“主业未必完全无忧”时做重投入战略转向；第二，雷军在这种转向中扮演的角色，到底是品牌型创始人还是实质性决策中枢；第三，本轮 AI 与当年造车相比，究竟只是重复过去的方法论，还是在战略层级上已经更进一步。

表 1: 小米造车与 AI 战略的核心对照框架

对比维度	造车战略	AI 战略
战略启动标志	2021 年 3 月 30 日正式宣布造车	2023 年 8 月 14 日科技战略升级, AI 全面赋能
初始资源承诺	首期 100 亿元, 十年 100 亿美元	2026 年 160 亿元, 未来三年 600 亿元
战略目标	获取“车”这一新终端与新增长极	获取“AI 中枢/下一代交互入口/系统定义权”
创始人角色	雷军亲自带队、兼任 CEO	雷军担任战略拍板者+关键人才引入者+资源配置者
组织模式	大团队、重制造、重供应链、重交付	核心模型团队更扁平, 强调研究效率与组织解耦
短期商业化路径	卖车、放量、毛利率、交付效率	提升硬件竞争力、Agent/Token 订阅、组织效率
成功后的估值含义	从手机公司升级为“人车家”生态公司	从硬件生态公司升级为 AI 平台/物理 AI 平台候选者

资料来源: 公司公告, 小米 2020、2023 年投资者大会, 中信建投证券

1.1 从“谨慎布局”到“全方位进军”，小米 AI 投入发生了什么转变？

图 1: 小米 AI 投入的转变



数据来源: 公司公告, 小米 2020、2023 年投资者大会, 小米官方公众号, 中信建投证券

第一阶段：长期打底，但战略位置未到最高级别。早在 2016 年，小米就组建 AI 视觉团队；2017 年 9 月成立小米 AI 实验室，逐步布局视觉、语音、声学、NLP、知识图谱、机器学习等能力。到 2023 年 4 月，小米正式组建专职大模型团队；截至 2023 年 8 月，小米人工智能团队已 3000 多人，覆盖 12 个技术领域、99 个细分赛道。在早期阶段，小米 AI 更多服务于手机影像、小爱同学、IoT 联动、语音交互等场景，本质是作为产品体验增强器存在。这一阶段的特点是：能力分散在各终端和业务条线之中，更多是“AI to 产品”，而非“AI as 公司主线”。

第二阶段：2023 年科技战略升级，AI 首次成为集团语言。真正的第一次跃升，是 2023 年 8 月 14 日。雷军在年度演讲中把科技战略总结为四条：深耕底层技术、长期持续投入、软硬深度融合、AI 全面赋能，并首次提出“软件 $\times$ 硬件”的战略公式。这一步的重要性在于，AI 从“研发部门的能力项”变成“公司未来叙事

的主语”。但当时的 AI 路线仍偏谨慎，主要强调“轻量化、本地部署、端侧优先”，典型载体是 MiLM 系列，目标更接近“在小米设备上让 AI 实用化”。也就是说，小米当时虽然已经明确 AI 重要，但并未完全转向“做全球一线基座模型”这条最重路径。

第三阶段：从端侧路线转向基座模型主战场。真正的第二次跃升发生在 2024 年下半年至 2025 年。雷军在 2025 年投资者日回顾时明确表示：“两年半前大模型在全球开始兴起的时候，实话实说，那个时候我们的决策就是专注 AI 应用……去年 7 月份，我们下决心进入大模型的主战场。”它直接说明小米 AI 投入的转变不是线性加码，而是技术路线的切换。这一节点之后，小米内部发生了至少三层变化。第一，从“用别人的大模型/做应用”转向“必须有自己的基座大模型”；第二，从端侧轻量化优先转向端云协同、通用大模型+Agent 优先；第三，从产品赋能导向转向平台能力导向。2024 年底，雷军亲自参与引入罗福莉，成为这一轮路线切换的标志性事件。罗福莉加入后，小米大模型从 MiLM 系列逐步切换到 MiMo 体系，模型更新速度明显加快。从 MiMo-7B 到 MiMo-V2-Flash，再到 MiMo-V2-Pro、MiMo-V2.5-Pro，短时间内完成从“入局者”到“全球开源第一梯队”的跨越。

第四阶段：AI 预算显性化，成为集团资源配置中心。到 2026 年，小米 AI 投入发生了更深层的变化：预算显性化、战略中心化、资源优先级上移。公司明确表示，2026 年 AI 领域至少投入 160 亿元，未来三年投入 600 亿元；同时，2026-2030 年研发投入将超过 2000 亿元。

图 2：小米集团聚焦 AI 战略，研发规模持续扩大



数据来源：公司公告，AI 投入金额为估算值，中信建投证券

AI 战略的启动方式与造车高度相似，但兑现节奏更快、资源耦合更深。造车时期，小米先是大额资本承诺，再逐步建团队、建工厂、推产品；AI 时期，小米也是先上升到集团战略，再通过人才、预算、模型发布节奏把资源迅速压上去。但与造车不同，AI 不需要等产线、渠道、产能完全建成才产生外溢价值，从一开始就能同步作用于手机、汽车、IoT 和企业效率，因此其对集团整体价值的穿透性更强。

1.2 AI 组织架构变化&创始人为核心领导的意义

1.2.1 雷军仍在一线且话语权高，意味着什么？

表 2: 雷军在造车与 AI 两大战略中的核心作用

维度	造车阶段	AI 阶段	含义
创始人角色	亲自带队、兼任 CEO	战略拍板者、人才招聘者、资源配置者	都深度参与，但参与方式不同
时间投入	雷军本人曾表示 1/2 时间用在汽车	更强调系统性组织能力，“专门靠我一个人肯定干不了”	AI 更依赖组织，而非个人单兵作战
团队组织	快速扩张、层级分明、工程化推进	Core 团队扁平化、无职级、研究驱动	反映赛道属性不同
创始人核心动作	招将、定产品、抓工厂、抓交付、抓舆论	定方向、挖关键人才、容忍长期投入、支持新组织形态	AI 更像“战略赋能”而非“业务代打”
对外表达	押上全部声誉，为小米汽车而战	未来三年 AI 投入 600 亿，AI/机器人/汽车/制造融合	都是创始人亲自背书的集团主线

资料来源：小米公众号，官方新闻稿，中信建投证券

第一，AI 这类重投入、慢回报战略，最怕“中途摇摆”，创始人带领的战略最重要的意义是“坚定”。造车和 AI 极大的相似之处在于：最大的组织风险不是技术难，而是公司能否在连续几年看不到清晰利润回报时，仍维持战略定力。这类事情如果只靠职业经理人推动，往往容易在预算会、KPI 考核、阶段性业绩压力中被稀释。小米的特殊性在于，雷军仍然是强势创始人，且对资本配置、重大技术路线、关键人才引进拥有极高话语权。罗福莉在访谈中最值得重视的一句话是：“雷总同意就行了。”这并不只是口语化表达，而是揭示了小米 AI 组织的真实运转逻辑：当技术路线存在高不确定性、很难用传统 KPI 衡量时，最终背书者是创始人而非层层流程。

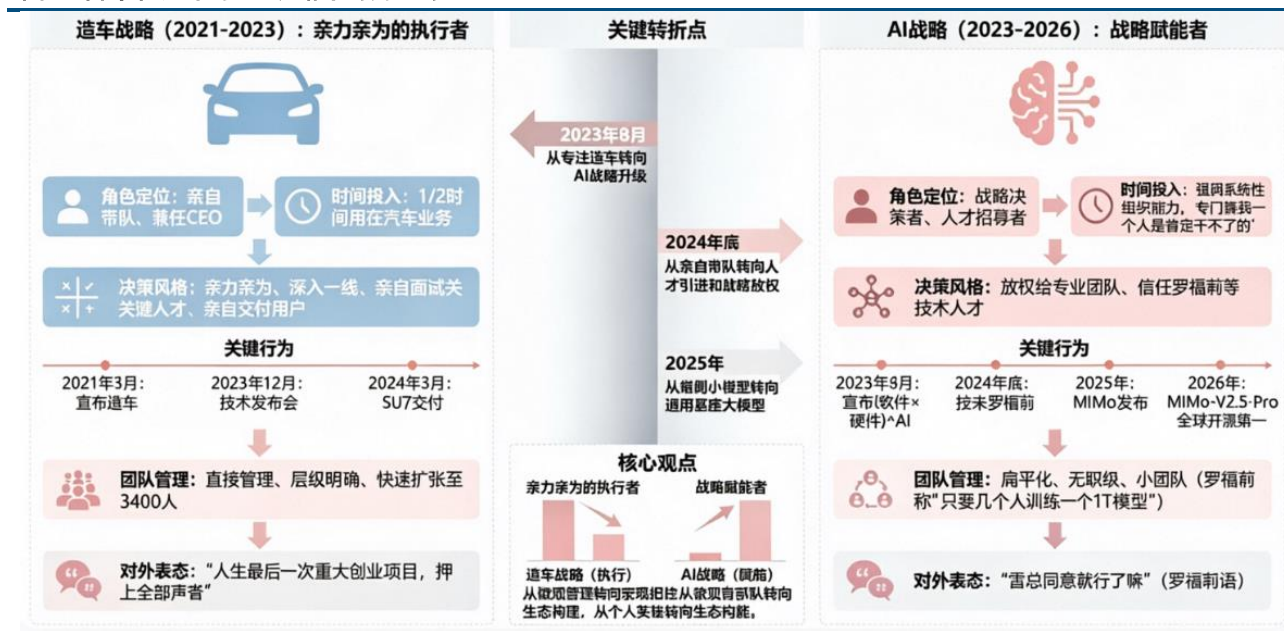
这和当年造车完全一致。雷军在 2025 年投资者日回顾造车时明确说，自己犹豫几个月后，最终决定亲自带队、志在必得；董事会与股东大会则承诺十年投资 650 亿元，从而让团队“心无旁骛”“避免短期行为”。这说明创始人高话语权最大的意义，是为高不确定性项目买来时间、耐心和组织豁免权。

