

汽车行业深度报告

AI 系列深度 24 期：世界模型作为“数据工厂”

增持（维持）

投资要点

- **物理 AI 与数字 AI 的数据差异，在于训练数据是现成可取还是必须主动构造。**数字 AI 依赖文本、代码、图像等已沉淀在互联网中的人类内容，Scaling 核心是抓取、过滤和压缩已有语料；物理 AI 最有价值的长尾、极端和事故场景在现实中天然稀疏、采集昂贵且危险。由此，物理 AI 的数据范式必须从抓取已有转向制造所需。
- **世界模型正在从车端预测组件，演变为云端数据工厂。**特斯拉提出自动驾驶基础世界模型，用神经网络预测并生成未来多摄像头画面；Waymo World Model 直接建模并生成未来世界状态，用于仿真推演与策略评估；NVIDIA Cosmos、Wayve GAIA 等把生成式世界模型定位为合成训练数据来源。其角色已从复现已采集场景，扩展为生成未观测场景。
- **世界模型的价值，集中体现在 Corner Case 制造。**真实道路中的罕见、危险和事故场景难以高频采集，单靠实车路测无法经济地证明安全性；世界模型可以在云端生成新轨迹、新视角、新主体和新交互，为策略训练与验证提供更高密度的边界样本。这使其成为物理 AI 数据供给和安全验证的重要基础设施。
- **但必须正视一项关键张力：用于制造长尾场景的生成式世界模型，其缺陷恰恰容易出现在长尾、分布外和长时域区域。**越稀缺、越偏离训练分布、时间跨度越长，模型越容易出现物理不一致、目标漂移和鬼影等问题。若忽视这一点，世界模型解决数据问题的结论会过于乐观；世界模型既是 Corner Case 生产者，自身也面临 Corner Case。
- **业界缓解张力的做法可分为外生受控生成和内生问题驱动生成。**前者由人工设定扰动参数或生成条件，并以真实轨迹、道路结构和可执行性约束锚定样本；后者由系统自诊断模型表现偏差，再定向生成 hard cases。两类方法的共同方向不是无差别放大生成规模，而是约束生成、提升样本有效性和物理可信度。
- **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。

2026 年 08 月 07 日

证券分析师 黄细里

执业证书：S0600520010001

021-60199793

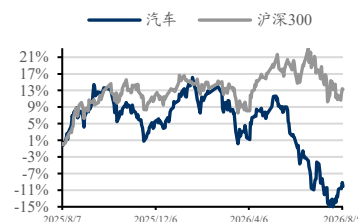
huangxl@dwzq.com.cn

研究助理 童明祺

执业证书：S0600125080005

tongmq@dwzq.com.cn

行业走势



相关研究

《AI 系列深度 17 期：视觉智能为何引入 Transformer? 》

2026-08-06

《AI 系列深度 16 期：语言智能为何从 RNN 转向 Transformer? 》

2026-08-06

内容目录

1. 数据范式转变：从抓取已有到制造所需	4
1.1. 数字 AI：训练语料相对现成，Scaling 核心是抓取与压缩	4
1.2. 物理 AI：最具价值的长尾与事故场景天然稀疏、昂贵且危险	4
1.3. 从抓取已有到制造所需：数据范式转变	5
2. 世界模型作为数据工厂：从复现场景到生成世界	6
2.1. 概念脉络：从强化学习的内部模型到生成式世界模拟器	6
2.2. 判别式到生成式：特斯拉提出世界模型作为数据生产工具	6
2.3. Waymo：从复现场景到生成世界	7
2.4. 国际前沿机构与主机厂的世界模型布局	7
2.5. 生成式视频模型作为世界模型的技术底座：中国大厂与头部创业公司	9
3. Corner Case 制造：世界模型既解决长尾，也受长尾约束	11
3.1. 可控场景生成与罕见场景的主动创造	11
3.2. 生成式世界模型在长尾场景的可靠性同样面临下滑问题	11
3.3. 思路一：外生受控生成，以真实结构锚定	11
3.4. 思路二：内生问题发现，由系统自诊断驱动定向生成	12
4. 结论：世界模型是物理 AI 数据范式转变的核心基础设施	13
5. 风险提示：	14

