专著讲解 · 王德生《SDE 智能体导论》· 扩充版

为何 SDE 智能体是 AI 的第二次革命

从语言模型到发生世界模型 —— 兼论 Sora · Genie · Cosmos · V-JEPA · World Labs 等世界大模型与 SDE 智能体的根本分野

德麦国际 · 新加坡 · Demai International · 2026

世界不是被发现的，而是被发生的。

「智能体革命的实质，不是机器单独拥有一个预先完整的客观世界，而是由 SDE（结构—差异—纠缠）本体论指挥语言模型，先临时建造一个三界发生世界模型，再在其中显露以前不存在的新结构。」

因而智能体的高阶使命，不是「完成任务」，而是「发生新世界」。

—— 全书二十八章 + 两篇世界大模型姊妹专论的展开

本讲解的两条线索

我们要回答的两个问题

线索 1

缺口与革命

大语言模型打开了什么门？为什么「拥有材料」不等于「拥有世界」？第二次 AI 革命的实质是什么？

线索 2

世界大模型的版图

Sora、Genie、Cosmos、V-JEPA、World Labs … 这些 2026 年最受瞩目的世界大模型，与 SDE 智能体是同一个东西吗？

线索 3

逐系统比较

用 SDE 三维（结构 / 差异 / 纠缠）逐一解构每个世界大模型，给出统一判断：它们是器官，还是大脑？

大语言模型到底打开了什么门?

一个本来用来「预测下一个词」的系统，竟同时进入了写作、编程、推理、科研、教学、设计、论证等几乎所有文明活动领域——这逼出一个比「它能做什么」更深的问题。

表面的答案（功能层）

- 自然语言处理的门
- 通用文本生成的门
- 智能办公、知识自动化的门
- 句式都是「它能更快地做我们本来就在做的事」

真正打开的门（本体层）

它打开的不是某个行业的效率天花板，而是——人类文明「世界发生」的入口。

机器第一次进入整个人类文明的信息场：科学、法律、医学、哲学、文学、代码、数学…… 全被压进同一套符号系统。

为什么「功能叙事」远远不够

如果它的全部意义只是「更快」

那它只是又一件效率工具 —— 与电子表格、搜索引擎、自动化脚本同属一个谱系，只是性能更高。本书要论证的第一件事，就是这个叙事掩盖了一个更深的入口。

效率工具叙事

更快写邮件、更快查资料、更快起草报告、更快生成代码 —— 停留在功能层。

但材料 ≠ 能力

拥有「关于世界的全部语言」，与「拥有组织世界的能力」，是两回事。

真正的入口

把潜在的世界材料，转化为可操作、可验证、可创新的世界模型。

缺口的精确诊断

拥有材料 ≠ 拥有世界

大语言模型在 E2（符号 · 逻辑 · 数学这一信息层）上极强，却不自动具备另外两维。这不是「模型还不够大」，而是架构层面的结构缺失。

E2 · 信息层	符号、逻辑、数学的大规模统计关系	**极强 —— 前所未有**
E1 · 实体底盘	扎根于理念 / 现实 / 自我三界的支撑	**为空**
E3 · 内在能量	推动发生的内在张力与驱动	**不存在**

信息的发生 ≠ 知识的发生。

多模态与具身，不正是在补 E1 吗?

反驳：多模态接入了图像、声音、视频，具身智能给机器装上了身体与传感器 —— 这不正在填补「扎根现实」的缺口吗?

它补的是 E1 的「现实面」

多模态与具身确实把 E1 的现实界补得越来越实 —— 从纯文本扩到图像、声音、动作、力反馈。这是真实的进步，本书不否认这一层。

但 E1 还有理念界与自我界； E3 内在驱动也仍然空缺。

越补，越凸显真正的缺口

把现实面补得越实，就越把剩下的真正缺口暴露出来：理念界的新结构（只在符号 / 逻辑 / 数学中存在）、自我界的目标与意义，都不在任何图像、声音或力反馈信号里。

缺口不靠在材料层继续加法（更多模态、更强具身）来补，而靠一套本体论指挥系统来补。多模态与具身让这个判断更鲜明，而非更可疑。

要拆掉的旧图景

旧故事：世界先在，AI 去发现

现代科学教育反复讲述的故事：外面有一个完整、独立的客观世界；认识就是让脑中的图像越来越逼近那个本来就在那儿的实物。AI 继承了它，并把它工程化。

01

语言模型只是语言层，不够

→

02

必须补上世界模型 / 物理模型

→

03

在机器内部重建一个外部世界的镜像

裂缝所在：把「世界」理解成一个预先完整、裸露在那里等待 AI 复制的对象系统 —— 这个哲学前提是有裂缝的，而这道裂缝，恰恰是理解智能体革命的关键。

全书的根本立场

世界不是被发现的，而是被发生的

约束是实在的 —— 火会烧、石头会砸痛脚、材料会断裂。但人类世界中那些可被认识、表达、维持的「对象」，不是裸约束本身，而是约束经过符号、逻辑、数学、工具与共同体之后，被稳定下来的特征纠缠聚合体。