第二，雷军的高话语权意味着：可以把“正确但短期不划算”的事做下去。AI 基座模型研发早期，ROI 很难量化。尤其对于小米这种原本以高效率、硬件周转著称的公司而言，做一个看上去短期不一定立刻变现、却要持续烧研发和算力的事情，本身就与历史基因存在不一致性。恰恰因此，雷军的一线话语权才更重要。2024 年底，他亲自参与人才招聘计划，把罗福莉招到小米，并在模型路线选择上与团队高度统一。从结果看，小米从 2024 年底确定通用大模型方向，到 2026 年 3-4 月 MiMo-V2-Pro 与 V2.5-Pro 跻身全球第一梯队，用时 15 个月左右，基本与全球核心玩家时间线一致。这种节奏如果没有创始人级别拍板、极可能在中间被无数次“保守化修正”。

第三，雷军在 AI 阶段的角色，已不同于造车时的“前线 CEO”，AI 阶段更值得重视的地方，是他在高话语权前提下选择放权给技术共同体。。造车时期，雷军是高度前置、强执行型的一线负责人。公司明确表示，雷军亲自带队，兼任智能电动汽车业务 CEO；他本人曾披露，自己 1/2 时间用在汽车业务，1/4 时间用在手机研发，1/4 关注集团运营和管理。在 2025 年投资者日中，雷军也明确回顾：造车是自己“犹豫了几个月”后决定亲自带队，并以“全力以赴，不留退路”方式推进。

AI 时期，小米大模型 Core 团队被描述为没有职级、没有小组、组织结构解耦，强调平权、创造力和研究氛围。这件事意义极大。因为造车更接近工程项目管理，强执行、强流程、强目标对齐更重要；而 AI 阶段，尤其是基座模型与研究组织，更适合由专业人才主导，创始人的核心价值转向给方向、给预算、给人、给战略安全边界。也就是说，雷军的话语权没有被用来强化控制，而是被用来建立一种适合 AI 研发的非传统组织形态。

图 3：雷军在造车与 AI 战略中的角色对比



数据来源：中信建投证券绘制

1.2.2 MiMo Core 团队：从“端侧 AI 能力补位”转向“集团基座模型中枢”

AI 团队组织定位的变化：MiMo Core 团队被放置于集团技术委员会之下，并给予罗福莉高度独立性，建立一个不隶属于单一终端部门的 AI 基座中枢；Plus 团队则承担更靠近应用和后训练的适配任务，形成研究突破与业务转化的分层。小米大模型团队最初于 2023 年 4 月组建，早期由栾剑带队，战略重点是轻量化与本地部署，这与当时小米强调端侧 AI、服务手机与 IoT 设备的路线高度一致。当时团队已有 6500 张 GPU 资源，并继续筹备自建 GPU 万卡集群，显示小米虽口径相对克制，但底层投入并不保守。真正的拐点出现在 2024 年底至 2025 年。罗福莉加入后，小米重组大模型体系：原负责人栾剑转而负责大模型 Plus 团队，不再从事基座模型预训练，而转向后训练与应用侧服务；罗福莉则组建并负责 MiMo Core 团队，专注基座模型与核心能力突破。

表 3: 小米大模型组织重构与分工

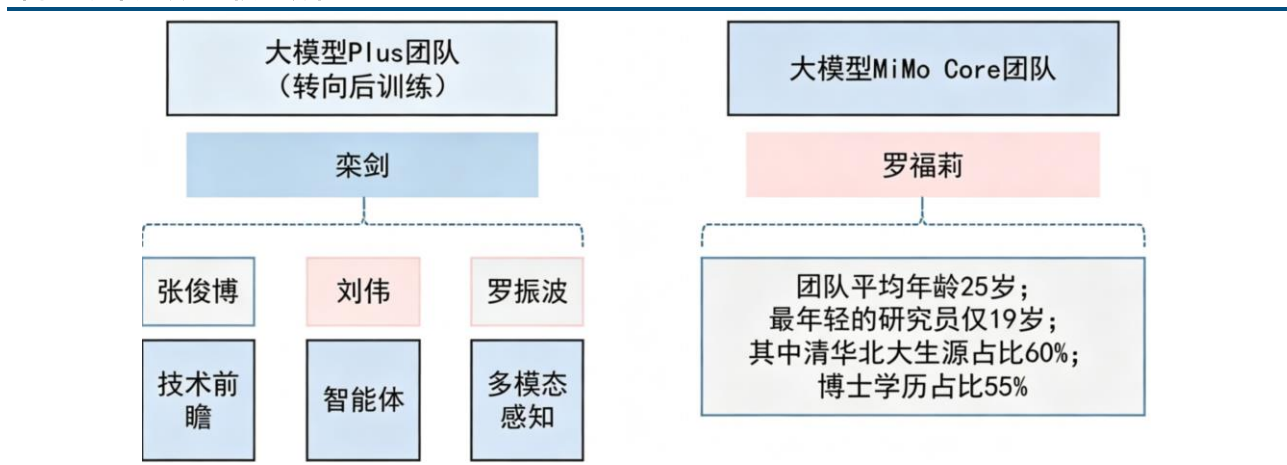
时间	组织动作	负责人	核心定位	组织含义
2023 年 4 月	组建大模型团队	栾剑	轻量化、本地部署	以端侧 AI 为主，服务手机/IoT 场景
2024 年底	罗福莉加入并组建 Core 团队	罗福莉	基座模型预训练、前沿能力突破	大模型从“能力补位”升级为集团级底座建设
2025—2026	原团队拆分为 Plus 与 Core 双线	栾剑、罗福莉	Plus 偏后训练与应用，Core 偏基座模型	形成“研究突破+业务适配”双轨制

资料来源：小米官方公众号，晚点，中信建投证券

Core 团队的独立性，是小米 AI 战略的关键信号，罗福莉拥有较高权限。MiMo Core 团队最值得关注的，不仅仅是“罗福莉是否加入”，而是她加入后拥有什么组织权限。Core 团队归属于集团技术委员会，不隶属于手机部或汽车部，由小米副总裁、集团技术委员会主席屈恒管理这条线，但罗福莉拥有很高独立性，并按自己的思路组建团队。屈恒本身是 2010 年加入小米的早期成员，现任副总裁兼集团技术委员会主席，同时分管集团质量委员会与信息技术部，2018—2023 年还担任生态链部总经理。这意味着 MiMo Core 并非挂靠于单一产品部门的实验室，而是被放置在一个能够跨手机、IoT、汽车进行技术统筹的集团级平台之下。

“无组、无强职级、强研究自由”，是 Core 团队最稀缺的组织资产。罗福莉表示，MiMo 团队目前已到 100 人规模，但这 100 人覆盖数据采集、数据质量、预训练、infra、后训练、开发、产品，以及语言、多模态、语音等方向；真正投入单代模型迭代的人数可能只有二三十人到三四十人。更关键的是，团队没有清晰分组、没有刚性 leader 边界，人员可在预训练与后训练间自然迁移，项目推动者也不对参与者拥有“绝对控制权”。

这类组织方式的价值在于，它高度匹配前沿模型研究的不确定性。罗福莉明确表示，团队没有严格 deadline，“模型训好了再发”；在训练中若发现问题，宁可停下数周排查，而不是为了发布时间硬推进。在大模型进入工程密集期后，这种“小团队极致追求正确性”的机制，实际上是对大厂常见 KPI 驱动研发方式的反向修正。