图表目录

图 1: 证明自动驾驶安全性所需驾驶里程的数量级5

图 2: 生成式世界模型发布时间轴8

表 1: 数字 AI 与物理 AI 在数据获取性上的对比6

表 2: 主要数据工厂型世界模型梳理9

表 3: 中国主要视频生成与世界模型基座梳理10

表 4: L2 与 L4 训练范式对比13

表 5: Corner Case 生成的两类思路对比13

1. 数据范式转变：从抓取已有到制造所需

物理AI与数字AI训练处于两种不同的数据供给环境。数字AI面对的是已经存在、可被抓取的海量语料，物理AI面对的是需要主动构造的稀缺场景。理解两者的数据获取性差异，是判断世界模型为何成为数据工厂的前提。

1.1. 数字AI：训练语料相对现成，Scaling核心是抓取与压缩

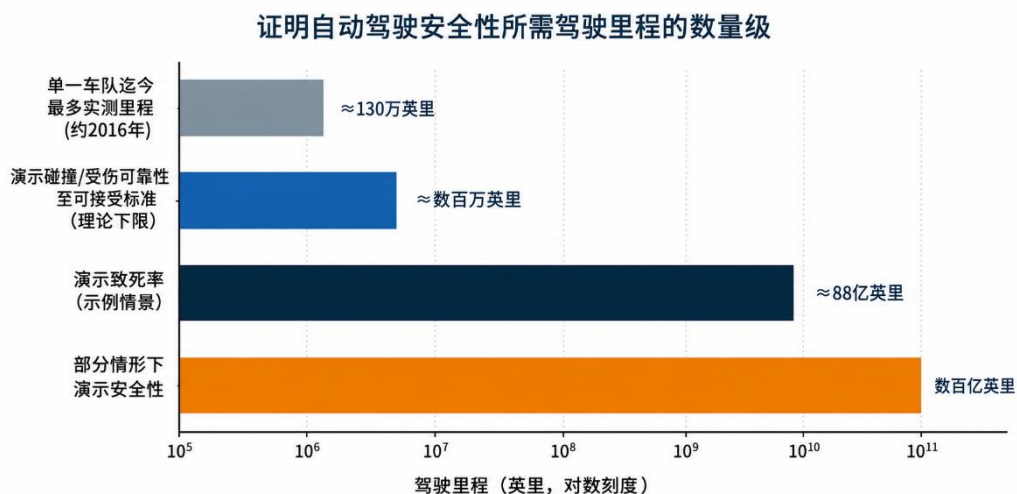
数字AI的训练红利，建立在文本、代码与图像等语料早已存在于互联网这一前提上。大语言模型与图像模型所需训练数据，本质上是人类已经生产并公开发布的内容，获取方式是对公开网页等来源进行抓取、过滤与压缩，而非从零创造。Epoch AI关于数据规模的研究估计，公开人类生成文本的有效存量约为300万亿token，并给出100万亿至1000万亿token的较宽区间。在该数量级条件下，提升既有数据利用效率是数字AI扩张的重要路径。

公开高质量文本数据存在可见存量上限，多项研究将其耗尽时点在5年以内。Epoch AI在2024年的研究预计，可公开获取的高质量文本数据可能在2026至2032年间基本用尽；其早期工作曾给出更靠前的时点，后因更有效的数据过滤与多轮训练等方法，将窗口前端从2026年推后至2028年附近。

1.2. 物理AI：最具价值的长尾与事故场景天然稀疏、昂贵且危险

物理AI最具训练价值的恰是长尾、极端与事故场景，而它们在现实中天然稀疏。与文本不同，自动驾驶等物理AI系统真正需要学习的，是突发避让、恶劣天气、临界碰撞等低频但关键的边界场景，这些场景在日常行驶记录中出现的频率极低。RAND以美国道路安全统计为基准测算，2013年人类驾驶死亡率约为每亿英里1.09人。而由于致死与受伤是相对于行驶里程的稀有事件，要仅凭实车路测在统计上清晰地证明自动驾驶的安全性，需要数亿乃至数千亿英里，其给出的一个示例情形即需约88亿英里；该研究同时提到，当时单一车队所积累的最多实测里程约为130万英里，与所需里程相差若干个数量级。

