

汽车行业深度报告

AI 系列深度 21 期: 物理 AI 如何从架构借鉴走向模型迁移?

增持 (维持)

2026 年 08 月 07 日

证券分析师 黄细里

执业证书: S0600520010001

021-60199793

huangxl@dwzq.com.cn

研究助理 童明祺

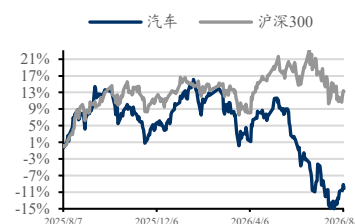
执业证书: S0600125080005

tongmq@dwzq.com.cn

投资要点

- **物理 AI 正建立在数字 AI 能力外溢的基础之上, 且借鉴层级正在上移。**早期迁移主要发生在架构理念层面, Transformer 注意力机制、BEV、端到端等方法被引入驾驶感知、轨迹预测和一体化规划; 近年则开始向模型层面迁移, Waymo EMMA 构建于 Gemini 之上, 理想 MindVLA、小鹏世界基座模型等车端 VLA 实践, 均体现数字 AI 基础模型能力向物理 AI 扩展。
- **可迁移红利并非线性外推, 而会随任务靠近真实物理世界而收窄。**越偏向语义理解、规划和多模态表征, 数字 AI 既有模型与训练范式的可复用性越强; 越接近实时三维感知、接触控制、安全攸关决策和高物理闭环, 直接迁移比例越低。总体看, 数字 AI 提供能力底座, 物理 AI 仍必须补足物理世界表征、真实数据和安全闭环。
- **物理 AI 相较数字 AI 的第一项系统差异在表征。**文本天然离散且带语义, 数字 AI 输入具备较强结构化基础; 物理世界则由连续光子、点云、运动和接触构成, 必须先被编码为可学习的三维空间表征。VectorNet、BEV+Transformer、Occupancy、三维高斯等技术, 本质上都是对物理世界进行模型可读表征的不同路径。
- **第二项差异在数据。**数字 AI 训练语料大量沉淀在互联网中, Scaling 核心是抓取、清洗和压缩已有数据; 物理 AI 最有价值的长尾、事故和极端场景天然稀疏、昂贵且危险, 单靠真实路测难以覆盖。世界模型因此被推向“数据工厂”位置, 用于重建、生成和放大稀缺场景, 为策略训练与安全验证提供可控环境。
- **第三项差异在学习范式。**数字 AI 模仿的是互联网级人类知识, 模仿上限对应超人类广度, 强化学习更多承担对齐和推理增强功能; 物理 AI 模仿对象常是平均人类司机, 而安全目标要求显著超越人类, 单纯模仿的上限更低。强化学习在高物理闭环任务中更像突破能力边界的关键环节, 而非后置修补。
- **我们认为, 数字 AI 底座模型发展路径已逐渐清晰、价值在于确定性, 但其难以实现物理世界的建模闭环; 我们认为物理 AI 将成为物理世界的 AI 解决方案, 从而在 AI 的当前发展阶段成为增量更大的方向, 智能驾驶作为最先跑通闭环的代表场景, 我们认为应重点关注。**
- **风险提示:** 技术迭代不及预期的风险, 大模型厂商 ARR 增速放缓的风险, 云服务商资本开支不及预期的风险, 智能驾驶及机器人商业化落地不及预期的风险。

行业走势



相关研究

《AI 系列深度 20 期: 如何看物理场景中数字 AI 的能力边界?》

2026-08-07

《AI 系列深度 18 期: 生成智能如何引入 Transformer?》

2026-08-07

内容目录

- 1. 数字 AI 能力向物理 AI 迁移，物理闭环决定兑现质量4
 - 1.1. 从架构理念到模型层迁移：借鉴层级持续上移.....4
 - 1.2. 越接近物理闭环，直接迁移红利越收窄5
- 2. 物理 AI 相较数字 AI：表征、数据与学习范式三重差异5
 - 2.1. 表征差异：物理世界表征是必经门槛.....5
 - 2.2. 数据差异：高价值数据稀疏，世界模型承担数据工厂角色.....6
 - 2.3. 范式差异：强化学习成为物理 AI 关键环节7
- 3. 结论：数字 AI 提供能力底座，物理 AI 补足物理闭环8
- 4. 风险提示9