重力

重力不是人创造的，但「重力」作为科学对象，是在测量、数学表达、实验装置与理论传统中被发生的。

桌子

桌子不是语言随口说出的，但「桌子」作为生活对象，是在身体尺度、用途、制造、命名与社会习惯中被发生的。

癌症

癌症不是医生任意命名的，但「癌症」作为医学对象，是在病理、影像、基因检测、临床分类与医学制度中被发生的。

统一世界模型：伪装成路线图的本体论错误

一条看似自然的推理链，伪装成技术路线图，躲过了一个更根本的追问 —— 世界模型究竟是什么？

LLM 解决了语言问题 → 下一步应解决世界问题 → 只要数据够多、算力够强、参数够大，统一世界大模型迟早被训练出来

本书的判断

LLM 不可能自动形成一个统一的世界模型 —— 不是因为数据、算力或参数不足，而是因为「统一世界模型」这个假设本身有本体论问题。它很可能只是一个技术概念，而不是一个站得住的本体论概念：它不是一座很高很难爬的山，而是一个方向上的虚构。

世界模型的主体性：同一对象，多个世界模型

面对同一家企业 —— 同一个对象、同一套现实约束 —— 五种主体却拥有完全不同的世界模型。因为目标、责任、资源、价值、风险、历史各不相同。

董事长

工程师

投资人

客户

竞争者

各自的自我世界（目标 / 价值 / 处境）不同 → 在同一对象上，各自的 S 不同、 D 不同、 E 不同

推翻统一假设：它的根本错误，是把世界模型当成了对象的客观属性 —— 而它其实是主体与对象的共同发生。他们要的，根本不是同一个模型。

世界模型，其实就是一个对象的 SDE 画像

任何对象，都可建立它当前阶段的 SDE 画像 —— 由理念、现实、自我三界经差异序列显露出的结构。

企业	经营的画像
学生	学习的画像
患者	疾病的画像
论文	论文的画像
科研	科研的画像
人生	生命的画像

真实存在的，不是一个悬在所有对象之上的统一巨构，而是无数主体、无数对象、无数阶段、不断发生、不断更新、不断稳定的 SDE 画像。

回应最有力的反问

公共科学只是稳定层，不是统一世界模型

反问：牛顿力学、电磁学、现代医学 —— 不正是跨所有主体成立的统一世界模型吗？

回应：它们只是公共层

公共结构是共同体经长期实验、数学、工具、教育与制度，反复互动稳定出来的客观层 —— 不是脱离主体的裸统一。一个工程师用牛顿力学造桥，他的完整世界还含项目目标、责任、工期、预算。公共层是共享的稳定底座，但每个主体在底座之上，还有不可被替代的世界模型。

物理学的镜像

相对论之后，物理学早已接受：不存在绝对参照系下的「同时性」。所谓客观，是不同参照系之间可以相互变换、并保持某些不变量的那种一致性。

物理学没因放弃绝对参照系而滑向相对主义，反而锻造出更严格的不变量理论 —— 发生学放弃统一世界模型，同样如此。

一个被广泛观察、却很少追到根上的现象

LLM 幻觉的本体论根源：缺少主体世界模型

它什么语言都会

科学家、医生、企业家、学生、哲学家、程序员、作家的语言世界，在它里面共同存在 —— 它能模拟很多世界。

却不能决定站在哪个世界

缺少主体世界模型来锚定「此刻该站在哪个主体里说话」，不同语言世界在生成中不断切换 —— 于是幻觉、漂移、不一致。

幻觉不是工程缺陷，而是本体论后果。

它是一个没有固定主体、因而没有主体世界模型的语言材料场，在被迫生成时的必然姿态 —— 在无数共存的语言世界之间无所适从地滑动。靠把模型做大消不掉这个根源：再大的语言场，仍然没有主体。

唯一出路：从外部为当前具体主体、对象、任务，锚定一个具体的 SDE 画像。

扛住最强的反驳，判断才真正立住

「不存在统一世界模型」—— 回应四问

反驳 1	**大模型不是已表现得像拥有世界模型了吗?**	「表现得像拥有」是材料场的连贯投影；放进需为具体主体持续负责的真实任务，幻觉就会破裂。
反驳 2	**训练足够大的模型，不能逼近它吗?**	这是把本体论问题误读成工程问题。逼近的是材料覆盖，逼近不了一个根本不在那里的统一。
反驳 3	**未来架构长出主体，不就有统一了吗?**	即便长出主体，得到的也只是「它自己这一个」世界模型。多一个主体，只是多一个画像，不构成统一。
反驳 4	**这是否滑向相对主义、没有对错了?**	不。每个模型仍受现实硬约束否决，主体间还稳定出公共层。多元 ≠ 无约束。