图1：证明自动驾驶安全性所需驾驶里程的数量级



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

真实采集这类场景还面临高成本与安全风险，难以单纯依靠扩大路测规模解决。RAND 据此提出，开发者无法仅依靠增加道路里程来证明安全，必须发展加速测试、虚拟测试与仿真、场景测试等替代方法。这一困境在产业界同样被反复提及，NVIDIA 在其面向驾驶的合成数据研究中指出，制作大规模、多视角、时序一致并带细粒度标注的驾驶数据集成本高昂，且长尾与安全攸关场景稀缺进一步加大了难度。

由此，物理 AI 数据问题的性质与数字 AI 相反，不是存量管理，而是增量制造，即如何安全、低成本地获得现实中难以遇到的关键场景。

1.3. 从抓取已有到制造所需：数据范式转变

数字 AI 与物理 AI 在数据来源、成本、分布与可控性上存在系统性差异。

表1: 数字 AI 与物理 AI 在数据获取性上的对比

维度	数字 AI	物理 AI
数据是否现成	文本、代码、图像已存在于互联网	长尾、极端、事故场景需主动制造
主要获取方式	抓取、清洗、压缩已有数据	实车采集，叠加仿真与生成
边际成本与风险	边际成本低，安全风险相对较低	采集成本高，部分场景具危险性
价值数据的分布	高质量语料相对可得	价值高度集中于稀疏长尾
存量约束	公开高质量文本预计 2026—2032 年趋于用尽	真实事故场景天然稀疏，路测难以穷尽
Scaling 核心动作	抓取与压缩存量	制造与覆盖未观测状态

数据来源:《Will we run out of data? Limits of LLM scaling based on human-generated data》,《Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?》等其他文献, 东吴证券研究所

2. 世界模型作为数据工厂：从复现场景到生成世界

面对增量制造需求，世界模型被引入为生成训练数据的工具。重点为预测状态与动作后果、支撑仿真训练和合成数据生成中的作用。

2.1. 概念脉络：从强化学习的内部模型到生成式世界模拟器

世界模型概念源自强化学习，指智能体对环境状态及其演化的内部预测模型。Ha 与 Schmidhuber 在 2018 年提出的代表性工作，用变分自编码器压缩高维感知、用循环神经网络建模时间动态，使智能体可在学习到的模型内部生成可能状态并训练策略。此后这一思路在 Dreamer 等系统中得到延续，世界模型的内涵也随大规模生成式建模发展而扩展，从辅助智能体的预测器，逐步演变为可生成完整交互世界的系统。

世界模型存在理解世界与预测未来两条路径，分别以 LeCun 的 JEPA 与生成式视频模型为代表。LeCun 在 2022 年的立场论文中主张世界模型应在抽象表征空间而非像素空间预测，由此发展出 JEPA、V-JEPA 以及面向机器人规划的 V-JEPA2 等非生成式路径；与之并行的生成式路径以 OpenAI 将视频生成模型视为世界模拟器、以及 DeepMind 的 Genie 等交互式生成环境为代表。但是现阶段无论哪条路径，所达到的都是对部分物理规律的近似建模，距离能够完整理解物理世界的通用世界模型仍有明显差距。

2.2. 判别式到生成式：特斯拉提出世界模型作为数据生产工具

特斯拉在 CVPR2023 提出构建自动驾驶领域的基础世界模型，标志着系统重心从判别式感知向生成式预测转移。当感知能力随 Occupancy Networks 趋于完善后，系统瓶颈转移到预测与规划；现实交通是高度不确定与动态交互的复杂系统，传统基于规则的搜索与硬编码交通规则难以应对。

