

以“数据飞轮 + 强化学习”范式为主体的企业 AI Infra 创新技术架构

杨琴

华为技术, 广东 深圳 518129

DOI:10.61369/ME.2026040030

摘 要 : 2025 年被行业普遍视作智能 Agent 规模化落地的关键年份, 人工智能技术摆脱传统对话框交互形态, 逐步转向可自主规划、调用工具、自主纠错的智能体落地模式。当前不少企业不再局限于利用提示词生成基础文本内容, 开始批量落地自研或外购智能体项目。结合麦肯锡 2025 年末行业调研数据, 全球超半数企业已在部分业务线常态化落地 AI 智能体产品, 近 39% 的头部企业可在生产环境并行部署十余套智能体协同系统。本文按照自主执行能力将市面商用 Agent 划分为流程编排型、半自主决策型、高度自主型三类, 逐一剖析不同层级产品落地特点与产业痛点; 文章指出大模型是智能体的决策中枢, 各类开发框架与配套工具承担落地执行职能, C 端通用 AI 依托海量公开预训练数据落地顺畅, 但 B 端企业受复杂业务流程、老旧非标系统、行业隐性私有知识约束, 通用大模型极易产生认知偏差与内容幻觉。依托银行风控、车险定损、电力信息抽取三类落地实例验证, 数据飞轮搭配强化学习能够针对性补齐垂域模型短板; 从数据工程、模型工程两大维度搭建迭代闭环, 数据端实现行业隐性经验标准化沉淀, 模型端分初、中、高三阶段循序渐进完成从应用落地到底层模型迭代的全流程建设。文末提出, 伴随 AI 深度融入企业核心业务, 企业的业务、应用、信息、技术四层架构将迎来全方位改造, 逐步形成人机协同运行的新型学习型组织。

关 键 词 : AI 智能体; 数据飞轮; 强化学习; 企业 AI 基建; 大模型微调

Innovative Architecture of Enterprise AI Infrastructure Based on the Paradigm of Data Flywheel plus Reinforcement Learning

Yang Qin

Huawei Technologies, Shenzhen, Guangdong 518129

Abstract : The year 2025 is regarded by the industry as a critical period for the large-scale commercial deployment of intelligent Agents. Breaking away from conventional chatbox interaction, artificial intelligence has evolved into intelligent agents capable of autonomous planning, tool invocation and self-correction. Instead of merely generating simple texts via prompts, numerous enterprises have invested in self-developed or outsourced Agent projects. According to McKinsey's year-end survey in 2025, over half of global corporations have deployed AI Agents on regular basis across partial business lines, and nearly 39% of leading enterprises operate more than ten interconnected Agent systems in production environments. Classified by autonomous performance, commercial Agents are divided into three categories: process-programmed type, semi-autonomous decision-making type and highly autonomous type, with respective application features and industrial drawbacks analyzed in this paper. Large language models serve as the decision core of Agents while relevant development frameworks and tools undertake executive functions. Benefiting from massive public pre-training datasets, consumer-oriented AI gains smooth implementation, whereas enterprise-oriented AI is restricted by complicated workflows, legacy non-standard IT systems and private industrial know-how, which readily lead to cognitive bias and model hallucinations. Verified by practical cases covering bank risk control, auto insurance loss assessment and power industry information extraction, the combination of data flywheel and reinforcement learning effectively remedies defects of vertical domain models. An iterative closed-loop system is constructed from data engineering and model engineering: implicit industrial experience is standardized via data engineering, and model construction proceeds in three phases to realize full evolution from practical application to underlying model optimization. In the end, deep integration of AI will reshape enterprises' business, application, information and technical architectures, facilitating the establishment of new learning organizations featured with human-AI collaboration.

Keywords : AI agent; data flywheel; reinforcement learning; enterprise AI infrastructure; LLM fine-tuning

一、企业 AI 不同 Agent 的自主化程度存在较大差距：

第一类：流程编排型 Agent。所有执行步骤按标准流程预先编排，工具调用都是硬编码写死的，Agent 自主性主要体现在文本生成，但是需要克制幻觉，模型精度有限。（目前 70% 以上的行业 AI 应用都是处于这个阶段）

- 知识问答型 Agent：2025 年开始企业普遍采用“开源大模型 + 提示工程 + RAG”的技术组合最小集构建“类 Chatbox”Agent，打造通用文档生成助手、行业场景知识问题助手等应用，为员工的工作知识和技术能力提升提供一个托底的技术支撑，实现企业内部的提质增效。

第二类：半自主规划 / 执行 Agent。在部分企业场景下具备自主思考和执行决策，并调用工具（工具预定义），Agent 执行精度达到企业级任务基本要求。（行业 AI 应用占比 20%）

- AI 编码 Agent：2023~2024 年 AI 编程仅能实现代码自动补全，到 25 年已经 Cursor、Claude Code 等编码工具让程序员用自然语言明确任务需求和范畴，单需要容忍 AI 可能出现的错误（70~80% 正确率）

第三类：高度自主 Agent。能够完全基于目标达成自主决策执行步骤，能够自主创造工具并使用，例如 OpenClaw 实现数字员工、自动驾驶等持续迭代迈向高度自主智能体。（仅不到 10% 的 Agent 接近此阶段）

- 自动驾驶：通过大量具备 L2 辅助驾驶能力的商用车开启“影子模式”（Shadow Mode），源源不断传回的“Corner Cases / 错题集”，成为喂养 L3~L5 自动驾驶大模型的的绝佳数据燃料，实现算法渐进式爆发的核心技术闭环。



二、企业 AI Agent：大模型能力是“大脑”，Agent 框架 / 工具是“手”