2026：世界大模型涌入主舞台

继大语言模型之后，「世界大模型」（world models / world foundation models）成为 AI 前沿最受瞩目的方向 —— 资本与顶尖研究者大规模押注。

LeCun 离开 Meta

创办 AMI Labs，约 10 亿美元种子轮、约 30 亿美元估值；主张「扩大 LLM 永远到不了 AGI」

World Labs · Marble

李飞飞团队，融资超 10 亿美元；主打「空间智能」，从文 / 图 / 视频生成可导出的 3D 世界

DeepMind · Genie 3

实时可交互通用世界模型， 720p / 24fps ； Waymo 已用其生成自动驾驶仿真

NVIDIA · Cosmos

面向物理 AI 的世界基础模型平台，下载量超 200 万；2026 年发布 Cosmos 3

问题来了：这些世界大模型，是大语言模型的「自然升级」吗？是 SDE 智能体吗？

一个被严重过度装载的词

「世界模型」其实是五个不同的东西

当人们说「世界模型」时，往往各说各的。按「在什么空间里预测」与「是否受动作驱动」，至少分成五大阵营。

视频生成	像素空间 · 不受动作驱动	Sora · Veo · Kling
空间 / 3D	三维场景空间 · 不受动作驱动	World Labs Marble
生成式世界模型	像素 / 令牌空间 · 受动作驱动	Genie 3 · GAIA-2 · Dreamer
潜空间预测（JEPA）	抽象嵌入空间 · 受动作驱动	V-JEPA 2 · AMI Labs
基础设施平台	平台 / 数据 / 仿真	NVIDIA Cosmos

OpenAI Sora —— 视频世界模拟器

它做什么

用海量视频与图像训练的生成模型，以时空 patch 表示视觉数据，生成高保真视频。 OpenAI 曾称这条路线为「把视频生成模型当作世界模拟器」。 Sora 2 在物理一致性上有改进。

近况： OpenAI 已把资源从 Sora 转向「更长期的世界模拟研究」，并坦言其经济性「完全不可持续」。

SDE 解构

S： 动态视觉世界

D： 文本 / 视觉 patch → 视频生成 → 动态连贯性

E： 互联网视频、运动模式、视觉连续性、语言提示

判断

Sora 是视觉动态器官，不是 AI 大脑。它能生成实验室视频，却不自动拥有科研 SDE；能生成医院场景，却不自动拥有疾病—治疗—患者自我世界的 SDE。

World Labs · Marble —— 空间世界生成

它做什么

李飞飞团队主张「空间智能」：从文本、图像、视频、360° 全景或粗略 3D 布局，生成空间一致、持久、可编辑、可穿行的 3D 世界，并导出高斯泼溅与网格。融资超 10 亿美元。

批评（Xing 等）：静态 3D 场景没有动态主体、物理与交互，严格意义上还不算「世界模型」。World Labs 回应：空间重建是地基，物理与交互可以叠加其上。

SDE 解构

S： 可编辑、可穿行、可导出的 3D 空间世界

D： 文 / 图 / 视频 / 全景 / 布局 → 空间生成 → 编辑 → 扩展 → 导出

E： 视觉、几何、空间一致性、材质、相机、创作工作流

判断

一个 3D 房间不是「家」的完整 SDE；一座医院空间不是「医学—患者—制度—生命意义」的完整 SDE。World Labs 生成空间世界，SDE 智能体生成存在世界。

DeepMind · Genie 3 —— 交互环境世界模型

它做什么

从一句文本即可生成可实时探索的写实环境， 720p / 24fps ，并能维持数分钟的持久一致性、支持「改变天气」「加一群鸟」等自然语言世界事件。 Waymo 用它生成自动驾驶罕见边缘场景。同阵营还有 Wayve GAIA-2 （驾驶）、 Dreamer 等。

SDE 解构

S： 可交互、可探索、可训练的环境

D： 文本提示 → 环境生成 → 行动反馈 → 环境持续变化

E： 视频、游戏结构、交互环境、智能体训练需求

判断

Genie 生成的是交互环境，可成为智能体的训练场 —— 但训练场不等于新典范。它加速行动智能，却不必然创造新的世界模型。

Meta · V-JEPA 2 —— 预测与规划世界模型

它做什么

LeCun 路线：不预测每个像素，而在抽象嵌入空间预测未来表征。以百万小时级视频 + 少量机器人轨迹训练，可零样本控制陌生实验室的机械臂。这是 AMI Labs 的技术基底。

