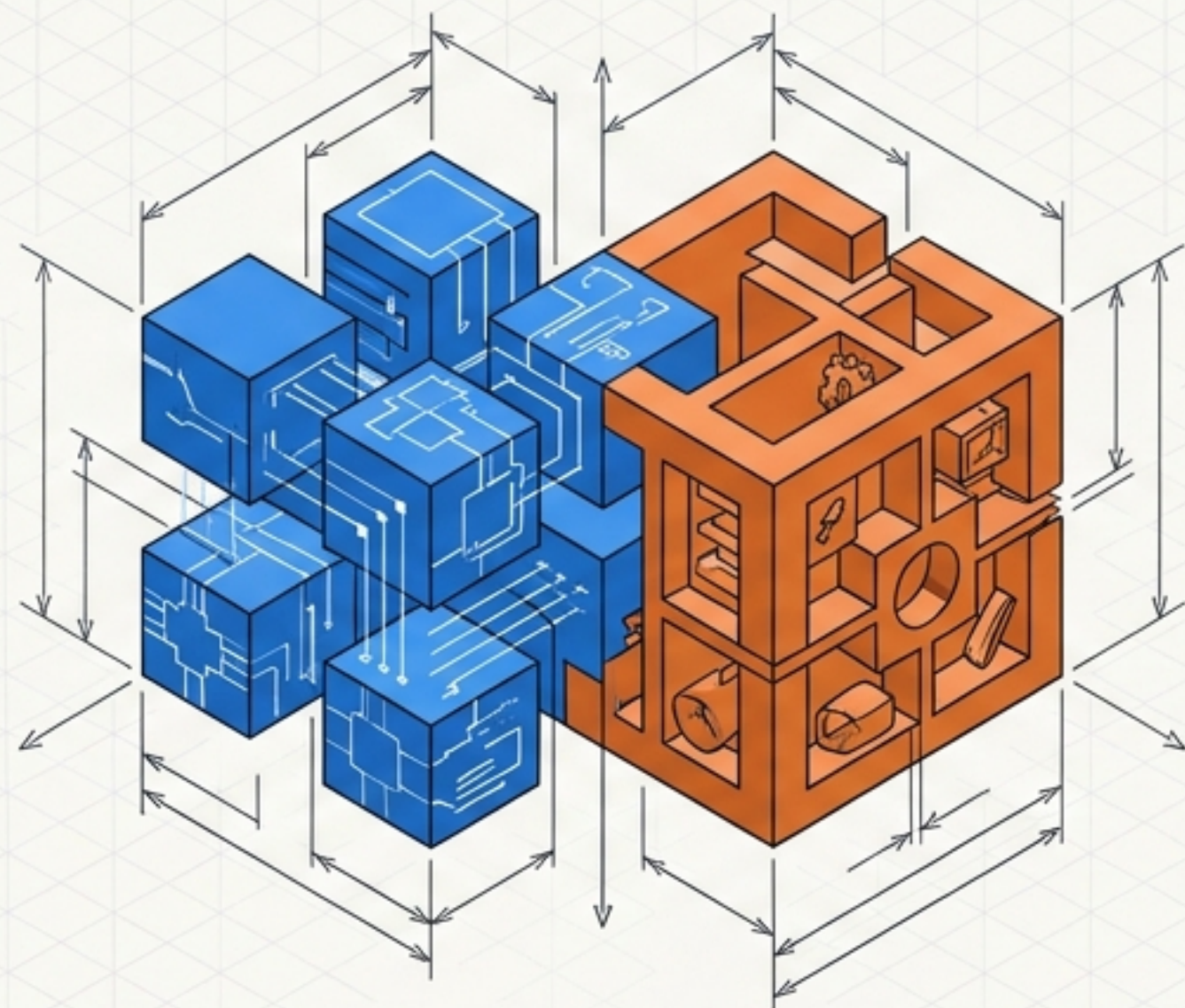




具身智能与物理AI：从VLA到World Models的范式演进

挑战、融合趋势与实践建议

Michael HUO

北美高校联盟 AI 创业俱乐部主席 / 唯爱AI基金会主席