图表目录

图 1: 物理 AI 对数字 AI 借鉴层级的跃迁4

图 2: 物理世界可学习表征演进.....6

图 3: 自动驾驶安全性所需测试里程.....7

图 4: 模仿学习上限的差异：数字 AI 与物理 AI.....8

表 1: 数字 AI 关键成果向智能驾驶的迁移路径5

表 2: 主流物理世界表征范式对比.....6

表 3: 世界模型作为数据工厂的代表项目7

1. 数字 AI 能力向物理 AI 迁移，物理闭环决定兑现质量

物理 AI Physical AI 通常指能够感知物理世界、并在三维空间中完成决策与行动的人工智能系统，自动驾驶是其最具代表性的落地场景之一。黄仁勋在 2025 年 CES 主题演讲中，将人工智能发展划分为感知 AI、生成式 AI、智能体 AI 与物理 AI 四个阶段，并提出通用机器人的 ChatGPT 时刻临近。

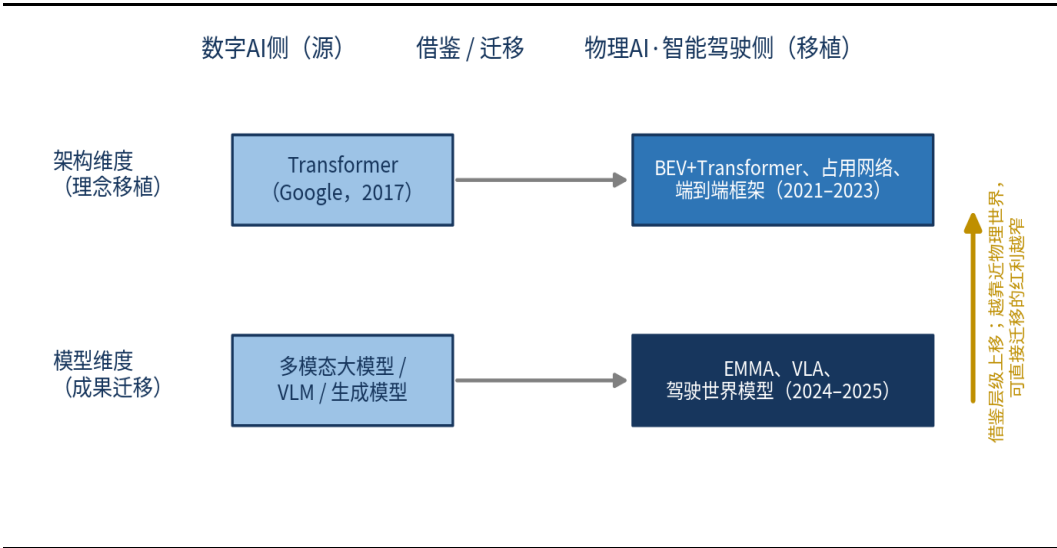
数字 AI 以大语言模型为代表，主要处理文本、代码、图像等数字信号；物理 AI 则需要在真实三维空间中感知、决策和行动。我们以智能驾驶作为切入点，复盘数字 AI 能力如何迁移到物理 AI，以及迁移过程中哪些环节仍需重构。

1.1. 从架构理念到模型层迁移：借鉴层级持续上移

智能驾驶对数字 AI 的借鉴大体经历两个阶段。第一阶段以架构理念移植为主：Transformer 确立的注意力机制与序列建模思想被引入驾驶感知与轨迹预测，催生了矢量化表示 VectorNet，Waymo，CVPR 2020，鸟瞰图 BEV 与 Transformer 结合的统一空间表征，以及以查询解码为核心的端到端一体化框架 UniAD。这一阶段迁移的是建模范式，而非具体模型权重。

第二阶段进一步表现为模型层面的直接迁移。代表性工作是 Waymo 于 2024 年提出的端到端多模态驾驶模型 EMMA，该模型构建在谷歌多模态大模型 Gemini 之上，将原始摄像头输入与文本指令映射到统一语言空间，再输出轨迹、目标检测与道路图等结果。与此同时，源自机器人领域的视觉—语言—动作 VLA 范式被引入车端，例如理想汽车 2025 年发布的 MindVLA，以及小鹏汽车 2025 年提出的约 720 亿参数世界基座模型。

图1：物理 AI 对数字 AI 借鉴层级的跃迁



数据来源：《Attention Is All You Need》，《EMMA: End-to-End Multimodal Model for Autonomous Driving》等其他文献，东吴证券研究所