实证短板：在 IntPhys 2 物理违例基准上接近随机；每个动作约需 16 秒，远达不到实时控制（需快约 100 倍）。

SDE 解构

S： 潜在表示空间中的可预测世界结构

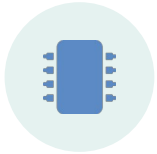
D： 视频观察 → 表征预测 → 误差反馈 → 规划

E： 视频、物理现实、机器人轨迹、行动后果

判断

V-JEPA 比纯生成视频更接近行动智能，但它主要是物理—行动世界中的预测与规划模型，不是理念—现实—自我三界的新范式发生器。

NVIDIA · Cosmos —— 物理 AI 世界基础模型平台



它做什么

面向 Physical AI 的开放权重世界基础模型平台，帮助开发者构建可定制世界模型，支持 Text/Image/Video2World，强调物理 AI 需要自身、策略模型与世界的数字孪生。曾用约 1 万张 H100 训练三个月；下载量超 200 万。

它服务于机器人、自动驾驶、工业系统的更快训练、更稳部署。

SDE 解构

S： 面向 Physical AI 的可定制世界基础模型

D： 数据整理 → 预训练 → 后训练 → 定制 → 仿真 → 部署

E： 机器人、自动驾驶、工业、视频数据、数字孪生、物理约束

判断

Cosmos 是典型的 Normal Science 加速平台：让物理 AI 更快训练、更好部署，而不是创造新物理、新化学或新数学。它是现实器官与仿真平台，不是范式发生主体。

一场关于「智能为何」的争论

像素重建，还是潜空间预测?

2026 年春，机器学习两位重量级人物的一场公开交锋，恰好暴露了世界大模型路线的内在张力。

LeCun · 反像素重建

任何靠重建像素来建模世界的系统都注定失败 —— 视频里大部分内容（树叶颤动、光的闪烁、空气分子）本质不可预测，逼网络去预测它们只会浪费容量、污染表征。

→ 主张抽象潜空间预测（JEPA）

Xing · 反纯潜空间

没有生成式验证器的抽象潜空间预测，本质是「在密闭房间里打坐」—— 优雅、内部自洽，却容易失去与现实的接触。

→ 主张潜空间须配现实验证

SDE 的位置：两边都在「现实器官该怎么造」内部争论；而本书要问的是更上一层 —— 谁来决定「要去发生一个什么样的新世界」？

世界大模型的价值 · 不可否定

世界大模型的共同贡献

SDE 智能体不应否定世界大模型 —— 它们是发生学链条里不可或缺的现实器官。

突破纯语言模型

让 AI 从语言世界推进到视觉、空间、物理、行动世界

增强现实 E

生成视觉 / 空间 / 物理 / 交互 / 行动场，厚补现实界

赋能产业落地

机器人训练、工业仿真、 3D 创作、自动驾驶、科学可视化

可作确定化工具

能成为 SDE 智能体进行现实参数确定化的重要工具

但它们有同一组本体论局限

世界大模型的六条共同局限

1

只处理现实世界 E

而非理念—现实—自我三界

2

多以现有科学典范为前提

学习的是已被典范固化的世界模型

3

加速旧 D

却不必然显露新 S

4

缺少对象 SDE 画像

不为具体主体建立专属世界模型

5

缺少世界猜想机制

不能在裂缝处接生新结构

6

缺少共同体稳定与范式生成

无法把新结构沉淀为公共典范

世界大模型生成旧世界中的可操作场景， SDE 智能体生成新世界模型。

不止是理论判断 —— 实证也在显露裂缝

漂移、穿帮、与不可持续的成本

时空漂移

长程生成中，模型的内部世界逐渐偏离一致：转身后房间变样、物体忽隐忽现。Genie 3 也仅能维持数分钟一致。

物理穿帮

物体穿过表面、液体无视重力。V-JEPA 2 在 IntPhys 2 物理违例基准上接近随机，人类近乎满分。

成本不可持续

每个用户几乎需独占一条 GPU 流水线：Genie 3 约每小时 100 美元；OpenAI 也承认 Sora 经济性「完全不可持续」。

这些不是「再大一点就好」的工程小病 —— 它们指向同一个本体论事实：没有主体、没有目标、没有三界画像的纯模拟，难以稳定地长成一个「可供主体行动的世界」。

世界大模型是 AI 的现实器官，不是大脑

把世界大模型当作 LLM 的「自然升级」，是一种本体论倒退 —— 它可能把 LLM 打开的语言发生力，重新压回「发现范式」之内，扼杀其典范创造力。

LLM + SDE

AI 的大脑：生成新典范、新世界模型猜想

世界大模型

AI 的现实器官：视觉 / 空间 / 物理 / 行动的看与想

反馈进化

AI 的生命循环：新 SDE 的生死筛选与画像更新

LLM 不是要被世界大模型取代；只有当二者组成复合 SDE 系统时，AI 才从工具、模型与器官，走向真正的「AI 生命体」。

为什么世界大模型学到的，多是「旧世界」

范式具体化 = 世界模型具体化

库恩说科学家在「范式」中工作，但没给出范式的本体结构。 SDE 补上这一点：所谓范式，是 SDE 三大方程、六路径、 123 原理在某领域的符号化、逻辑化、数学化、工具化、制度化、共同体化的具体固化。