图 4: 小米大模型团队组织架构


数据来源：晚点公众号，中信建投证券

1.3 投入对比：非云业务视角下的投入强度加码

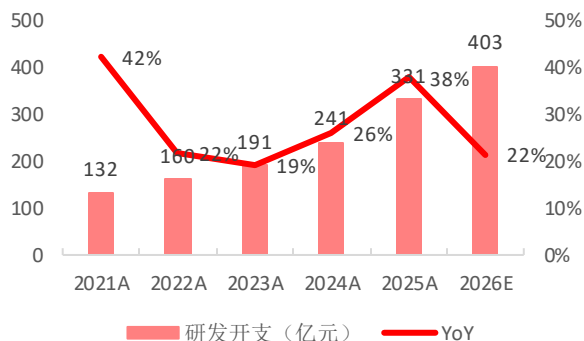
若仅以绝对资本开支规模比较，小米在 2025 年的 Capex 为 182 亿元，显著低于阿里约 1238 亿元、腾讯约 792 亿元以及字节约 1500 亿元，看似并不属于中国 AI 投入最激进的一梯队。但小米并非依赖公有云资本开支逻辑去堆算力，其投入强度更接近参考“产业化 AI 平台公司”。

小米的 Capex 结构：汽车主导，AI 加速。小米 2025 年全年资本开支达到 182 亿元，同比增长 73.2%，其中汽车分部 Capex 为 120 亿元，同比增长 196.1%，意味着全年约 66% 的资本开支直接投向智能电动车及相关创新业务。2024 年资本开支为 104.8 亿元，其中手机×AIoT 分部 64 亿元、汽车分部 41 亿元，资本重心由传统硬件向汽车与 AI 迁移。管理层明确表示，2025 年 Capex 增量主要来自汽车和 AI。小米的投入核心在于：一端是汽车工厂与智能制造，另一端是 AI 基础设施、万卡集群、自研模型与终端 Agent 能力。相比许多互联网公司把 Capex 投向统一云底座，小米更多是在搭建“人车家”生态所需的专用基础设施层。

公司明确表示：未来三年 AI 领域投入超过 600 亿元，其中 2026 年单年 AI 相关研发和资本投入超过 160 亿元；这 160 亿元包含当期研发费用、当期 Capex 以及历史 Capex 摊分三部分（当期研发费用超过 70%）。

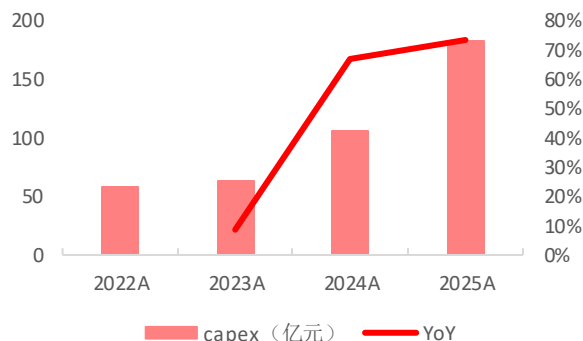
从研发上看，公司研发投入已从 2021 年 132 亿元、2022 年 160 亿元、2023 年 191 亿元、2024 年 241 亿元持续抬升，2025 年全年进一步达到 331 亿元，同比增长 37.8%，AI 及相关业务的比例约为 1/4（约 80 亿元）；过去五年（2021—2025 年）累计研发投入 1055 亿元，小米给出 2026—2030 研发投入超过 2000 亿元的中长期承诺。

图 5：小米历年研发费用



数据来源：公司公告，中信建投证券

图 6：小米历年 capex 投入



数据来源：公司公告，中信建投证券

二、基座模型：MiMo 万亿级模型，进入开源模型第一梯队

小米基座大模型的演进路径，核心不是简单追求参数扩张，而是沿着“端侧轻量化起步—云边端矩阵化扩展—开源推理基座突破—Agent 与全生态底座化”的路线持续升级。2023 年是技术立项与 MiLM 验证期，2024 年是 MiLM2 产品化落地期，2025 年是 MiMo 开源基座跃迁期，2026 年是 MiMo-V2/V2.5 冲击全球第一梯队的能力释放期。

表 4：小米基座大模型阶段性演进（2023-2026）

时间	代表模型/事件	关键参数/能力	阶段特征	战略含义
2023 年 4 月	大模型专职团队成立	团队由栾剑负责，雷军推动，投入“不设上限”	正式立项	标志小米从 AI 应用走向基座模型自研
2023 年 8 月	MiLM-1.3B/MiLM-6B 亮相	1.3B 端侧跑通，6.4B 在 C-Eval/CMMLU 同参数第一	端侧验证期	确立“轻量化、本地部署、端云协同”路线
2024 年 5 月	MiLM 完成备案	面向汽车、手机、智能家居开放	合规与量产准备期	从实验室走向规模化产品部署
2024 年 11 月	MiLM2 发布	0.3B-30B 矩阵，10 大能力平均提升超 45%，最长窗口 200k	云边端矩阵化	完成从单点模型到全场景矩阵的升级
2025 年 4 月	MiMo-7B 开源	7B 推理模型，25T tokens 训练，强化数学与代码	开源基座起点	小模型高推理效率，切入开源推理赛道
2025 年 12 月	MiMo-V2-Flash	309B 总参/15B 激活，256K 上下文，150 tok/s	开源旗舰期	以 MoE 架构实现速度、成本与 Agent 能力平衡
2026 年 3 月	MiMo-V2-Pro/Omni / TTS	1T+总参、42B 激活、1M 上下文	Agent 底座化	模型目标从对话升级到复杂任务执行
2026 年 4 月	MiMo-V2.5 / V2.5-Pro 开源	V2.5-Pro 1.02T/42B，V2.5 310B/15B，AA 开源并列第一	全球第一梯队冲刺	小米进入全球开源大模型第一梯队

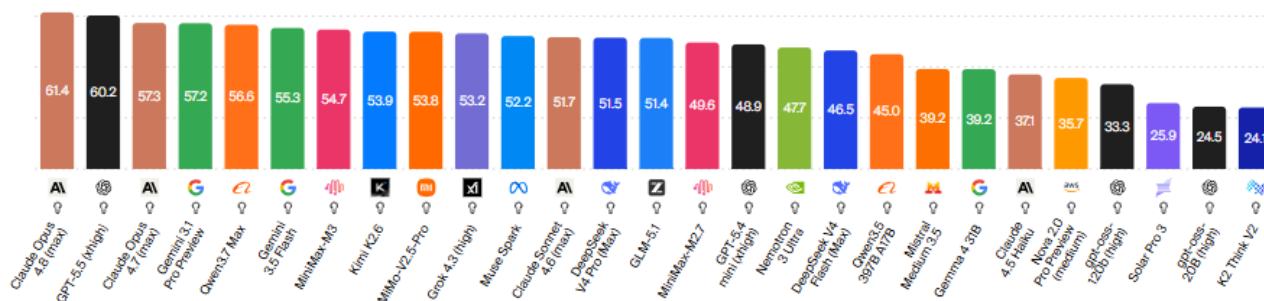
资料来源：小米官方公众号，中信建投证券

2026 年：MiMo-V2 与 V2.5 将小米推入全球第一梯队。2026 年是小米大模型能力集中释放的一年。2026 年 3 月，小米正式发布面向 Agent 时代的旗舰基座模型 MiMo-V2-Pro，总参数超过 1 万亿，激活参数 42B，支持 100 万 token 上下文。同期还发布 MiMo-V2-Omni 与 MiMo-V2-TTS，小米移动端 Agent Xiaomi miclaw 亦进入小范围封测，开始验证模型从“对话能力”向“系统级执行能力”的升级路径。

从指标上看，MiMo-V2-Pro 已不再是面向单一终端的模型，而是面向真实 Agent 工作场景优化的旗舰底座。资料显示，其在 Artificial Analysis 综合榜单一度位列全球第 8、国内第 2，并在 PinchBench、ClawEval 等 Agent 基准上接近 Claude Opus 4.6。MiMo-V2-Pro 此前以匿名模型“Hunter Alpha”在 OpenRouter 走红，说明其能力已被海外开发者在真实环境中验证，而不是仅停留在内部自测。

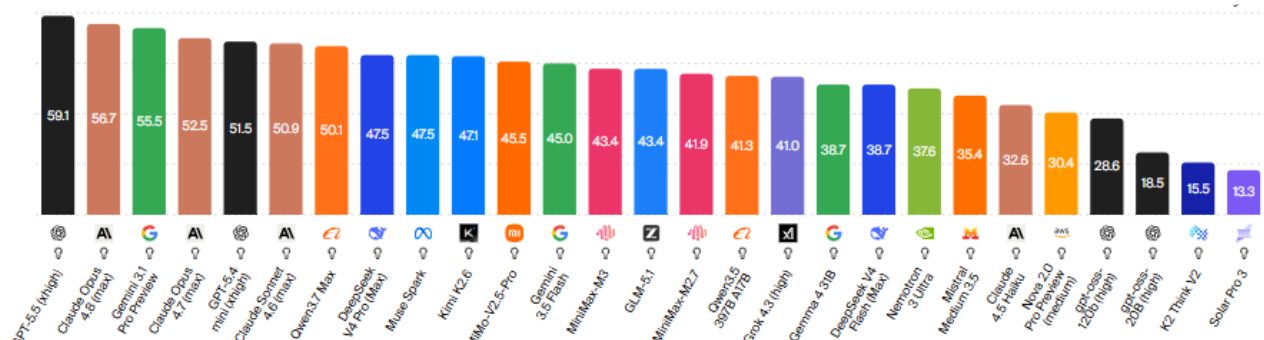
随后在 2026 年 4 月，小米进一步推出并开源 MiMo-V2.5-Pro 与 MiMo-V2.5。其中，V2.5-Pro 为 1.02T 总参数、42B 激活参数，基于混合注意力架构，预训练 27T tokens，支持 1M 上下文；V2.5 为 310B 总参数、15B 激活参数的原生全模态模型，训练 48T tokens。官方数据显示，MiMo-V2.5-Pro 在 Artificial Analysis 综合智能指数上以 54 分位列全球开源并列第一，并在 GDPVal-AA、ClawEval 等榜单中登顶。