表1: 数字 AI 关键成果向智能驾驶的迁移路径

数字 AI 侧成果，提出方/年份	智能驾驶侧迁移对应	代表工作，年份，	借鉴层级
Transformer, Google, 2017	统一 BEV 空间表征	特斯拉 BEV 方案 2021、 BEVFormer 2022	架构维度
Transformer 查询解码机制	端到端一体化框架	UniAD CVPR 2023	架构维度
多模态大模型 Gemini, Google, 2023	端到端多模态驾驶模型	Waymo EMMA 2024	模型维度
VLA 范式，源自机器人领域，	车端 VLA 驾驶模型	理想 MindVLA、小鹏世界基 座模型 2025	模型维度
生成式/扩散模型	驾驶世界模型	Wayve GAIA-1 2023、英伟达 Cosmos 2025	模型维度

数据来源：《Attention Is All You Need》，《Planning-oriented Autonomous Driving》等其他文献，东吴证券研究所

1.2. 越接近物理闭环，直接迁移红利越收窄

我们认为，沿着从架构理念到具体模型的借鉴层级上移，可直接复用的成果在形态上越来越完整，但其与物理世界的耦合也越来越紧密。文本、代码等数字信号与驾驶所需的三维时空信号之间存在结构性差异；越接近车辆控制、实时三维感知与安全攸关决策环节，数字 AI 已有成果可直接迁移的比例越低，需要针对物理世界重新设计或重新训练的部分越多。借鉴层级在上移，但越靠近物理闭环，可直接迁移的红利区间可能趋于收窄。

2. 物理 AI 相较数字 AI：表征、数据与学习范式三重差异

在共享底层技术的同时，物理 AI 与数字 AI 在输入表征、数据获取与训练范式三个方面存在系统性差异。

2.1. 表征差异：物理世界表征是必经门槛

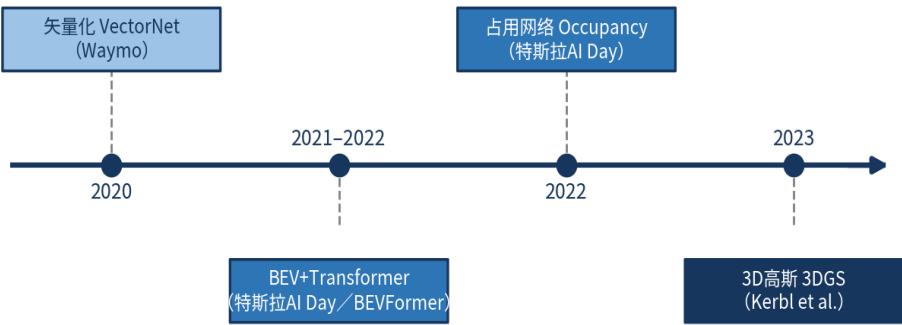
数字 AI 处理的文本天然是离散的、带语义的符号 token，可被模型直接读取，输入端已经具备较强结构化基础。物理世界的原始输入则是连续的光子、点云等信号，并不存在天然、通用、可被模型直接处理的物理世界符号化 tokenization 方案。

因此，智能驾驶技术演进中包含了数字 AI 无需经历的一段历程：将三维物理空间构建为可学习的模型底座。代表性的表征范式包括矢量化表示 VectorNet、BEV 与 Transformer 结合的鸟瞰栅格表征、占用网络 Occupancy Network，以及三维高斯泼溅 3D Gaussian Splatting。

我们认为，数字 AI 面对的是已被人类语言编码过的世界，而物理 AI 需要先把物理世界编码成模型可学习的表征。表征构建能力由此成为物理 AI 区别于数字 AI 的基础

门槛。

图2：物理世界可学习表征演进



数据来源：《VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation》，《3D Gaussian Splatting for Real-Time Radiance Field Rendering》等其他文献，东吴证券研究所

表2：主流物理世界表征范式对比

表征范式	代表工作，提出方/年份	空间组织方式	主要特点
矢量化表示	VectorNet, Waymo, 2020	折线向量集合	紧凑高效，利于行为预测
BEV 栅格	BEVFormer 在 2022 年、特斯拉, 2021	鸟瞰二维栅格	多相机统一，利于规划
占用网络	Occupancy, 特斯拉 AI Day, 2022	三维体素占用	表达任意形状，不依赖类别
三维高斯	3D Gaussian Splatting, Kerbl 等, 2023	可微高斯点云	高保真重建与实时渲染

数据来源：《VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation》，《BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers》等其他文献，东吴证券研究所