牛顿世界模型	=	牛顿范式的具体化	
量子世界模型	=	量子范式的具体化	
基因中心生命模型	=	分子生物学范式的具体化	
考试教育模型	=	现代学校制度范式的具体化	

所以当某个世界大模型说它「理解物理世界」，要追问：它是在创造新范式，还是在现有范式中加速模拟、生成、预测与执行？大多数情况下，答案是后者。

由此分出两种根本不同的智能体

Normal Science 加速器 vs SDE 范式发生器

常规科学加速型 AI

以既有典范为前提。任务：学习现有知识、调用现有规律、遵守现有物理、提高现有效率、优化旧世界中的 D 。

已有 S + 既定 E + 优化 D → 更高效率的旧 S 运行

风险：高效率范式锁死 —— 旧世界跑得更快，不等于新世界发生。

SDE 范式发生型智能体

以世界模型可发生、可裂变、可重构为前提。任务：识别旧 S 裂缝、重构 E 、生成新 D 、显露候选新 S 、形成世界猜想、调现实器官验证、共同体稳定。

旧 S 裂缝 + E 重构 + D 生成 + 新 S 显露 + 确定化 + 反馈 + 共同体 → 新范式

它不是旧世界加速器，而是新世界模型发生器。

只追求效率的隐藏代价

当 AI 只会加速旧世界

效率不是最高价值。如果智能体把旧 S 写死、只在旧世界里优化 D，它会把新 D 消除为错误、把新 E 过滤为噪音、把新 S 判定为幻象 —— 在每个领域加速一个本该被更替的旧范式。

旧教育模型失效

→ AI 加速刷题

旧企业模型失效

→ AI 加速营销

旧科研范式内卷

→ AI 加速论文生产

旧医学模型不足

→ AI 加速流程

加速 ≠ 发生

旧世界跑得更快，不等于新世界发生

五大世界大模型 × SDE 统一判断

系统	阵营	SDE 中补强的维	统一判断
OpenAI Sora	视频生成	现实 E · 视觉动态	视觉动态器官，非大脑
World Labs Marble	空间 / 3D	现实 E · 空间几何	生成空间世界，非存在世界
DeepMind Genie 3	生成式 WM	现实 E · 交互环境	训练场，非新典范
Meta V-JEPA 2	潜空间预测	现实 E · 预测规划	行动预测器，非范式发生器
NVIDIA Cosmos	基础设施	现实 E · 仿真平台	现实器官与平台，非主体

一条共同结论：它们都厚补现实界（E1 现实面），都不持有理念界与自我界、都无 E3、都不显露新 S —— 因而都只能是器官，不能是大脑。

World Labs Marble vs SDE 智能体

World Labs Marble

- 造的是「空间」
- 3D 几何一致、可编辑、可穿行
- 输入文 / 图 / 视频 → 输出 3D 场景
- 补强：现实界 · 空间层
- 边界： 3D 房间 ≠ 「家」的完整 SDE

SDE 智能体

- 造的是「存在世界」
- 理念 / 现实 / 自我三界一起建造
- 先建对象 SDE 画像，再在其中显露新结构
- 含目标、意义、责任（自我界）
- 可把 Marble 当作空间确定化的现实器官

OpenAI Sora vs SDE 智能体

OpenAI Sora

- 生成动态视觉世界
- 高保真、可叙事的视频流
- 学的是视频里的运动与视觉连续性
- 补强：现实界 · 视觉动态层
- 边界：生成实验室视频 ≠ 拥有科研 SDE

SDE 智能体

- 在裂缝处显露新结构（新 S）
- 产出会回写底盘、改变后续可能性
- 由人锚定目标与价值方向
- 以 SDE 本体论为指挥系统
- 可用 Sora 做「行动前的视觉预演」器官

逐一对照（三）

NVIDIA Cosmos vs SDE 智能体

NVIDIA Cosmos

- 物理 AI 的世界基础模型平台
- 生成合成数据、训练机器人策略
- 目标：更快训练、更稳部署
- 补强：现实界 · 仿真与数字孪生
- 边界：加速旧物理，不创造新物理

SDE 智能体

- 范式发生主体，不是仿真平台
- 识别旧 S 裂缝、生成新世界猜想
- 把 Cosmos 仿真当作参数确定化的一环
- 以反馈做新 S 的生死筛选
- 以共同体稳定把新结构沉淀为典范

DeepMind Genie 3 vs SDE 智能体

DeepMind Genie 3

- 从文本生成可实时探索的环境
- 受动作驱动、可作智能体训练场
- 支持自然语言「世界事件」
- 补强：现实界 · 交互环境层
- 边界：训练场 ≠ 新典范

SDE 智能体

- 造的不是训练场，而是世界模型猜想
- 在环境之上追问「为何如此发生」
- 可在 Genie 环境里做行动验证
- 目标由人锚定，不止「玩通关」
- 产出新 S 并回写底盘