研究方向：物理AI、具身智能与AI智能体系统

2026年7月18日

核心战略目标



绘制物理AI挑战雷达

理解当前领域面临的主要技术与物理瓶颈。



解构架构融合趋势

分析VLA与World Models的互补关系及未来架构。



展望2026-2030路线图

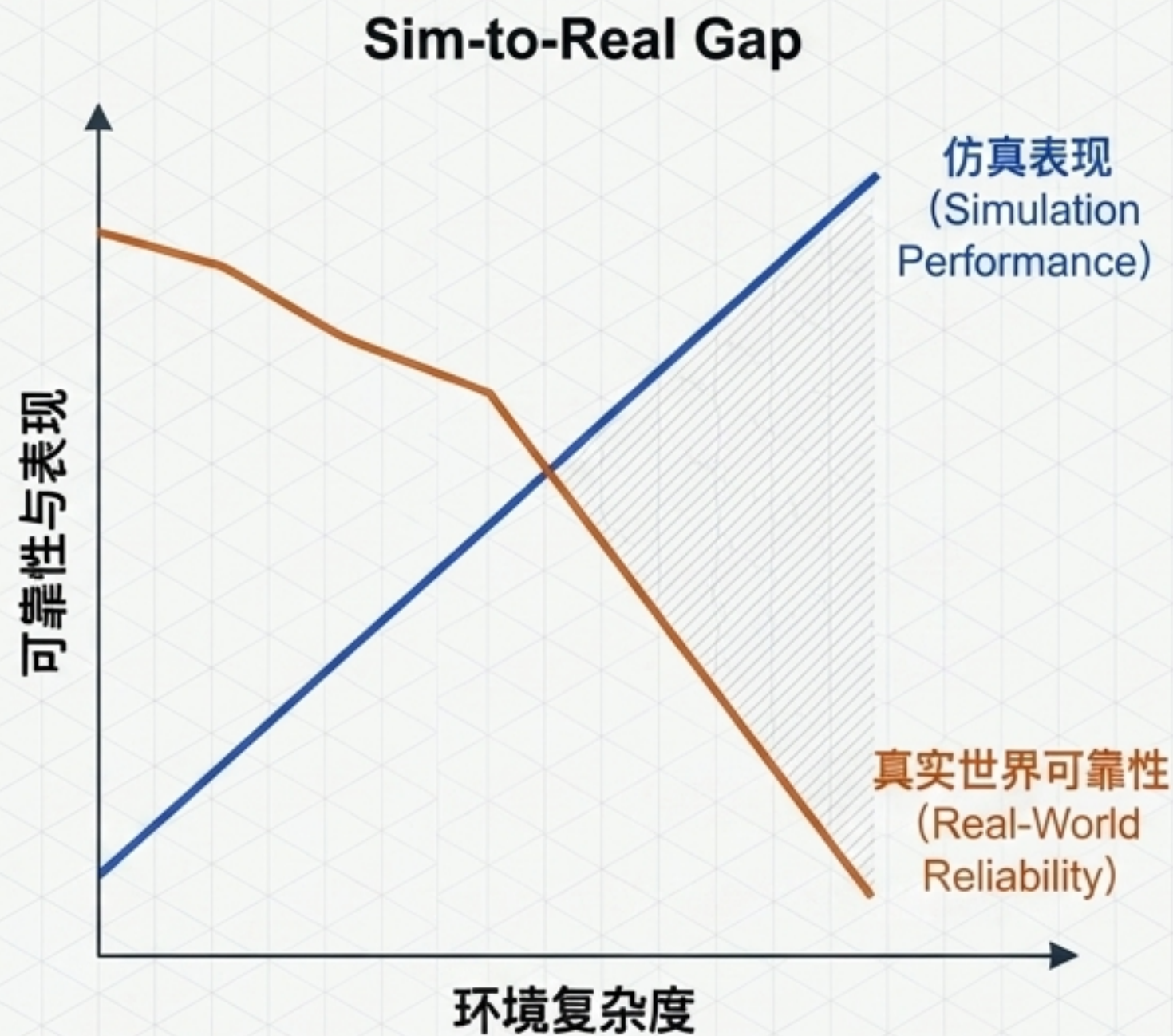
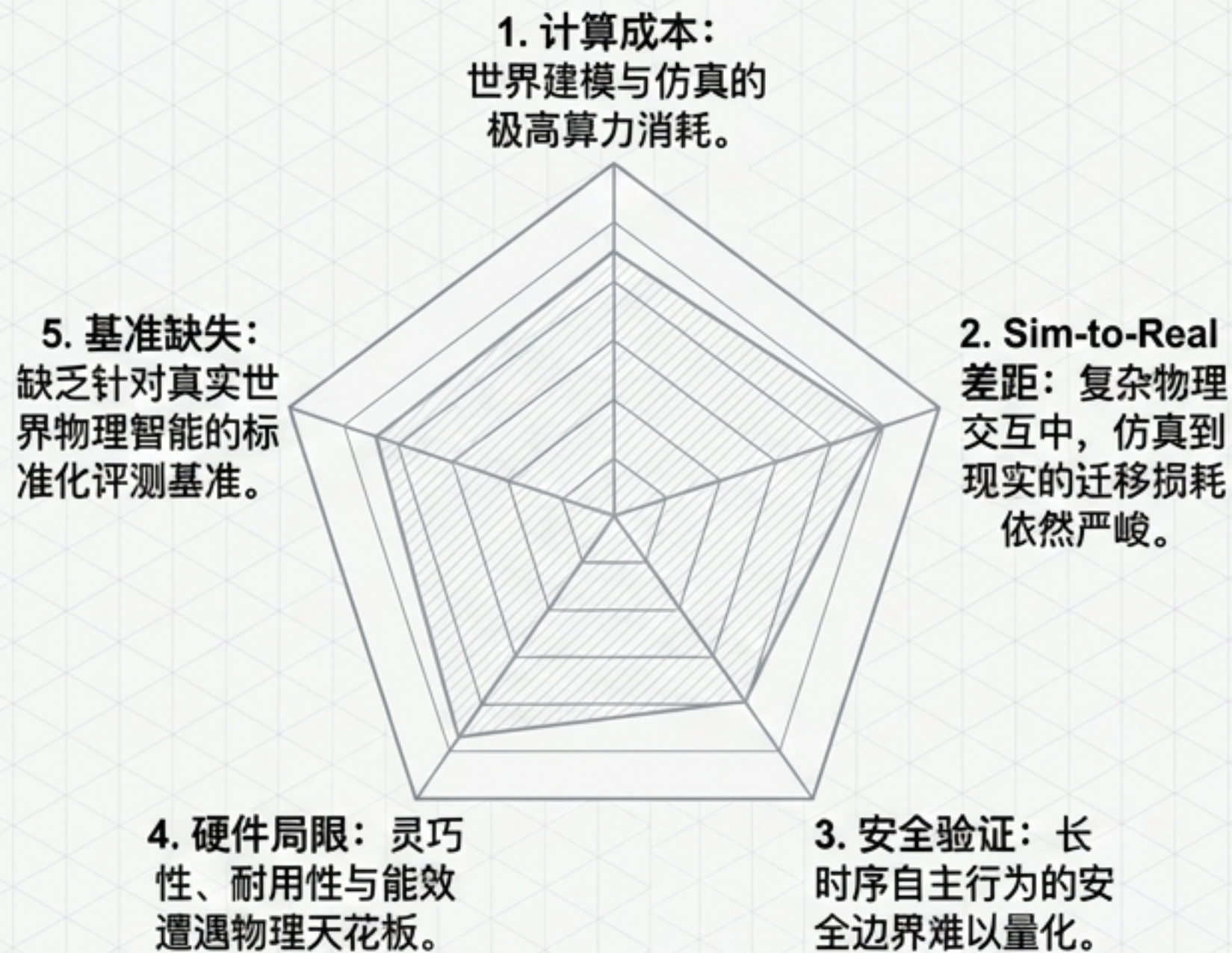
把握从受控环境到非结构化通用场景的演进时间表。