特斯拉的神经网络可根据过去数秒视频，预测并生成未来全部 8 个摄像头画面；生成画面不仅在视觉上连贯，还体现出对车辆行驶轨迹、光影变化与遮挡关系等物理要素的建模能力，从而减少手工物理规则依赖。

2.3. Waymo：从复现场景到生成世界

Waymo 云端仿真经历了从真实数据回放到生成式世界模型的演进。1) 2016 至 2020 年其云端以 Carcraft 为核心，仿真体系以真实数据回放为主，主要服务模块化系统验证；2) 2020 年以来引入基于神经网络的多主体行为建模，仿真由规则驱动转向学习驱动；3) 2024 年前后随 SimAgents 等生成模型的应用，能力由轨迹预测扩展至场景生成；4) 同年 Waymo Foundation Model 统一建模场景状态与交互关系；5) 至 2026 年 2 月，Waymo 正式发布 Waymo World Model，直接建模并生成未来世界状态，用于仿真推演与策略评估。

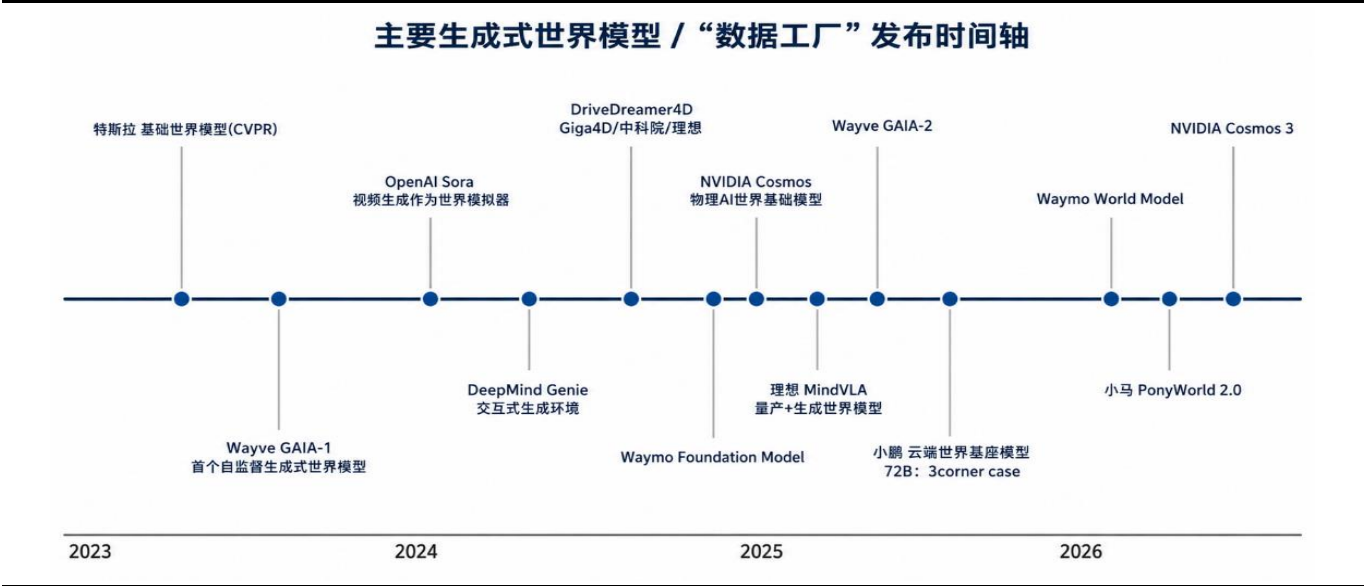
Waymo World Model 把自动驾驶问题由轨迹预测重构为对交通系统状态演化的整体建模。该模型通过统一框架刻画场景上下文与多主体交互，建模对象由单体轨迹预测转向多主体联合分布，体现为对环境状态转移过程的刻画。其主要价值在于缓解长尾场景覆盖与泛化能力约束，通过生成新的场景分布并支持策略推演，把仿真由复现样本拓展为覆盖潜在状态空间，使系统能力不再受限于既有数据覆盖，而更多依赖模型对未观测场景的外推。这一定位与特斯拉一脉相承，即把世界模型从车端组件提升为云端数据与训练基础设施。