2.2. 数据差异：高价值数据稀疏，世界模型承担数据工厂角色

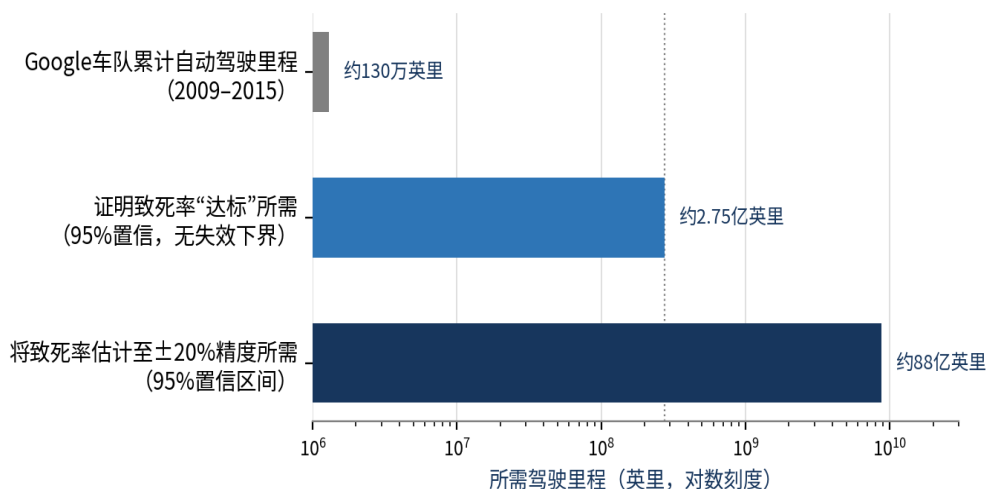
数字 AI 的重要红利在于训练语料已经大量沉淀在互联网中，文本、代码、图像等数据获取成本相对较低，规模化的核心动作是抓取、清洗与压缩已有数据。

物理 AI 最具价值的数据恰恰是长尾、极端与事故场景，而这类数据天然稀疏、采集成本高，且可能涉及安全风险。兰德公司在 2016 年的情景测算说明了路测里程的量级问题：以 2013 年美国人类驾驶的安全水平为基准，每 1 亿英里约 1.09 起致死事故；若要在 95%置信度下证明自动驾驶致死率达标，在无失效情形下约需 2.75 亿英里无失

效里程；若要将致死率估计到 $\pm 20\%$ 的精度，95%置信区间，则约需 88 亿英里。以百辆车规模的车队计算，后者需要数百年。作为对照，谷歌自动驾驶车队在 2009—2015 年累计自动驾驶里程约为 130 万英里。

为缓解高价值数据稀疏问题，业界发展出以世界模型 **World Model** 生成合成数据的路径。代表性项目包括 Wayve 的 GAIA-1 与 GAIA-2，英伟达面向物理 AI 的世界基座模型平台 Cosmos，腾讯混元 3D 世界模型 1.0，以及小鹏云端世界基座模型。

图3：自动驾驶安全性所需测试里程



数据来源：《Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?》，东吴证券研究所

表3：世界模型作为数据工厂的代表项目

项目/模型	机构	时间	定位
GAIA-1	Wayve	2023	面向自动驾驶的生成式世界模型
GAIA-2	Wayve	2025	增强可控性与地域多样性的合成数据生成
Cosmos	英伟达	2025	面向物理 AI 的世界基座模型平台
混元 3D 世界模型 1.0	腾讯	2025	三维世界生成，WAIC 2025 发布，
世界基座模型	小鹏	2025	约 720 亿参数云端驾驶大模型

数据来源：《GAIA-1: A Generative World Model for Autonomous Driving》，《Cosmos World Foundation Model Platform for Physical AI》等其他文献，东吴证券研究所