制定执行级实践方案

为个人与组织提供即插即用的实操建议。

2026年的关键摩擦点



行业共识：泛化能力与非结构化安全的博弈

在 WAIC 2026 大会上，业界专家达成一项核心共识：
虽然 World Models 显著提升了系统的泛化能力，
但整个领域在非结构化环境中，仍极度缺乏成熟的安全保障与长期可靠性解决方案。

泛化是基础，但在开放世界中，缺乏安全验证的智能无法转化为真正的生产力。

范式诊断：竞争还是互补？

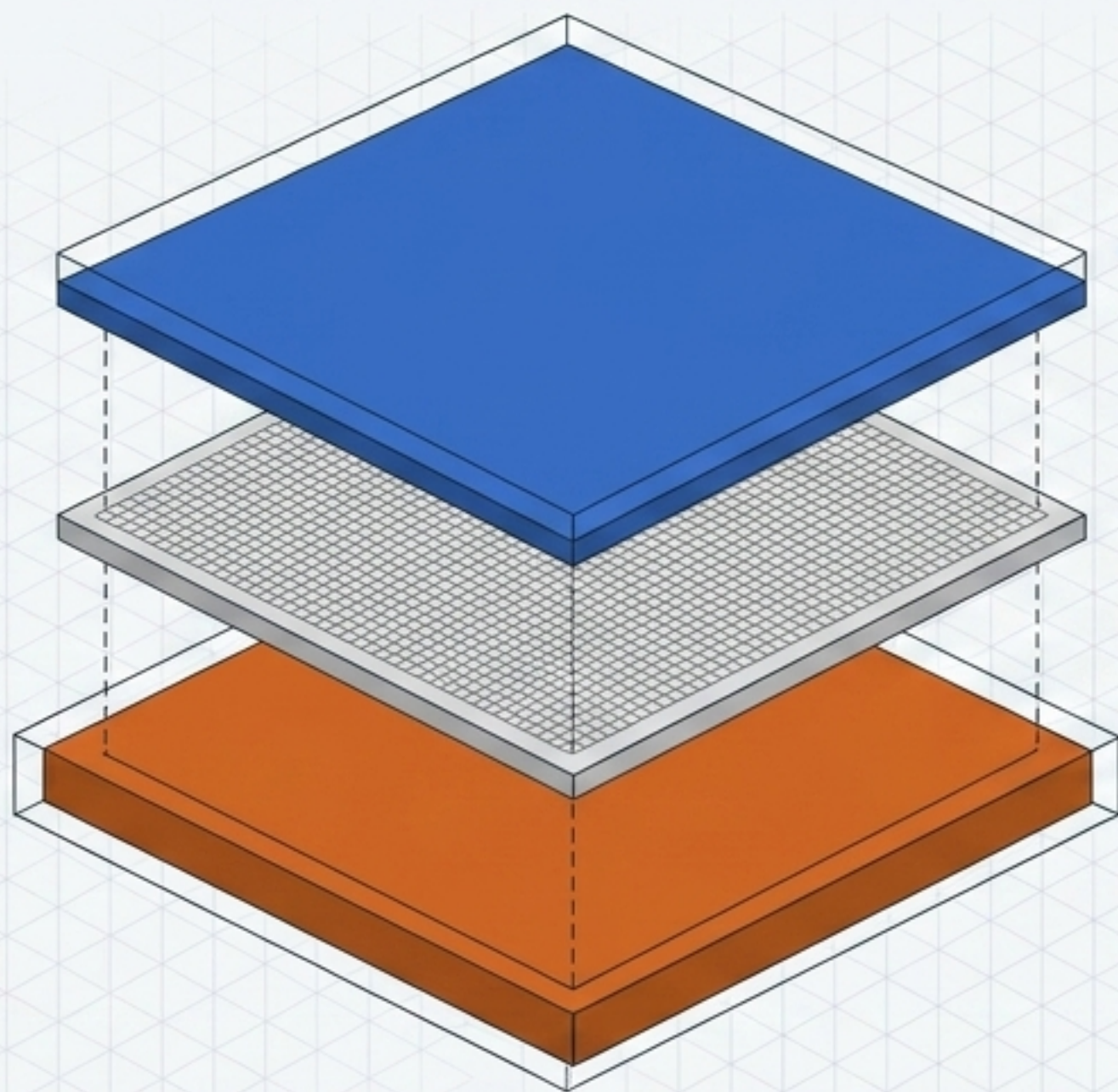
来源于 AGI-2026 与 WAIC 2026 的技术研判结论。

	[VLA]	[World Models]
核心优势	高层推理与语言理解	物理预测与因果关系建模
推理层级	语义级逻辑 (Semantic Logic)	时空级逻辑 (Spatial-Temporal Logic)
作用域	任务意图解析与用户交互	底层物理交互与环境规划