图 7: Artificial Analysis Intelligence Index



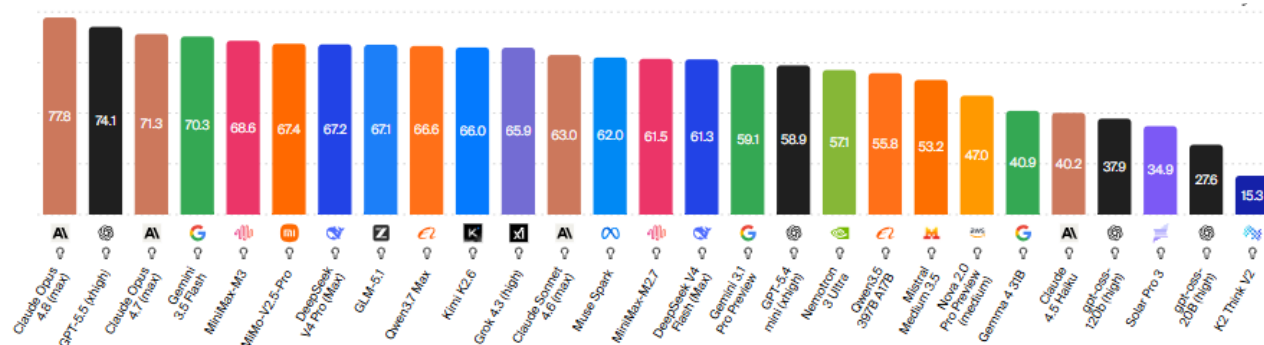
数据来源: Artificial Analysis, 中信建投证券 (截至2026.5.31)

图 8: Artificial Analysis Coding Index



数据来源: Artificial Analysis, 中信建投证券 (截至2026.5.31)

图 9: Artificial Analysis Agentic Index



数据来源: Artificial Analysis, 中信建投证券 (截至2026.5.31)

表 5: 2025-2026 年小米 MiMo 系列与行业头部水平对比

模型	发布时间	参数结构	上下文	主要优势	行业相对位置
MiMo-7B	Apr-25	7B	未明确	AIME 202555.4, 超 o1-mini	高效推理开源模型 突破口
MiMo-V2-Flash	Dec-25	309B 总参 / 15B 激活	未明确	开源 Top2, 代码/Agent 强, 115-150 tok/s	开源高性价比旗舰
MiMo-V2-Pro	Mar-26	1T 总参 / 42B 激活	1M	Agent、工具调用、长程规划	全球前十、国内前列
DeepSeek-V3	Dec-24	671B/37B 激活	128K 以上	数学、代码、性价比强	全球性价比标杆
Claude Opus 4.6	Feb-26	未披露	1M	顶级代码与 Agent	全球头部闭源模型
GPT-5.4	2026/3 前后	未披露	1M	通用能力、计算机使用	全球头部闭源模型

资料来源: 各公司官网, 官方公众号, 中信建投证券

从 **MiMo-7B** 到 **MiMo-V2-Pro**, 小米只用了不到一年时间就走完了许多厂商数年的参数扩张路径。2025 年 4 月 MiMo-7B 开源, 标志着小米正式从 MiLM 品牌切换到 MiMo 品牌; 2025 年 12 月 MiMo-V2-Flash 发布时, 已是 309B 总参数、15B 激活参数的 MoE 模型; 2026 年 3 月的 MiMo-V2-Pro 则直接跃迁至 1T+总参数、42B 激活参数, 总参数规模约为 Flash 的 3 倍以上, 并把上下文从 256k 扩展到 1M。

小米并没有在从 **Flash** 到 **Pro** 的过程中彻底推翻前一代技术路线, 而是延续“混合注意力+MTP+MoE”的主框架做规模放大。MiMo-V2-Flash 采用 5:1 的 Hybrid SWA 架构, 即每 5 层 SWA+1 层 GA 交替, 窗口大小 128, 在长上下文场景可将 KV Cache 和注意力计算压缩至传统全注意力的近 1/6; 配合 3 层 MTP, 推理可提速 2.0-2.6 倍, 单机吞吐达 5000-15000 token/秒, 单请求吞吐 115-150 token/秒。它证明了小米有能力把“参数不极致大”的模型做成“速度快、成本低、适配强”的 Agent 底座。

MiMo-V2-Pro 则将这条路线进一步推向“高强度 Agent 工作负载”。其核心参数为 1T+总参数、42B 激活参数、1M 上下文, 在 Artificial Analysis Intelligence Index 发布当周位列全球第八、国内第二; 在 OpenRouter 匿名测试阶段, “Hunter Alpha”一周突破 1 万亿 token 调用量, 后续累计处理量超过 2 万亿 token, 并在调用量榜单中居首。

但从技术上看, Pro 真正的提升不只来自参数放大, 更来自后训练范式。**MiMo-V2-Pro** 强调三项关键能力: 一是 MOPD 多教师在线策略蒸馏, 缓解大模型在代码、数学、Agent 等多域训练中的“跷跷板效应”; 二是真实环境的 agentic RL 训练, 覆盖四大类场景、超过 12 万真实交互环境; 三是与北大联合研发的 ARL-Tangram 训练基础设施, 提高资源弹性, 带来平均动作完成时间 4.3 倍改善、RL 训练速度最高 1.5 倍提升、外部资源节省最高 71%。

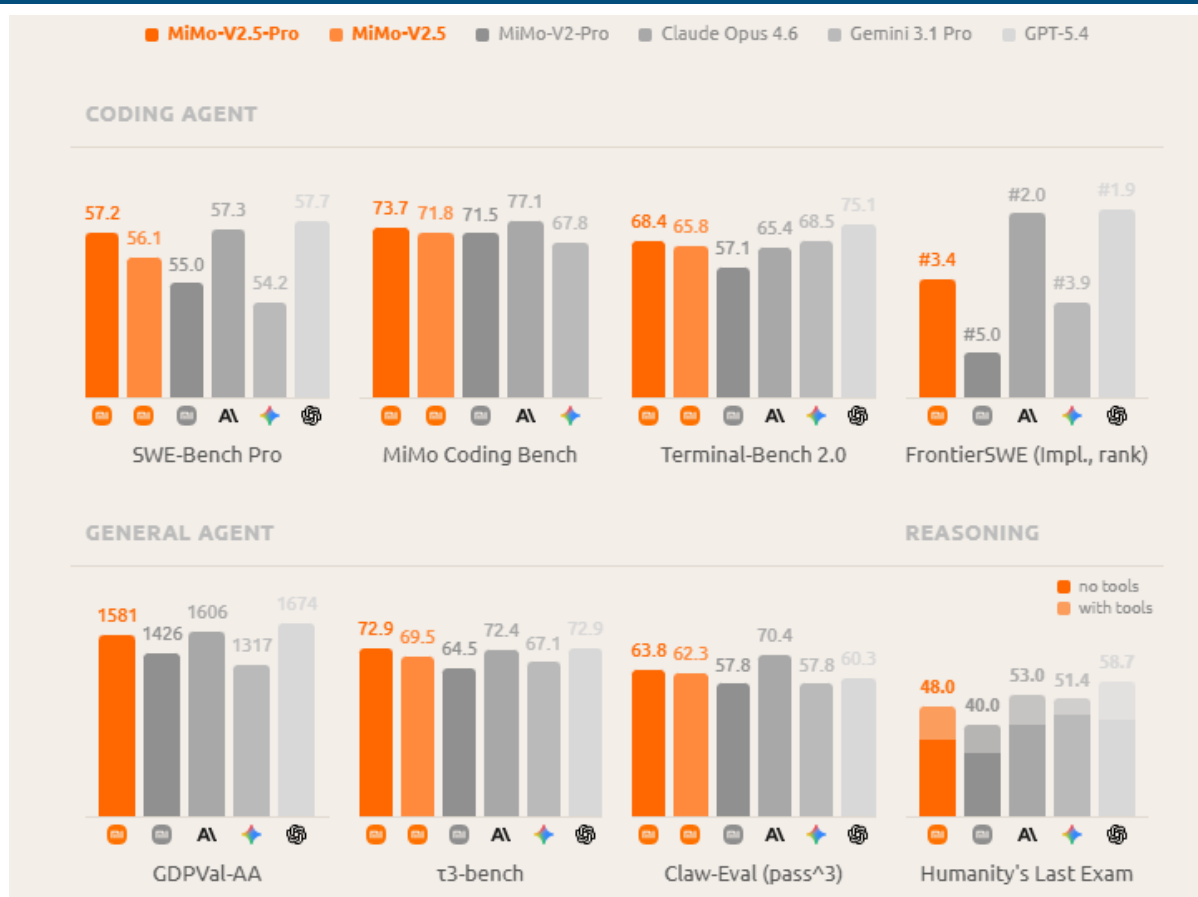
MiMo-V2.5-Pro 较 **MiMo-V2-Pro** 再次升级, 进入第一梯队。从结构上看, V2.5-Pro 仍采用 MoE 与混合注意力, 但官方模型卡显示其技术形态更清晰: 70 层 (1 层 dense+69 层 MoE), 其中 10 层全注意力、60 层 SWA, SWA/GA 交错比例为 6:1, KV Cache 在长上下文下可进一步压缩至传统全注意力的近 1/7, 且集成 3 层 MTP, 输出吞吐大致可提升 3 倍。