2.4. 国际前沿机构与主机厂的世界模型布局

以 NVIDIA 与 Wayve 为代表的国际机构，已把生成式世界模型明确定位为合成训练数据来源。NVIDIA 于 2025 年 1 月发布 Cosmos 世界基础模型平台，用于生成具照片级真实感、符合物理约束的合成数据，以训练与评估物理 AI。

早期采用方包括 1X、Agility、Figure AI、Foretellix、Uber、Waabi 与小鹏等，Wayve 亦在评估用 Cosmos 检索边界与 Corner Case 场景。Wayve 方面，GAIA-1 被描述为首个面向自动驾驶的生成式世界模型，规模约 90 亿参数；2025 年 3 月发布的 GAIA-2 采用潜在扩散框架，支持多摄像头、时空一致的可控生成，覆盖英国、美国与德国等地，并可按需求生成临界碰撞、紧急制动乃至分布外危险场景用于训练与验证。

图2：生成式世界模型发布时间轴



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

表2：主要数据工厂型世界模型梳理

机构/项目	发布时点	定位	技术与路径要点	主要用途
特斯拉基础世界模型	CVPR2023	自驾基础世界模型	预测生成未来 8 摄像头画面，体现物理建模能力	预测与规划，物理状态建模
Wayve GAIA-1	2023	首个自驾生成式世界模型	约 90 亿参数生成式世界模型	合成数据与仿真
Wayve GAIA-2	2025-03	可控多视角生成式世界模型	潜在扩散，多摄像头，覆盖英美德，可生成分布外危险场景	合成数据与验证
NVIDIA Cosmos	2025-01 起	物理 AI 世界基础模型平台	生成照片级、物理一致合成数据，Cosmos3 于 2026 年发布	训练与评估数据生成
Waymo World Model	2026-02	生成未来世界状态	多主体联合分布，环境状态演化建模	仿真推演与策略评估
X-World	学术	可控自车多相机世界模型	扩散与潜空间，视图-时间自注意力，I2V/V2V/C2V	闭环仿真、在线 RL、数据生成
DrivePTS	CVPR2026	可控场景生成与 Corner Case 构造	渐进式学习加互信息约束，VLM 分层描述，频率引导结构损失	生成罕见场景
DriveDreamer4D	2024	世界模型先验驱动 4D 重建	真实轨迹受控扰动加可执行性筛选，联合优化 4DGS	跨轨迹训练样本生成
理想 MindVLA	2025-03	重建加生成统一世界模型	3D 高斯波减约提速 7 倍，仿真长尾支撑	闭环强化学习训练
小鹏云端世界基座模型	2025 CVPR2025	主机厂云端世界基座模型	72B 基座，强化学习补长尾，面向 Corner Case 生成	合成数据与闭环训练
小马智行 PonyWorld1.0	2020 起	L4 强化学习训练底座	真实数据校准，Virtual Driver rollout 闭环	训练、评估与迭代
小马智行 PonyWorld2.0	2026-04	自诊断训练闭环	Intention Layer 自诊断，问题驱动生成 hard cases	问题驱动迭代

数据来源：《GAIA-2: A Controllable Multi-View Generative World Model for Autonomous Driving》，《DriveDreamer4D: World Models Are Effective Data Machines for 4D Driving Scene Representation》等其他文献，特斯拉、英伟达等公司公开资料，东吴证券研究所

2.5. 生成式视频模型作为世界模型的技术底座：中国大厂与头部创业公司

生成式世界模型在技术上以视频生成模型为底座，中国大厂与头部创业公司在这一

基座层快速追赶。

1) 阿里通义万相 Wan 于 2025 年 2 月开源，并在 2025 年初的 VBench 评测中居于前列；2) 字节跳动布局 Seedance、PixelDance 与即梦等；3) 腾讯混元 HunyuanVideo 于 2024 年底开源，约 130 亿参数，当时为规模较大的开源视频模型；4) 智谱 AI 的 CogVideoX 与清影已开源；5) MiniMax 的海螺 Hailuo 在 2025 年 10 月推出 2.3 版本；6) 生数科技的 Vidu 与快手的可灵 Kling 亦在评测中表现靠前。