2.3. 范式差异：强化学习成为物理 AI 关键环节

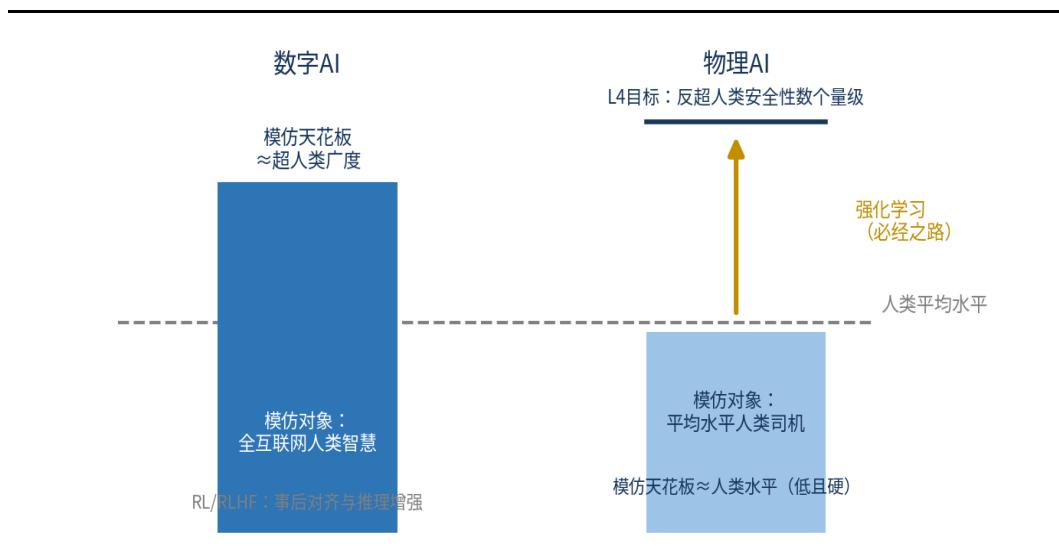
数字 AI 的预训练以下一词元预测为代表，本质上也是一种模仿，但其模仿对象是整个互联网上的人类知识，模仿学习可覆盖的知识边界更宽，因此该路径能够持续释放能力；强化学习与基于人类反馈的强化学习 RL/RLHF 更多承担对齐和推理能力增强角

色。需要说明的是，强化学习在数字 AI 中的作用近年来同样上升：DeepSeek 于 2025 年发布、并于同年 9 月在 Nature 发表的 R1 工作显示，纯强化学习可在不依赖监督微调的情况下激发模型推理能力；OpenAI 的 InstructGPT 亦表明，经 RLHF 训练的较小模型，其输出可被更多评价者偏好于规模更大的基础模型。

物理 AI，以 L4 级自动驾驶为代表，所模仿的对象通常是平均水平的人类司机，而其安全性目标却是显著超越人类。以 Waymo 披露的运营数据为例，其无人驾驶里程已累计超过 1.7 亿英里，与人类基准相比，涉及严重伤害及以上的碰撞减少约 92%、安全气囊弹出事件减少约 83%。

我们认为，当模仿对象为平均水平人类司机、而目标是大幅超越人类安全性时，单纯模仿学习的上限相对更低。在此意义上，强化学习在物理 AI 中更接近突破能力边界的关键路径，而非后置修补。

图4：模仿学习上限的差异：数字 AI 与物理 AI



数据来源：《Training language models to follow instructions with human feedback》，《DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning》，Waymo 官网，东吴证券研究所

3. 结论：数字 AI 提供能力底座，物理 AI 补足物理闭环

物理 AI 正建立在数字 AI 能力外溢的基础之上，且借鉴层级持续上移：从 Transformer 等架构理念的移植，发展到以 EMMA 构建于 Gemini 之上、车端 VLA 等为代表的模型层面直接迁移。我们认为，越靠近物理世界本身、越接近实时三维感知与安全攸关决策环节，数字 AI 已有成果可直接迁移的比例越低，可迁移红利区间可能趋于收窄。

物理 AI 相较数字 AI 存在系统性差异，可分为三方面：在表征上，物理世界缺乏天然、通用的符号化方案，需要先构建可学习的三维表征；在数据上，高价值长尾与极端场景天然稀疏，世界模型被部分机构用作合成数据的数据工厂；在范式上，由于模仿对象为平均水平人类司机而安全目标远超人类，强化学习被多方视为不可或缺的环节。

4. 风险提示

技术迭代不及预期的风险：数字 AI、物理 AI 技术进展低于预期，任务或应用场景拓宽速度低于预期。

大模型厂商 ARR 增速放缓的风险：若客户增长、续费率或客单价不及预期，模型厂商 ARR 增速可能放缓。

云服务商资本开支不及预期的风险：若自由现金流承压，或 AI 算力需求及投资回报低于预期，云服务商可能下调资本开支。

智能驾驶及机器人商业化落地不及预期的风险：若产品可靠性、成本下降或用户需求低于预期，相关产品商业化进程可能放缓。

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

买入：预期未来 6 个月个股涨跌幅相对基准在 15%以上；

增持：预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间；

中性：预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间；

减持：预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持：预期未来 6 个月内，行业指数相对强于基准 5%以上；

中性：预期未来 6 个月内，行业指数相对基准-5%与 5%；

减持：预期未来 6 个月内，行业指数相对弱于基准 5%以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街 5 号

邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>