V2.5-Pro 相较 V2-Pro 最关键的变化, 其实不在参数, 而在训练目标与行为控制。V2.5-Pro 的提升主要来

自：训练数据升级、强化学习重构、工具调用优化、以及新增的“harness awareness”能力。所谓 harness awareness，即模型能够识别自己所处的 Agent 框架、CI 环境或测试脚手架，并据此调整输出格式、工具调用方式、错误处理路径和上下文组织方式，尽可能减少“模型输出本身正确、但在执行框架内失败”的问题。

这种变化直接反映在指标上。V2.5-Pro 在 ClawEval 上达到 63.8%/64%，单轨迹仅使用约 70k token，较 Claude Opus 4.6、Gemini 3.1 Pro、GPT-5.4 在相近能力水平下少用 40–60% token；较 Kimi K2.6 少用约 42% token。与此同时，V2.5-Pro 在 GDPVal-AA 上达到 1581，较 V2-Pro 的 1426 显著提升；在 Humanity’s Last Exam 上从 40.0 升至 48.0，SWE-Bench Pro 从 V2-Pro 的 55.0 升至 57.2。

图 10: MiMo-V2.5-Pro&MiMo-V2-Pro benchmark



数据来源：小米 MIMO 官网技术解析，中信建投证券

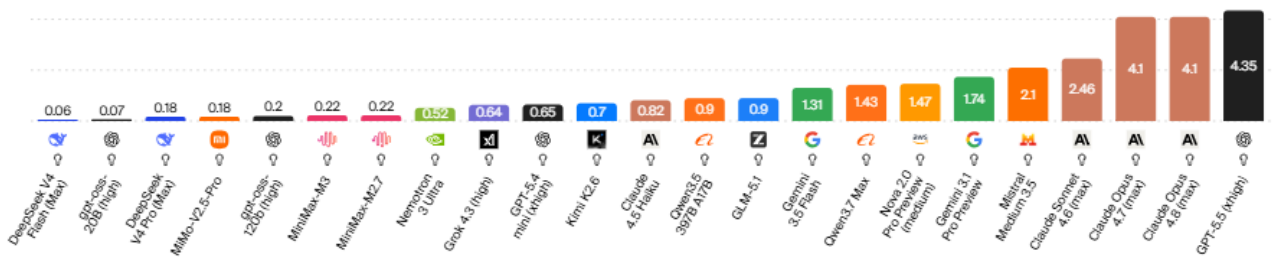
表 6: MiMo-V2.5-Pro 较 MiMo-V2-Pro 的核心提升

指标	MiMo-V2-Pro	MiMo-V2.5-Pro	变化
Artificial Analysis Intelligence Index	49	54	+5
Agentic Index	63	67	+4
GDPVal-AA	1426	1581	+155
SWE-Bench Pro	55.0	57.2	+2.2
Claw-Eval (pass^3)	57.8	63.8	+6.0
Humanity's Last Exam	40.0	48.0	+8.0

资料来源：小米 Mimo 官网技术解析，Artificial Analysis，中信建投证券

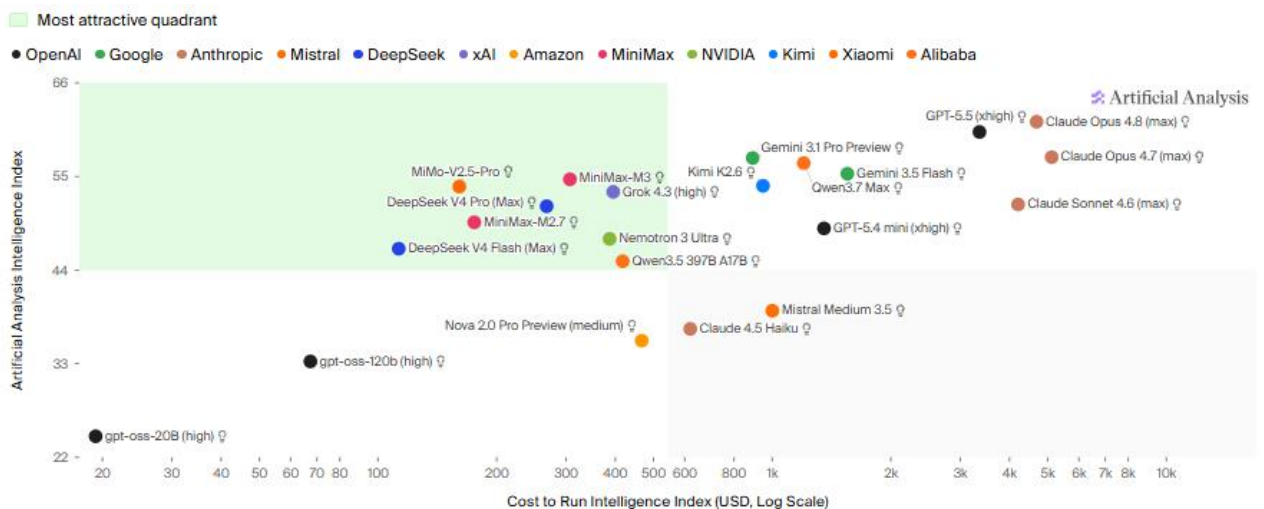
更重要的是其经济性和 Token 使用效率：在能力相近情况下，相比 Claude Opus 4.6、Gemini 3.1 Pro、GPT-5.4，可减少约 40%-60% token 消耗。对企业级调用场景而言，是直接影响商业化成本结构的核心变量。

图 11: Price of leading AI models-blended price



数据来源：Artificial Analysis，中信建投证券（截至2026.5.31）

图 12: Intelligence vs. Cost to Run Artificial Analysis Intelligence Index



数据来源：Artificial Analysis，中信建投证券（截至2026.5.31）

三、AI 应用落地方向：人车家全生态的 Agent 化进程

小米 2025—2026 年 AI 应用落地的核心，是把 Agent 从手机端的任务执行器，扩展为贯穿手机、汽车、家庭与机器人四类物理终端的统一调度中枢。这一进程的关键变化，在于 AI 不再停留在“问答式入口”，而是开始获得系统权限、跨端记忆、多模态感知、设备编排与物理执行能力，从而推动小米“人车家全生态”由互联网络升级为可被统一驱动的行动网络。

从落地节奏看，小米已经形成较清晰的三层推进路径：第一层是手机 Agent 化，以 Xiaomi miclaw 为代表，验证系统级执行闭环；第二层是车和家的空间 Agent 化，分别通过 XLA/OneVL 认知智驾与 Miloco 全屋智能方案，把 AI 从屏幕带入驾驶空间和居住空间；第三层则是具身执行端，借助 Xiaomi-Robotics-0、TacRefineNet 与工厂实习机器人，将 AI 最终延伸至真实工业动作执行。

图 13：小米围绕人车家全生态建设应用 AI 大模型与 Agent 技术的技术架构



数据来源：小米 2026 年投资者日，中信建投证券

3.1 手机端先行：miClaw 初步试水

2026 年 3 月，Xiaomi miClaw 启动小范围封测，首批支持小米 17 系列机型；4 月起封测扩展至平板、PC、Mac 和有屏音箱，并成为国内首批通过中国信通院手机端智能助手评估的产品。小米并没有先从云端聊天入口切入，而是直接把 Agent 嵌入系统层，验证“AI 能否代替用户完成真实任务”。

miClaw 的能力结构非常完整，官方将其概括为四层：系统底层能力、个人上下文理解、生态互联、自进化。其中最关键的是系统底层能力。miClaw 以系统应用身份运行，通过 ROM 推送安装，具备系统 UID 与系统分区权限，将手机能力封装为 50+ 系统工具和生态服务，并通过“用户输入—模型推理—工具调用—结果回传—继续推理”的循环闭环来完成任务。