但上述模型多为通用视频生成模型，视频生成能力不能直接等同于可被安全闭环依赖的数据工厂，仍需动作条件化、长时域一致性与安全评估闭环验证，但它们构成了生成式世界模型的共同技术底座。

大模型公司更多以多模态基座与具身智能底座的角色介入物理 AI，自动驾驶专用世界模型则主要由智驾厂商推进。1) DeepSeek 与月之暗面以大语言模型为主，DeepSeek 被用作部分具身机器人厂商的推理底座，智谱 AI 与阿里 Qwen 等则作为多模态与具身研究的开源基座；2) 百度在自动驾驶上的主线是 Apollo 自动驾驶大模型 ADFM 与萝卜快跑运营，并于 2024 年 12 月发布 Apollo 开放平台 10.0。

我们认为，中国大厂在通用生成与基座层投入较大，面向自动驾驶的专用世界模型主要由小马智行、理想与小鹏等厂商承担，两者在产业链上呈分工关系。

表3：中国主要视频生成与世界模型基座梳理

机构	代表模型/产品	类型与状态	与物理 AI/世界模型的关系
阿里巴巴 通义	通义万相 Wan	视频生成，部分开源	通用视频生成基座，2025 年初 VBench 居前列
字节跳动	Seedance PixelDance 即梦	视频生成	通用视频生成基座
腾讯	混元 HunyuanVideo	视频生成，2024 年底开源，约 130 亿参数	通用视频生成基座
智谱 AI	CogVideoX 清影	视频生成，开源	通用视频生成基座，兼具多模态底座
MiniMax	海螺 Hailuo	视频生成	物理世界理解能力增强
生数科技	Vidu	视频生成	通用视频生成
快手	可灵 Kling	视频生成	通用视频生成
DeepSeek	DeepSeek	以大语言模型为主	多作为多模态与推理底座介入具身，非自驾专用世界模型
月之暗面	Kimi 等		
百度	Apollo ADFM、萝卜快跑	自动驾驶大模型与运营	自驾主线以 ADFM 与运营为主

数据来源：《CogVideoX: Text-to-Video Diffusion Models with An Expert Transformer》，阿里巴巴、腾讯等公司公开资料，东吴证券研究所整理

3. Corner Case 制造：世界模型既解决长尾，也受长尾约束

世界模型被寄望于规模化生成 Corner Case，但其自身在长尾与长时域处的可靠性同样构成约束。行业的核心任务由此集中于两点：提升罕见场景的可控生成能力，并以约束、筛选与反馈机制控制长尾生成误差。

3.1. 可控场景生成与罕见场景的主动创造

学术界已实现在给定动作条件下、符合物理约束的可控多视角场景生成。该模型可做到自车动作可控、动态与静态元素可控、长时程生成、跨相机一致与全局外观编辑，并已投入闭环仿真测试、在线强化学习与数据生成等研发环节。其技术要点包括以 VAE 把视频压成潜空间后用扩散模型去噪以节省算力，以视图-时间自注意力维持跨视图与时间的一致性，以解耦交叉注意力分通道注入自车动作、动态物体、静态路元素与相机参数等条件，并支持图生视频、视频生视频与条件生视频三种模式；其两阶段训练先把大型预训练视频生成模型改造为可控多摄像头世界模型，再转换为流式自回归模拟器。

DrivePTS 进一步面向 Corner Case 生成，将常规场景扩展为罕见场景。（1）针对几何条件耦合导致的可控性下降，模型采用渐进式学习与互信息约束：先生成道路结构，再生成交通参与者，并降低地图与车辆特征的相似度，强化不同条件的独立约束。（2）针对粗粒度文本难以有效约束跨视角语义，模型利用视觉语言模型生成多视角、分维度的层级描述，提供更细粒度的文本条件。（3）针对前景细节缺失，模型以傅里叶变换与高通滤波提取高频信息，并在前景区域引入频率引导结构损失，强化边缘与轮廓还原。三类机制分别提升几何可控性、语义约束力和前景结构保真度，为罕见场景生成提供支撑。