结论：二者非零和博弈，而是构建下一代物理AI的完美拼图。

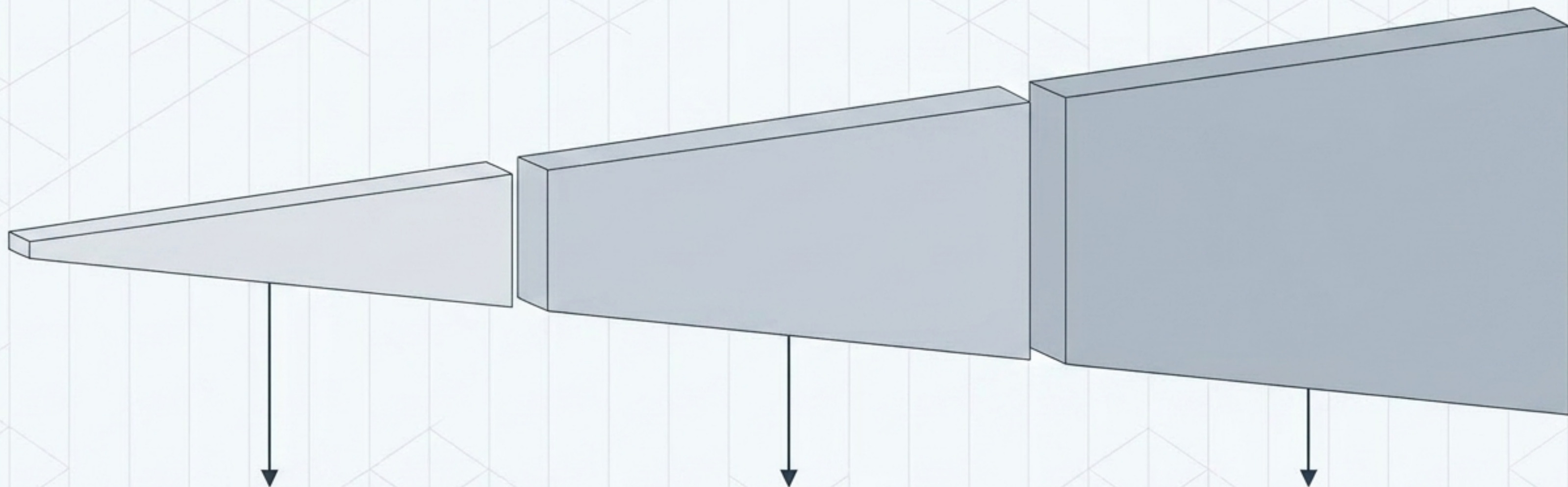
终极形态：混合系统架构

未来3-5年主导物理AI的标准范式设计。



- 顶层：任务理解层 (VLA / LLM)**
负责高层任务理解、多模态输入与自然语言交互。
- 中层：桥接协调层 (Coordination Layer)**
负责解析高层意图，将其转化为可执行的物理参数边界。
- 底层：物理执行层 (World Models)**
负责底层物理预测、因果推理、微秒级运动规划与动作输出。