与此同时，miClaw 并非简单工具总线，而是开始形成用户级状态记忆。其采用三级智能记忆管理，通过轮次压缩、token 压缩、本地缓存等方式，在连续执行 20 步复杂操作时仍能保留初始目标，且可节省 50%—90% 的 token 开销。这使其更接近“任务代理”而非“单轮助手”。

表 7: miclaw 基本能力

类别	代表能力
通信	短信读写、联系人搜索、通话记录查询
日历	日程读取、事件创建
文件	文件读写、内容搜索、图片加载
网络	联网搜索、网页内容抓取
系统	启动应用、精准定位、语音播报
系统工具	闹钟、录音机、笔记、相册等
智能家居	米家设备查询与控制
任务	定时任务、系统事件通知、闹钟
扩展	开放协议工具、技能、子智能体

资料来源：小米技术官方公众号，中信建投证券

假设一个场景，手机收到通知："购票成功，G1234 北京→成都 周五 08:30"。Miclaw 不仅仅是把出发日期、车次报给用户（这是很表层的东西），他会感知到这个事件，关联到用户实际会做什么，远高于普通 AI 能提供的能力。

表 8: 假设收到购票成功信息，miclaw 可能会做的步骤

步骤	AI 做了什么	调用的能力
感知	获取购票短信	获取短信事件
关联	查日历发现周五有会	日程读取
关联	查天气发现成都有雨	天气查询
关联	算家到车站通勤时间	定位服务
判断	推理得出：会要调、出发要提前、需要带伞	大模型推理
行动	建出差日程+设 6:30 闹钟	日程创建+闹钟
行动	出发日语音播报行程摘要	语音播报

资料来源：小米手机 miclaw 测评。中信建投证券

除此之外，miclaw 的生态互联能力也是核心重点。米家平台连接设备数已突破 10 亿台，miclaw 已实现完整米家协议客户端，可以读取设备状态、发送控制指令，并把设备能力翻译为模型可理解的自然语言描述。这意味着 miclaw 是以手机为入口的人车家生态 Agent 中枢雏形，在场景上，它已可以完成如提前开启车内空调、整理手机文件、联动家电设置家庭模式、会议期间自动静音与暂停扫地机等任务。

当然，当前 miclaw 仍处于验证期。小米明确提示其在稳定性、功耗表现、复杂场景执行成功率上仍需持续优化，且不推荐普通用户在主力机上升级。这说明小米现阶段对手机 Agent 的目标不是快速商业化，而是先验证系统级执行、权限治理与跨端记忆三件事。这与其后续将 miclaw 与超级小爱融合、并进一步深度改造 HyperOS/AIOS 的方向高度一致。

3.2 汽车：智驾与座舱是 AI 反哺业务最直接的利润场景

汽车是小米 AI 能力最容易形成估值重估的场景。原因在于，AI 对汽车的价值不仅体现在体验升级，更体现在高阶智驾、座舱交互、OTA 效率、品牌溢价与软件收入空间五个方面。

小米汽车的智驾演进节奏非常清晰：2023 年 12 月技术发布会首次公开变焦 BEV、超分辨 OCC、感知决策一体模型；2024 年 3 月 SU7 量产交付时已标配高速 NOA、主动安全、代客泊车、辅助泊车；2024 年 5 月启动 10 城城市 NOA 试点、8 月推进全国开通；2024 年 8 月底推出“全国都能开”先锋体验版；2024 年 9 月 OTA 1.3.0 在十城核心区域开放城市领航；到 2025 年 2 月 26 日，端到端全场景 HAD 实现全量推送，完成“车位到车位”贯通。2026 年 3 月，小米发布 XLA 认知大模型架构，并在新一代 SU7 上首发搭载；5 月又在其基础上发布并开源 Xiaomi OneVL，尝试统一 VLA、世界模型与潜空间推理三条路线。

XLA 的关键不只是“升级了智驾模型”，而是重新定义了车端 Agent 的输入与推理结构。XLA 中的“X”代表原生多模态输入，它把激光雷达、视觉、导航、声音、4D 毫米波雷达乃至机器人物理 AI 交互数据整合进同一认知框架。其底座模型是 Xiaomi MiMo-Embodied，该模型同时在辅助驾驶数据与室内机器人数据上训练，小米明确提出这是全球首个打通“辅助驾驶+具身机器人”的统一基座。

座舱侧，小米依托澎湃 OS 的天然优势，把“手机生态能力”复制进车内。SU7 搭载五屏联动、手机投屏、跨设备复制、后排 Pad 扩展等能力，车机可实现 1.49 秒极速启动、座舱升级最快 3 分钟、整车系统升级最快 30 分钟，这是传统车企很难具备的消费电子式迭代效率。

3.3 家庭空间 Agent 化，AI 将“可连接”推进到“可理解、可主动服务”

家庭场景是小米最具生态体量优势的 Agent 落地空间。2025 年 11 月，小米发布 Xiaomi Miloco (Xiaomi Local Copilot)，并在 2026 年 AWE 上进一步展示其“高智能的家”场景化能力。Miloco 的本质，是把智能家居的控制逻辑从“手动设规则”改造成“AI 根据视觉与上下文自主编排规则”。截至 2026 年 Q1 末，小米 AIoT 平台连接设备数首突破 11 亿台，拥有 5 件及以上连接设备用户达 2360 万，正是这种设备密度，赋予 Miloco 成为家庭 Agent 中枢的现实基础。

Miloco 的技术架构清晰：其以米家摄像头为视觉信息源，以 MiMo-VL-Miloco-7B 端侧视觉语言模型为核心，将任务拆为规划与视觉理解两个阶段，并通过 MCP 协议打通米家生态、Home Assistant 及第三方 IoT 平台。与传统“温度>30℃开空调”式机械规则不同，Miloco 允许用户通过自然语言定义模糊目标，再由模型完成规则生成与设备调用。

2026 年 AWE 展示了其最具代表性的主动服务逻辑：当系统识别到用户盖上被子、放下手机等动作后，可自动执行关灯、拉窗帘、调低空调风速等多设备联动；也可在“有人读书”时自动开灯、“小孩玩手机”时联动音箱提醒、“有人睡觉但未盖被子”时调整空调温度。家庭空间有望实现视觉感知+语义理解+设备执行的完整闭环。

表 9: Miloco 相对传统智能家居规则引擎的能力跃迁

比较维度	传统智能家居	Xiaomi Miloco
触发方式	手工设定阈值和条件	自然语言定义意图 + 视觉理解自动推理
感知能力	依赖单点传感器	以摄像头为核心，具备全屋视觉问答与事件识别
规则生成	用户手动配置	模型自动生成复杂联动规则
隐私处理	多依赖云侧	端侧视频理解，视觉数据可不外流
生态扩展	受限于单平台协议	支持米家、HomeAssistant、第三方 IoT 接入
服务模式	被动响应	主动服务、上下文感知、家庭记忆方向演进

资料来源：小米官方公众号，中信建投证券

Miloco 的战略含义有两层。第一，它是 MiMo 在真实生活空间的首个应用落地，标志着小米把大模型能力从数字内容延伸到物理空间。第二，它为未来与 miclaw 的融合预留了接口。投资者日披露，miclaw 也会赋能 Miloco，多模态感知+家庭记忆，未来存在一个中枢贯通全域设备，实现共享记忆、跨端调度与主动智能。这意味着 Miloco 并非独立产品，而是家庭空间中的子 Agent，未来会被统一纳入更上层的全生态 Agent 框架中。

3.4 机器人与制造：AI 外溢到物理世界，是长期想象力所在

如果说手机、汽车、家居解决的是“屏幕内”和“车内”的交互效率，那么机器人与制造则对应小米 AI 长期最具想象力的方向，即物理世界执行能力-让 Agent 拥有自主物理操作能力 。小米在人形机器人与四足机器人方向早在 2021 年发布 CyberDog、2022 年发布 CyberOne 时就开始布局。到 2026 年，小米已明确把具身智能视作 MiMo 的重要延展。

2026 年 2 月，小米发布触觉驱动的精细抓取模型 TacRefineNet 与首代机器人 VLA 大模型 Xiaomi-Robotics-0；3 月，人形机器人 CyberOne 已进入小米汽车工厂。从实际数据上看，在自攻螺母上件工站实现 3 小时持续作业、双侧同时安装成功率 90.2%、满足 76 秒产线节拍。

图 14: 小米最新一代机器人（2026 投资者日）



数据来源：小米 2026 投资者日 PR，中信建投证券

从模型侧看，Xiaomi-Robotics-0 是一个 47 亿参数具身 VLA 模型，在 LIBERO、CALVIN、Simpler Env 三大仿真 Benchmark 中相对 30 个对比模型取得 SOTA 成绩。其核心设计是用 MoT 架构将 VLM 与多层 Diffusion Transformer 解耦，使视觉语言理解与连续动作生成分开处理，从而降低真机推理延迟。