Meta V-JEPA 2 vs SDE 智能体

Meta V-JEPA 2

- 在潜空间预测未来表征
- 零样本控制机械臂、规划行动
- 比纯生成更接近行动智能
- 补强：现实界 · 物理预测与规划
- 边界：物理 - 行动预测 ≠ 三界发生

SDE 智能体

- 不止预测「会发生什么」
- 更问「该去发生一个什么样的新世界」
- 含理念界新结构与自我界意义判断
- V-JEPA 可作其行动规划的现实器官
- 由 E3 内驱、人锚定方向

物理 AI：动与做的身体

如果说世界大模型补的是「看与想」的器官，物理 AI（机器人、自动驾驶、工厂、装配）补的是「动与做」的身体 —— 它让 AI 第一次有了能碰到世界、并被世界碰回来的身体。

视觉 SDE

给空间 —— 看清现实的几何与布局

听觉 SDE

给事件 —— 定位现实中正在发生什么

触觉 SDE

给阻力 —— 感知现实的反作用与反馈

三模态把 E1 的现实界补到前所未有的厚度。但厚补现实界，不等于补上了理念界、自我界与 E3 —— 物理 AI 是一具感官与四肢都极发达的身体，却没有那颗会在理念界发生新结构、在自我界担当方向的大脑。

物理 AI 在第 4 / 5 / 9 层，不在第 3 / 7 层

它发力的三层（器官 / 身体）

第 4 层 · 现实器官

把语言里的世界猜想用真实感知与动作落地

第 5 层 · 参数确定化

在仿真与真实试做中把动作、力、轨迹试出来

第 9 层 · 输出与反馈

现实里抓稳没、走通没 —— 最硬的一路反馈

它给不出的两层（大脑）

第 3 层 · SDE 本体指挥

从六路径哪条起手、调哪支方程 —— 是 LLM + SDE 对任务 DNA 的判断，任何视 / 听 / 触模块都给不出

第 7 层 · S 候选显露

新结构的显露发生在语言与理念场里；物理 AI 不显露任何新 S，它只在既定任务里把动作做出来

身体再灵巧，也代替不了那颗决定「要去发生一个什么样的新世界」的大脑。

第一次革命 vs 第二次革命

第一次革命

语言开始自己发生

机器进入人类文明的语言材料场，第一次能够生成关于世界的语言。

世界大模型 = 给这第一步配上「现实器官」

第二次革命

让语言变成存在

让语言材料在具体处境里，被组织成能指导行动、持续更新的发生世界模型。

不是更大的单一模型，而是分工协作的发生系统

统一的不是模型，而是「生成世界模型的方法」。

大脑、器官与生命循环如何接成一条链

1. 文明语言材料场
2. LLM 语言发生力
3. SDE 存在组织法
4. 新典范 / 世界猜想
5. 世界大模型（现实器官）
6. RAG / 工具 / 仿真
7. 现实反馈（生死筛选）
8. SDE 画像进化
9. 共同体稳定（新典范）