2026-2030 演进路线图



2026-2027: 落地启航

场景：受控工业环境

特征：混合系统架构开始广泛部署，解决特定场景下的高价值痛点。

2028-2029: 跨越深水区

特征：安全验证体系大幅完善，边缘计算能力提升使得系统计算需求显著降低。

2030: 走向通用

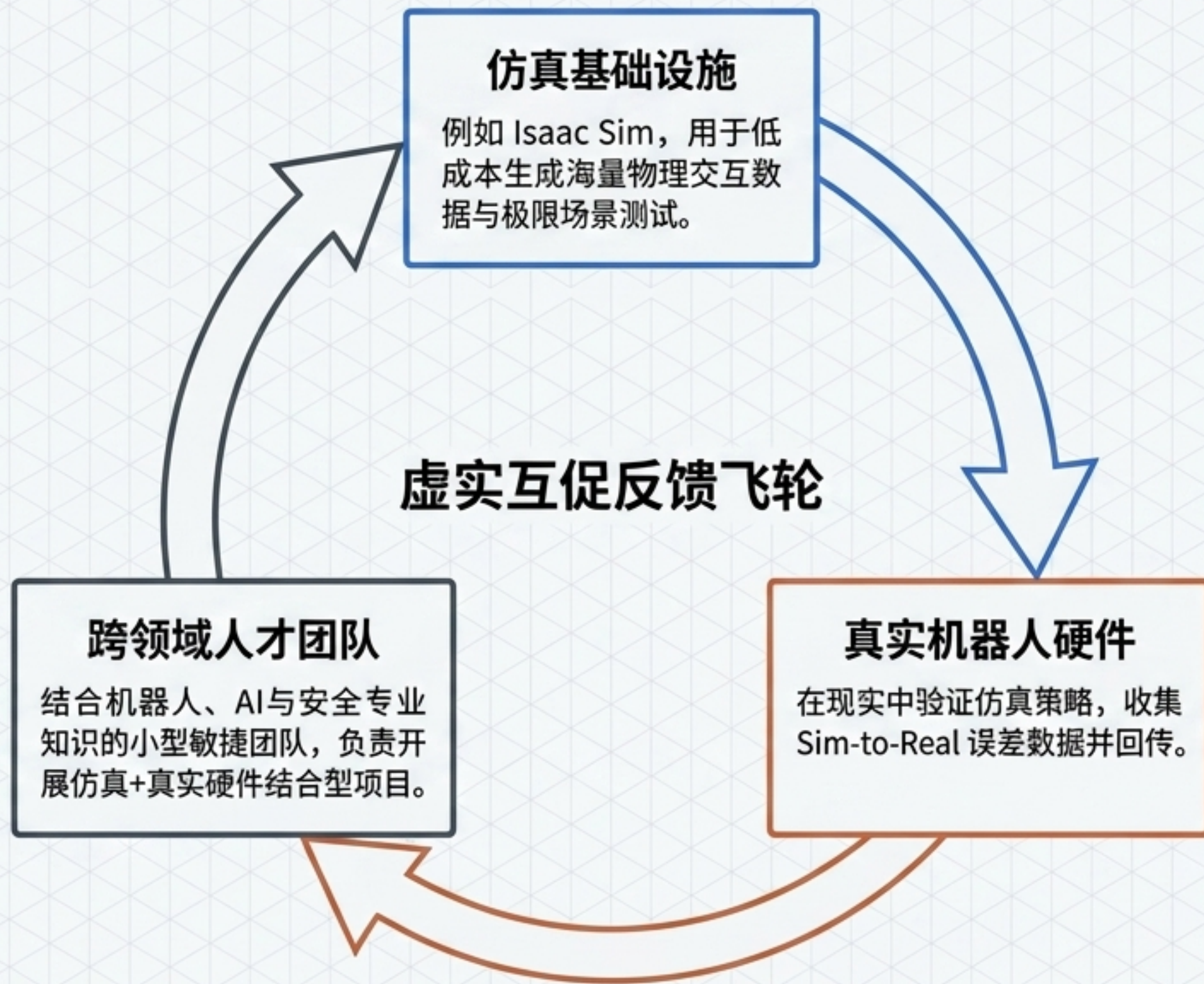
场景：非结构化家庭与公共环境

特征：通用型人形机器人实现规模化运作，物理AI深度融入日常生活。

战略执行手册

	个人开发者	企业与组织
知识与战略	重点学习世界建模概念与混合架构理论。	摒弃立即追求完全通用性，从受控环境中的试点项目切入。
工具与工程	强化 VLA 与 World Model 系统的提示词工程（Prompt Engineering）能力。	采用双线投资策略：同步加码仿真基础设施与真实物理机器人硬件。
生态与团队	持续追踪 WAIC、CoRL、RSS 等顶会前沿论文。	建立结合机器人学、AI与安全专家的跨界团队，并与顶尖世界模型研究机构开展产学研合作。

组织的进化引擎：闭环迭代系统



物理AI的六大核心公理

1. 奠基

LLM 与多模态模型已为具身智能打下坚实底层基础。

2. 局限

纯 VLA 模型在复杂物理环境中的天花板已经显现。

3. 范式转移

World Models 是 2026 年最具颠覆性的技术变革。

4. 现实部署

人形机器人的商业化与物理世界部署正在真实发生。

5. 主导架构

VLA + World Models 的混合系统将在近期占据绝对主导。

6. 长期挑战

跨越非结构化环境的安全与技术深水区依然任重道远。

转变正在发生。



无论你是个人开发者、学术研究者，还是企业决策者，
现在正是布局世界建模技术与混合架构的最佳窗口期。