硬件侧，小米灵巧手路线也具有鲜明工程特征。2026 年 3 月底，小米披露其灵巧手采用全掌触觉采集手套、仿生汗腺液冷散热、抓握寿命 15 万次以上等方案，其中仿生汗腺系统每分钟蒸发 0.5mL 水、提供约 10W 主动散热能力。

图 15: Xiaomi-Robotics-0 模型与真机后训练流程示意图

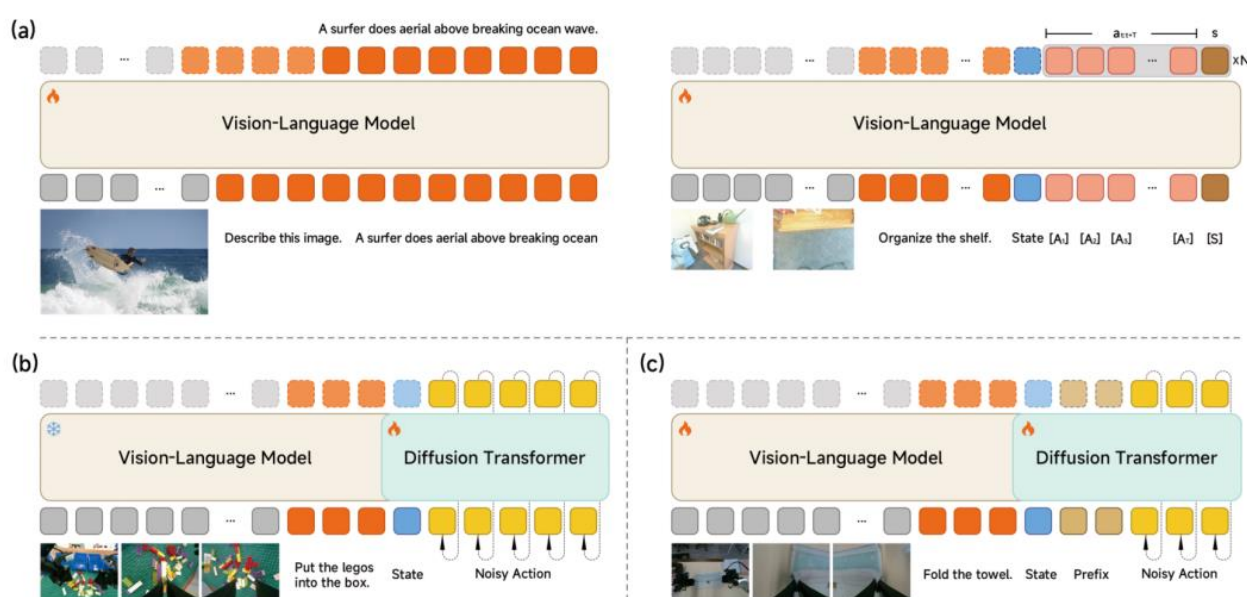


Figure 3 Model & Training. (a) During the first step of pre-training, we train the VLM on both vision-language data (left) and robot trajectory data (right). Vision-language data are trained via a next-token-prediction objective. We adopt the training paradigm in Choice Policies [51] to train the VLM for action prediction on the robot trajectory data. (b) In the second step of pre-training, we freeze the VLM and train the diffusion transformer for generating actions via flow-matching. (c) During post-training for asynchronous execution, we prepend clean action prefix to the noisy action tokens.

数据来源：论文《Xiaomi-Robotics-0: An Open-Sourced Vision-Language-Action Model with Real-Time Execution》，中信建投证券

机器人落地仍处于早期，目前进展“只是开始”，未来 3-5 年才可能实现工厂全面落地，管理层表示，未来五年希望有大批量具身机器人进入小米工厂工作。因此，具身智能对小米短期业绩贡献有限，但对 AI 战略估值意义较大，因为它把小米与大多数纯模型公司区分开来：后者停留在数字空间，小米则正沿着手机—汽车—家庭—工厂的链条持续进入物理世界。

投资建议与盈利预测

造车阶段，市场给小米加估值，更多是因为它多了一条大赛道和第二增长曲线。AI 阶段，如果市场真正买账，改变的不只是某个分部估值，而是整个集团的属性认知。如果市场认知到：小米不是硬件拼盘，而是一个能用 AI 重新组织终端生态价值的公司，那么传统 SOTP 并不完整，而是应该转向“整体平台重估”。

表 10: 造车与 AI 对小米估值框架的影响差异

维度	造车带来的变化	AI 带来的变化
对收入结构	新增汽车收入与毛利贡献	可能提升各业务收入质量与服务收入占比
对利润表	先亏后盈，路径较清晰	先压利润，回报分散体现在毛利、费用率、订阅与 ASP
对估值模型	给予汽车业务 PS/DCF 溢价	可能推动整体从硬件 PE 向平台型溢价迁移
对市场叙事	“手机公司成功跨界造车”	“物理 AI 平台公司雏形”

资料来源：中信建投证券绘制

AI 最现实的回报路径，未必是订阅，而是先体现在产品力与组织效率。从商业模式上，小米现阶段并不急于把 AI 单独变现，当前更关注模型、应用与生态能力成熟，Token Plan、订阅制、高端产品溢价、API 服务等均在探索中，但尚未将其设定为核心 KPI。雷军在 2025 年投资者日表示：如果大模型做到数一数二、让体验更好，哪怕多赚一个点毛利，商业模式就闭环了。这说明小米内部并不把 AI 短期 KPI 设定为“先做出一个大订阅生意”，而是先把 AI 作为硬件溢价工具、生态黏性工具与内部效率工具。这个逻辑和造车初期也很像。造车前期不是先追求利润，而是先证明产品定义能力、品牌拉升能力、用户迁移能力；AI 同理，先证明手机、车、家居因 AI 而明显更好用，再谈大规模直接收费。

AI 比造车更难短期定价，但长期想象空间更大。造车的好处是可见度更高，销量、ASP、毛利率都能跟踪；AI 的难点在于大量价值短期难进入报表，所以更容易出现“看得到战略，看不到业绩”的定价错配。但也正因为如此，一旦 AI 真正打通系统级入口、订阅和硬件溢价，其整体重估幅度可能高于单一汽车业务。小米已经成为少数同时具备自研模型、系统权限、终端入口、制造场景的全栈型 AI 玩家。

尽管短期传统主业在大周期下承压，存储的价格仍在快速上涨，成本周期拐点可能在 2027 年以后，但受益于前期存储备货节奏、公司主动调整产品结构、以及 4 月起部分产品涨价拉动，手机毛利率预计得到较好控制，此外汽车新品周期+产能爬坡有望形成催化。另一方面，公司仍然积极投入 AI，在更长远的时间看将是最核心的竞争力之一。

预计小米集团 2026/2027 年收入 4846/6077 亿元(yoy+6%/+25%)，调后净利润 286/393 亿元(yoy-27%/+37%)，我们给予公司核心主业 26 年 20x PE，汽车 26 年 2.5x PS，目标价 38.7 港币，给予“买入”评级。

表 11: 小米业绩拆分预测 (百万元)

1810. HK CNY, mm	2021A	2022A	2023A	2024A	2025A	2026E	2027E
总营收	328,309	280,044	270,970	365,906	457,287	484,571	607,676
YOY	34%	-15%	-3%	35%	25%	6%	25%
分部收入							
手机	208,869	167,217	157,461	191,759	186,440	168,108	198,985
YOY	37%	-20%	-6%	22%	-3%	-10%	18%
IoT	84,980	79,795	80,108	104,104	123,200	133,088	146,397
YOY	26%	-6%	0%	30%	18%	8%	10%
互联网	28,212	28,321	30,107	34,115	37,440	39,063	42,067
YOY	19%	0%	6%	13%	10%	4%	8%
汽车				32,728	106,070	138,970	214,619
毛利润	58,261	47,577	57,476	76,560	101,806	101,692	123,960
YOY	59%	-18%	21%	33%	33%	0%	22%
毛利率	17.7%	17.0%	21.2%	20.9%	22.3%	21.0%	20.4%
手机毛利率	11.9%	9.0%	14.6%	12.6%	10.9%	8.0%	8.3%
IoT毛利率	13.1%	14.4%	16.3%	20.3%	23.1%	23.1%	23.0%
互联网毛利率	74.1%	71.8%	74.2%	76.6%	76.5%	75.3%	76.0%
汽车毛利率				18.5%	24.3%	20.7%	20.0%
经营费用	38,887	42,465	43,451	55,041	73,000	83,084	93,886
% Rev	11.8%	15.2%	16.0%	15.0%	16.0%	17.1%	15.5%
销售开支	20,981	21,323	19,227	25,389	33,214	35,559	40,107
% Rev	6.4%	7.6%	7.1%	6.9%	7.3%	7.3%	6.6%
研发开支	13,167	16,028	19,098	24,051	33,132	40,257	45,576
% Rev	4.0%	5.7%	7.0%	6.6%	7.2%	8.3%	7.5%
行政开支	4,739	5,114	5,127	5,601	6,653	7,267	8,204
% Rev	1.4%	1.8%	1.9%	1.5%	1.5%	1.5%	1.4%
净利润	19,283	2,503	17,474	23,578	41,567	21,132	30,017
YOY	-5.1%	-87.0%	598.1%	34.9%	76.3%	-49.2%	42.0%
净利率	5.9%	0.9%	6.4%	6.4%	9.1%	4.4%	4.9%
调整后净利润	22,040	8,518	19,272	27,234	39,167	28,577	39,270
YOY	69.4%	-61.3%	126.2%	41.3%	43.8%	-27.0%	37.4%
调整后净利率	6.7%	3.0%	7.1%	7.4%	8.6%	5.9%	6.5%
调整后主业净利润	-	11,618	25,922	33,534	38,237	28,062	30,346