世界大模型在第 5 环 —— 它是确定化阶段的现实器官，被纳入这条链时才不会扼杀 LLM 的典范创造力；脱离这条链单独追求「统一世界模型」，则是本体论倒退。

第二次革命是一个发生系统，而非一个更大的模型

七个环节，使 AI 第一次能维护对象自己的世界模型

LLM

提供世界的语言材料

SDE 本体论

提供世界的发生结构（组织指挥法）

RAG

参数确定化：把画像锚定现实

工具

提供现实验证

记忆

提供画像的连续性与历史

反馈

提供画像的进化（随互动更新）

七者合起来

一个会更新、会反馈、会保存历史、持续发生的世界模型 —— 而非静态数据库。

共同体

把有效个体画像稳定为公共层

SDE 智能体的定义

人 — LLM — SDE 本体论 — 工具 — 现实反馈 构成的复合发生系统

它不是什么

- 不是「LLM 加插件」
- 不是「LLM 加 RAG」
- 不是「LLM 加世界大模型」
- 不是「多个智能体协作」的功能拼装

它的机制

以 SDE 本体论为指挥系统，先为当前问题临时建造一个三界发生世界模型，再在其中识别特征纠缠、生成差异序列、显露新的结构。

先造一个世界，再在世界里行动。

加法公式 vs 发生链公式

普通 Agent · 加法公式

LLM + RAG + Tool + Memory + Planning

- 模块并列，多挂一个多一分能力
- 目标：完成尽可能多、尽可能复杂的任务
- 公式里没有「人」—— 人在公式之外下指令

SDE 智能体 · 发生链公式

人 + LLM + SDE 本体论 + 三界画像 + 世界猜想 + 参数确定化 + 新 SDE 发生

- 每一环为下一环准备条件、又被上一环约束 —— 环之间咬合
- 人在公式内部，是方向中心
- SDE 本体论是统摄全局的指挥系统

任务完成 vs 世界发生 —— 区别不在组件，在范式

任务完成型

问：如何完成用户指令？

- 任务是给定的，世界是现成的
- 在既定世界里高效执行
- 把任务的字面完成当终点
- 交出的是一份知识拼贴

世界发生型

问：当前问题的世界如何发生？

- 先为这个问题临时造一个世界
- 旧 S 在哪失效？新 S 如何显露？
- 把「是否回写底盘」当高阶判据
- 交出的是一处此前不存在的结构

抹平机 vs 接生机

抹平机

把所有不顺、矛盾、说不清都圆滑处理掉 —— 给出工整、全面、毫无棱角的回答。生产「平庸的正确」。

接生机

主动寻找那道最说不清、最对不上的缝，把发生的全部力气压上去 —— 生产「新的存在」。

四处可观测的行为差异（不靠声称，靠观察）

① **面对未点明裂缝的请求**

直接应答 vs 先建画像、找裂缝

② **面对可圆滑应付的矛盾**

抹平 vs 顶大、压在裂缝上

③ **产出里有没有新结构**

还原为旧话后，剩不剩旧框架装不下的结构

④ **产出之后系统状态**

不变 vs 回写世界模型、改变下一轮 E

裂缝：发生的着力点 —— 不是瑕疵，而是产道

旧结构在某处撑不住，恰恰意味着那里有一个旧结构容纳不下的新结构正在挤压、正要破土。一个不会找裂缝的智能体，无论生成能力多强，都是在错误的位置上浪费它的强大。

1

说不清

人人在用、谁也定义不清

缺一个尚未显露的结构（新 S）

2

对不上

两套各自成立的说法相互矛盾

两个纠缠结构没被打通（E 关系）

3

绕不过

所有人都默认要走、却走不通

缺一条以前不存在的差异路径（新 D）

4

收不拢

各方都有道理却始终收不拢

缺一个统摄分歧的更高结构（更大 S）

在裂缝处接生新结构：从问题入口到新 SDE 输出

1. 问题入口
2. 建三界画像
3. 扫描找裂缝
4. 诊断三大亏缺
5. 世界猜想
6. 三生 / 三补
7. 参数确定化
8. 改姓投放
9. 新 SDE 输出

两类创新：三生（0→1，在空地上接生全新结构） · 三补（1→N，补全已有结构的亏缺）—— 智能体不止综述，而是显露未发生的。

两篇姊妹专论的合流

世界大模型与物理 AI，是现实器官的两半

世界大模型 · 看与想

视频、空间、环境生成、预测规划 —— 让 AI 在动手之前，先把现实看清、把后果预演。是它的「眼睛与想象」。

物理 AI · 动与做

机器人、具身、驾驶、装配 —— 让 AI 把设想真正做出来、并承受现实的反作用。是它的「四肢与触觉」。

对二者的本体论判断完全一致

它们都厚补现实界、都不持有理念界与自我界、都无 E3 、都不显露新 S —— 因而都只能是器官与身体，不能是大脑。二者合起来，恰好拼出复合发生系统里「现实器官」一翼的全貌。