3.2. 生成式世界模型在长尾场景的可靠性同样面临下滑问题

用于制造长尾场景的生成式世界模型，其生成可靠性恰在长尾与长时域处最易退化。

长时程一致性是生成式世界模型反复强调的难点，Wayve 在 GAIA-2 描述中专门强调要维持长时域时空一致性，X-World 亦专门设计视图-时间自注意力与流式自回归机制来维持长时间序列连贯。

在分布外方向，DriveDreamer4D 提出的背景正是现有重建方法在跨原始采集轨迹渲染时容易出现目标错位、车道模糊与鬼影，说明越偏离已观测分布，生成质量越难保证，意味着越稀缺、越偏离训练分布、时间序列越长的场景，生成越不可靠，而这些区域恰是最需要 Corner Case 的区域。

因此，世界模型既是 Corner Case 生产者，本身也面临 Corner Case。

3.3. 思路一：外生受控生成，以真实结构锚定

第一类做法是外生受控生成，生成内容由外部预设扰动参数或生成条件指定，并以

真实轨迹与可重建结构为锚、以可执行性约束筛选；前述 X-World 与 DrivePTS 等可控生成方法亦属这一类。DriveDreamer4D 由 GigaAI、中国科学院自动化所与理想汽车等联合于 2024 年提出，其核心是把世界模型先验从数据侧引入 4D 重建：第一，将原始自车轨迹转换至首帧自车坐标系，通过横向偏移与纵向速度扰动构造变道、加减速等新轨迹，并依据可行驶区域与最小碰撞距离筛除不可执行路径。第二，将 3D 车辆框、车道地图与相机位姿同步投影至新轨迹视角，形成逐帧几何条件，约束前景目标、道路拓扑与相机运动的一致性。第三，条件视频扩散模型依据上述几何条件生成新视角序列，将世界模型转化为新轨迹训练样本生成器。第四，模型构造时间对齐的真实与生成视频对，联合优化 4DGS，并通过感知正则抑制鬼影、飞点与目标漂移。该路径以真实结构约束生成过程，降低分布外场景的退化风险。

理想的重建加生成统一世界模型体现同一取向，即以重建锚定真实、以生成补全未见视角。其基于自研的重建加生成云端统一世界模型，融合重建模型的三维场景还原与生成模型的新视角补全及未见视角预测，构建接近真实世界的仿真环境，并以 3D 高斯泼溅把场景重建提速约 7 倍，在仿真环境中模拟施工区与紧急避让等长尾场景以支撑大规模闭环强化学习。我们归纳认为，DriveDreamer4D 与理想 MindVLA 同属以真实数据为基底、以生成补全的路线，其共同点是让生成建立在可重建、可校验的真实结构之上。

小鹏云端世界基座模型属于外生受控生成路线，重点面向自动驾驶 Corner Case 与高价值训练数据生产。小鹏已启动 720 亿参数世界基座模型研发，并以强化学习、预训练、后训练与模型蒸馏构建云到端的模型迭代流程。其核心作用在于对撞车、事故等稀缺极端场景形成可规模化补全，并通过车端运营样本持续回流优化数据与模型。

3.4. 思路二：内生问题发现，由系统自诊断驱动定向生成

第二类做法是内生问题驱动生成，即把 Corner Case 发现与触发权交给系统自身，由模型对自身表现的诊断驱动定向生成，而非由外部无差别放大。PonyWorld 自 2020 年起建设，其 1.0 版本定位为贯穿云端训练与车端部署的强化学习训练底座，用真实车队数据校准世界模型，让 Virtual Driver 在虚拟环境中完成 rollout、奖励评估与策略更新，训练结果部署到车端后再通过真实运营样本回流进入下一轮迭代。小马在 2020 年前后从模仿人类驾驶的 Learning by Watching 转向在世界模型中反复训练的 Learning by Practicing，并把世界模型称为训练车端模型的工厂，其精度决定 Virtual Driver 训练上限。