数据来源: 公司公告, 中信建投证券预测

表 12: 小米三表预测 (百万元)

资产负债表						利润表					
	2023	2024	2025	2026E	2027E		2023	2024	2025	2026E	2027E
总流动资产	199,053	225,709	254,811	294,550	309,347	营业收入	270,970	365,906	457,287	484,571	607,676
总非流动资产	125,195	177,447	253,285	275,663	302,616	营业成本	213,494	289,346	355,481	382,878	483,716
总资产	324,247	403,155	508,096	570,213	611,963	销售费用	19,227	25,389	33,214	35,559	40,107
总流动负债	115,588	175,385	192,405	229,063	223,875	管理费用	5,127	5,601	6,653	7,267	8,204
总非流动负债	44,398	38,565	49,367	50,987	52,435	研发费用	19,098	24,051	33,132	40,257	45,576
总负债	159,986	213,950	241,773	280,050	276,310	投资公允价值变动	3,501	1,051	13,312	849	0
股本	1.41	0.41	0.43	0.43	0.43	分占按权益法入账之投资亏损	46	277	326	175	200
储备	163,995	188,737	266,218	290,163	335,653	其他收入	740	1,667	2,737	1,985	2,130
优先股	0	0	0	0	0	其他收益	1,697	-10	2,720	1,286	1,332
归属母公司股东权益	163,995	188,738	266,219	289,897	335,386	经营利润	20,008	24,502	47,901	22,904	33,736
少数股东权益	266	467	105	266	267	财务收入净额	2,003	3,624	1,746	2,116	2,116
股东权益合计	164,262	189,205	266,323	290,163	335,653	可转换可赎回优先股	0	0	0	0	0
总权益及负债	324,247	403,155	508,096	570,213	611,963	除所得税前利润	22,017	28,127	49,648	25,020	35,852
						所得税	-4,537	-4,548	-8,081	-3,888	-5,835
						净利润	17,474	23,578	41,567	21,132	30,017
						调后净利润	19,272	27,234	39,167	28,577	39,270
						调后核心业务净利润	25,922	33,534	38,237	28,062	30,346
现金流量表						关键比率					
经营活动所得现金净额	41,300	39,295	34,142	32,136	35,786	营业收入YOY	-3%	35%	25%	6%	25%
投资活动所得现金净额	-35,169	-35,386	-71,679	-12,622	-15,136	净利润 YOY	598%	35%	76%	-49%	42%
融资活动所得现金净额	-505	-3,999	30,766	2,352	2,265	GM	21.2%	20.9%	22.3%	21.0%	20.4%
汇率变动	398	120	23	0	0	OPM	7.4%	6.7%	10.5%	4.7%	5.6%
现金及现金等价物变动	27,607	33,631	33,661	26,914	48,780	NPM	6.4%	6.4%	9.1%	4.4%	4.9%
年末现金及现金等价物	33,631	33,661	26,914	48,780	71,695						

数据来源: 公司公告, 中信建投证券预测

风险分析

新产品出货量不及预期；手机市场下行风险，大盘出货量下跌，新品推广不及预期的风险；存货减值拨备的风险；政策管理风险，国内互联网业务发展受限；海外产品竞争加剧等；IoT 周期属性较强，需求端疲软；宏观经济下行导致 3C 类产品需求下降；欧洲能源紧缺的影响；汇率波动风险；地缘政治风险导致部分国家业务受损；互联网业务受宏观经济影响，广告业务持续受到冲击；造车等新产品业务造成的研发费用呈长期上升态势，带来利润端压力。

分析师介绍

于伯韬

海外研究首席分析师，FRM，香港大学金融学硕士，2020 年加入中信建投海外研究团队，2018-2020 年就职于长江证券研究所海外团队，2025/2024/2021/2020 年新财富港股及海外方向第五名，2022 年新浪金麒麟港股及海外市场最佳分析师第三名，2020 年新浪金麒麟港股及海外市场新锐分析师第一名。

评级说明

投资评级标准		评级	说明
报告中投资建议涉及的评级标准为报告发布日后6个月内的相对市场表现，也即报告发布日后的6个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数作为基准；新三板市场以三板成指为基准；香港市场以恒生指数作为基准；美国市场以标普500指数为基准。	股票评级	买入	相对涨幅 15%以上
		增持	相对涨幅 5%—15%
		中性	相对涨幅-5%—5%之间
		减持	相对跌幅 5%—15%
		卖出	相对跌幅 15%以上
	行业评级	强于大市	相对涨幅 10%以上
		中性	相对涨幅-10-10%之间
		弱于大市	相对跌幅 10%以上

分析师声明

本报告署名分析师在此声明：（i）以勤勉的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，结论不受任何第三方的授意或影响。（ii）本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

法律主体说明

本报告由中信建投证券股份有限公司及/或其附属机构（以下合称“中信建投”）制作，由中信建投证券股份有限公司在中华人民共和国（仅为本报告目的，不包括香港、澳门、台湾）提供。中信建投证券股份有限公司具有中国证监会许可的投资咨询业务资格，本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格证书编号已披露在报告首页。

在遵守适用的法律法规情况下，本报告亦可能由中信建投（国际）证券有限公司在香港提供。本报告作者所持香港证监会牌照的中央编号已披露在报告首页。

一般性声明

本报告由中信建投制作。发送本报告不构成任何合同或承诺的基础，不因接收者收到本报告而视其为中信建投客户。

本报告的信息均来源于中信建投认为可靠的公开资料，但中信建投对这些信息的准确性及完整性不作任何保证。本报告所载观点、评估和预测仅反映本报告出具日该分析师的判断，该等观点、评估和预测可能在不发出通知的情况下有所变更，亦有可能因使用不同假设和标准或者采用不同分析方法而与中信建投其他部门、人员口头或书面表达的意见不同或相反。本报告所引证券或其他金融工具的过往业绩不代表其未来表现。报告中所含任何具有预测性质的内容皆基于相应的假设条件，而任何假设条件都可能随时发生变化并影响实际投资收益。中信建投不承诺、不保证本报告所含具有预测性质的内容必然得以实现。

本报告内容的全部或部分均不构成投资建议。本报告所包含的观点、建议并未考虑报告接收人在财务状况、投资目的、风险偏好等方面的具体情况，报告接收者应当独立评估本报告所含信息，基于自身投资目标、需求、市场机会、风险及其他因素自主做出决策并自行承担投资风险。中信建投建议所有投资者应就任何潜在投资向其税务、会计或法律顾问咨询。不论报告接收者是否根据本报告做出投资决策，中信建投都不对该等投资决策提供任何形式的担保，亦不以任何形式分享投资收益或者分担投资损失。中信建投不对使用本报告所产生的任何直接或间接损失承担责任。

在法律法规及监管规定允许的范围内，中信建投可能持有并交易本报告中所提公司的股份或其他财产权益，也可能在过去 12 个月、目前或者将来为本报中所提公司提供或者争取为其提供投资银行、做市交易、财务顾问或其他金融服务。本报告内容真实、准确、完整地反映了署名分析师的观点，分析师的薪酬无论过去、现在或未来都不会直接或间接与其所撰写报告中的具体观点相联系，分析师亦不会因撰写本报告而获取不当利益。

本报告为中信建投所有。未经中信建投事先书面许可，任何机构和/或个人不得以任何形式转发、翻版、复制、发布或引用本报告全部或部分内容，亦不得从未经中信建投书面授权的任何机构、个人或其运营的媒体平台接收、翻版、复制或引用本报告全部或部分内容。版权所有，违者必究。

中信建投证券研究发展部

北京
朝阳区景辉街 16 号院 1 号楼 18 层
电话：（8610）56135088
联系人：李祉瑶
邮箱：lizhiyao@csc.com.cn

上海
上海浦东新区浦东南路 528 号南塔 2103 室
电话：（8621）6882-1600
联系人：翁起帆
邮箱：wengqifan@csc.com.cn

深圳
福田区福中三路与鹏程一路交汇
汇电金融中心 35 楼
电话：（86755）8252-1369
联系人：曹莹
邮箱：caoying@csc.com.cn

中信建投（国际）

香港
中环交易广场 2 期 18 楼
电话：（852）3465-5600
联系人：刘泓麟
邮箱：charleneliu@csci.hk