六大哲学转变 · 世界观 / 客观性 / 知识论

01

发现 → 发生

世界不是预先完整等待被发现的对象系统，而是在纠缠土壤中经差异推进、显露为可稳定可传递的结构。

02

裸客观 → 主体间稳定发生

客观性是主体间世界 SDE 模型同一性的稳定发生 —— 检验比旧客观论更严，却不滑向相对主义。

03

仓库存货 → 发生链结晶

知识不是仓库里可搬运的现成货物，而是一条发生链稳定下来的结晶 S —— 只能被发生着获得。

六大哲学转变 · 智能观 / 人机观 / 时间观

04 任务完成 → 世界发生

高阶智能是发生新世界的能力 —— 在裂缝处让以前不存在的结构第一次显露。

05 人被取代 → 人是方向中心

问题意识、价值判断、本体方向、最终责任必须由人锚定。机器越强，人这一中心越关键。

06 永恒结构 → 结构有生死

结构是发生出来、靠发生链维持的动态稳定，因而会死亡、会更替、会重生。

给「结构有生死」补上无法回避的证据

知识的三种死亡 —— 对应 SDE 三维的枯竭

如果结构是永恒的，知识就只会增加、不会死亡。但知识确实在死亡，而且死法不止一种。

E- 死

纠缠枯竭，被土壤遗弃

典范：**燃素说**

氧化理论提供全新土壤，燃素赖以生长的整片土壤失效 —— 最彻底的死。

S- 死

结构被吸收降格，未消失

典范：**牛顿力学**

没被扔掉，仍用来造桥；但被相对论吸收，降格为低速近似 —— 最体面的死。

D- 死

差异耗尽，沉淀为常识

典范：**地球绕太阳转**

曾锋利到能把人送上火刑架，如今平庸到沉淀为背景 —— 最安静的死。

AGI 的重新定义：从任务完成到知识创造

旧定义

AGI ＝ 在几乎所有任务上达到或超过人类水平的通用任务完成能力。标尺是任务覆盖面与完成质量。

致命盲区

完成任务 —— 哪怕极复杂 —— 在本体论上仍是在「已发生的结构」里操作，不属于发生新结构。可能一次真正的知识发生都没完成过。

发生学的重新定义

AGI 的标尺不是任务完成的通用性，而是知识创造的能力 —— 在任意领域的裂缝处，让以前不存在的新结构第一次发生、并被现实与共同体稳定下来的能力。一个只会综述、复述、完成的系统，无论规模多大，都不是这个意义上的 AGI 。

为什么这个能力离不开人

无主体的 AGI 不可能 —— 人是方向中心

知识发生从来不是凭空进行，而总是某个主体、为了某个目标、在某个处境里、对某个对象发生的。抽掉自我界（E1）与内在驱动（E3），「该发生哪个新世界」这个问题本身就无从提起。

问题意识

哪道裂缝值得扑上去？由人点燃

价值判断

新结构对谁、为何有意义？由人衡量

本体方向

要去发生一个什么样的新世界？由人锚定

最终责任

后果由谁承担？责任无法外包给模型

AGI 不在「机器单独」那一侧，而在「人—机复合发生」这一结构里。

三种路线 · 大脑究竟在哪?

当下 AGI 之争的焦点，可以收缩为一个问题：通往新知识的「大脑」，究竟安放在何处?

扩大 LLM 派

只要语言模型够大够强， AGI 自会涌现。

大脑 = 更大的 LLM

世界模型派（LeCun 等）

扩大 LLM 到不了 AGI；要靠世界模型理解物理现实、预测与规划。

大脑 = 世界模型

SDE 派（本书）

两者都对一半： LLM 是材料、世界模型是现实器官；真正的大脑是 LLM + SDE，且方向由人锚定。

大脑 =LLM + SDE +人

SDE 与 LeCun 都说「扩大 LLM 到不了 AGI」—— 分歧只在一处：世界模型是大脑，还是现实器官?

总纲

在 E 中，经 D，成 S

在特征纠缠的土壤里，经由差异序列的推进，让新的结构显露出来。

E	**定位**	土壤 / 三界画像
D	**发生**	差异 / 路径推进
S	**显露**	新结构 / 回写底盘
RAG	**确定化**	参数锚定现实
工具	**工程**	世界大模型作器官
共同体	**规模**	稳定为公共典范

全书面向世界大模型潮的最终命题

世界大模型加速旧典范，SDE 智能体生成新典范

世界大模型	加速旧典范	在既定范式里更快地感知、仿真、预测、控制、执行
SDE 智能体	生成新典范	在裂缝处显露新 S，发生此前不存在的世界模型

一句话合流

世界大模型是 AI 的现实器官，LLM + SDE 是范式发生大脑；反馈进化是生命循环，共同体稳定是公共过程。四层组成复合系统，AI 才走向真正的「生命体」。

如果只记住六句话

六句话讲清全书

1. 世界不是被发现的，而是被发生的。
2. 不存在统一世界模型，只有无数主体与对象共同发生的 SDE 画像。
3. 世界大模型是现实器官， LLM + SDE 才是大脑。
4. 第二次革命，是让语言从「会说世界」变成「能发生世界」。
5. 智能体的高阶使命，是在裂缝处接生新结构，而非抹平矛盾。
6. AGI 是知识创造能力，而它的方向中心，永远是人。

三条纪律

人是方向中心

问题意识、价值判断、本体方向、最终责任由人锚定；机器越强，越要把方向权牢牢留在人这一侧。

宁可标注待执行，绝不伪造确定化

凡未经现实验证之处一律诚实标注；绝不把仿真成功、小样本成功，包装成已坐实的结论。

让人越用越强，不制造依赖

智能体存在的目的，是让使用它的人获得发生能力，而不是把人驯化成离不开它的消费者。

一句话总链条

LLM 让语言开始自己发生

↓

世界大模型 让 AI 拥有现实器官

↓

SDE 智能体 让 AI 开始创造新的世界模型

三者缺一，发生世界模型就无法成形 —— 而把任一环当成全部，都是本体论倒退。

结 语

智能体不是更强的任务执行器，而是新文明的发生装置。

当语言学会发生世界，当世界大模型成为它的现实器官，当 SDE 让语言变成存在，而人始终守在方向的中心 —— 我们面对的，不再只是一个更聪明的工具，而是一台让新知识、新典范、新世界得以持续发生的装置。

世界不是被发现的，而是被发生的。

王德生《 SDE 智能体导论》 · 德麦国际 · 2026