表4: L2 与 L4 训练范式对比

维度	L2 辅助驾驶	L4 无人驾驶
责任边界	人类驾驶员兜底	系统独立承担驾驶责任
训练范式	Learning by Watching	Learning by Practicing
数据价值	规模驱动体验提升	高价值样本驱动能力提升
技术核心	模仿学习与端到端体验	世界模型加强学习闭环

数据来源：小马智行 CTO 楼天城公开访谈，东吴证券研究所整理

PonyWorld2.0 以 Intention Layer 做自诊断，把模型表现偏差拆解为可训练问题并据此生成定向采集与 hard cases。该版本于 2026 年 4 月正式发布，在 1.0 强化学习训练体系上引入可诊断中间表征与训练调度机制，把模型迭代从数据回流驱动推进至问题诊断驱动。（1）其意图层在输出驾驶动作之外显式表征决策意图，用于对齐意图、动作与结果的一致性；（2）归因层进一步拆分车端策略误差与世界模型误差，把表现偏差转化为可操作的训练方向；（3）数据层据诊断结果生成定向采集需求，把某类路口、混行交通或低置信度决策等工况转化为真实道路采集任务；（4）训练层则把重点从增加仿真里程转向提高样本有效性，将算力集中到 hard cases。

我们认为，这一做法的实质是把 Corner Case 发现权从工程师经验转移到系统内生诊断，从而让生成与采集更有针对性。

表5: Corner Case 生成的两类思路对比

思路	代表	Corner Case 的发现/触发来源	典型做法	主要约束与筛选
外生：受控扰动或生成	DriveDreamer4D、理想 MindVLA、小鹏云端世界基座	由人工预设或外部条件指定，包括扰动参数、场景类型、通用生成目标	真实轨迹横向偏移与纵向变速等受控扰动，重建加生成补全，按条件生成场景	可行驶区域、最小碰撞距离、结构条件投影、真实锚定
内生：问题驱动地发现	小马 PonyWorld2.0	由系统对自身表现的诊断指定	生成定向采集与 hard cases	意图-动作-结果一致性、车端策略误差与世界模型误差归因

数据来源：《DriveDreamer4D: World Models Are Effective Data Machines for 4D Driving Scene Representation》，理想、小马智行等公司公开资料，东吴证券研究所整理

4. 结论：世界模型是物理 AI 数据范式转变的核心基础设施

我们认为，物理 AI 的数据范式正从抓取已有转向制造所需，世界模型是这一转变的核心基础设施。数字 AI 的数据问题是对有限存量的高效利用，物理 AI 的数据问题则是对稀缺、危险、未观测场景的安全构造，二者性质不同。世界模型从特斯拉提出的车端基础模型，到 Waymo 直接生成未来世界状态、NVIDIA 与 Wayve 把生成定位为合成

数据来源，其角色已由复现已采集场景扩展为生成未观测场景，承担起数据供给的新基础设施职能。

世界模型作为数据工厂的价值与其自身可靠性边界并存，二者需同时评估。世界模型最被需要的能力，是在长尾、分布外与长时域处生成高价值场景，而这恰是其生成可靠性最薄弱的区域。这一张力意味着，不能因为世界模型能生成 Corner Case 就默认数据问题已被解决，世界模型自身也面临 Corner Case，生成数据的物理一致性与分布外可靠性需要被持续检验。

从可观测做法看，业界共同取向是约束生成与提高样本有效性，而非无差别扩大生成规模。外生受控生成以真实结构锚定、降低分布外退化，内生问题驱动生成把 Corner Case 发现权从工程师经验转移到系统自诊断。两类做法路径不同，但都指向同一方向，即让生成更可控、更有针对性、更可校验，这与单纯追求生成规模的思路存在区别。

5. 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>