- 大模型能力是“大脑”：理解任务意图、推理与规划、指挥工具调用、执行过程反思

- Agent 框架 / 工具是“手”：工具调用执行、环境感知外部信息、记忆管理等

2024~2025 年全球主流的基础大模型能力主要的发展方向是 Agentic 智能体能力，包括任务完成能力、规划质量、工具使用、鲁棒性 / 反思、长程记忆等。大模型 Agentic 能力是决定能否高效、稳定地把任务做成的基础。在开放性的 2C 场景下，Agent 业务逻辑和调用工具非常简单，因为互联网的基础设施是高度标准化的，可以基于公开的、人类常识性的大规模文本预训练出来的（Pre-trained）。随着基础大模型 Agentic 智能体能力快速迭代进化，大量 2C 的 Agent 应用体验得到极大的提升，涵盖个人助理、

生产力工具、创意娱乐、生活服务等多个领域。例如 Agent 能够提高个人生产力，自主处理邮件、安排会议、跨应用同步数据，甚至能代表用户进行初步的市场调研。Agent 打理用户的生活服务与消费，可以能够直接在浏览器中模拟人类操作，完成订餐、订票、比价购物及自动填写复杂表单。

但是在企业 2B 专业化的应用场景与 2C 的通用化、生活化的应用场景存在差别。大模型不理解企业专业性的业务场景与处理逻辑，存在领域认知鸿沟。

- 企业业务流程复杂，存在各种前置条件分支的企业业务流程；

- 企业内部的文档和语言环境充斥着大量的缩写、代号以及复杂的跨表格关联，缺乏对业务上下文深刻理解的大模型极易断章取义，大模型极易产生幻觉；

- 企业复杂的内部 IT 工具调用，企业内部运行着大量早期开发的 ERP、OA、CRM 等系统，缺乏标准 API 接口的 IT 业务软件等，调用存在强约束与容错率极低；

- 行业历年积累的行业 know-how（行业暗知识）都隐藏在私有数据内。基础大模型在预训练阶段，垂直行业的训练数据不够，真实业务场景的后训练数据也很少，大模型企业级的业务理解和规划能力无法“先天养成”，需要在企业内部使用过程中，通过微调 / 强化学习“后天培养”。

三、“数据飞轮 + 强化学习” 范式为企业 AI 注入“灵魂”

- 银行：通过强化学习 CoT 能力，提升企业风控助手查准率和召回率。原来大模型缺乏信贷风控管理业务逻辑，预测精度很低，通过模型蒸馏 + GRPO，基于业务反馈优化垂域思维链，提升预测精度了 70%，。大模型调优技术成为金融行业场景应用推广的标准范式。

- 保险：通过 GRPO 多模态强化学习，在辅助定损环节，让多模态大模型能够更好的做多图联合思考推理能力，对驾驶员上传的照片证据序列、维修清单来进行图像分类、零件识别、损伤区域 / 类型 / 尺度的识别准确率提升 20%。

- 电力：电网运行中从供电、交付、运维到运营过程中，大量专业的关键信息需要识别，抽取和标注。通用基础大模型无法理解领域专业知识，及背后意图。通过难易聚类 + 多维奖励机制，可以更接近实际的提升学习能力。最终在信息抽取准确度和信息关联理解方面均有 30% 的提升，增强信息归类识别方面专业理解力。

四、用“数据工程 + 模型工程”一起转动 AI 迭代飞轮

既然提升大模型的企业级能力不能依赖一个更大的通用预训练模型来实现。需要让大模型在企业真实场景里先用起来，并持续产生价值和数据。大模型拿到最真实的企业场景的后训练数据之后，才能建立起预训练和后训练之间的循环迭代。让用户不停使用，模型不同获得反馈，模型后训练优化迭代才能越用越聪明。否则 AI 没有被用起来，就没有真实数据，模型就没有办法变

强，Agent 就更难用起来。企业需要数据工程、模型工程和算力工程三种能力协同，打通垂直场景下的数据采集和模型训练的统一标准和迭代循环的飞轮，跨越 2B 的认知鸿沟。

（一）数据工程：数据飞轮把 DI 变成 DIKW-P/C/Env

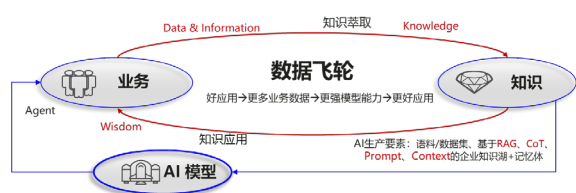
- 隐性经验显性化 Data&Informaton → Knowledge：对准业务难题 / 挑战，识别关键数据，把散落在作业记录，老专家脑子里的实战经验，转化成高质量的知识，

- 隐性专家经验变成显性化成数据、文档和 CoT 思维链数据等

- 显性知识工程化：关键要打破组织壁垒，通过统一的作业平台，加强业务和技术团队协同，在医疗行业叫“医（医生）/工（工程师）结合”，在金融领域叫“业（业务部门）/技（科技团队）融合开发”，共同对知识分层分级，分类标引，加强语义关联，形成高质量的知识体系

- 工程知识 AI 化 DIKW → Prompt/Context/Environment：知识（Knowledge）要最终变成作为大模型的生产要素，包括提示词工程、Centext 上下文、RAG 知识库、CoT 数据集、行业私有数据训练语料等，把知识转化成大模型智能（Wisdom）

- 完成数据飞轮闭环，需要统一的工具链流水线，拉通多个部门，打通多个业务环节



（二）模型工程：三阶九步，实现从“模型应用”到“模型定制”，再到“模型底座演进”技术跨度

企业可以按节奏分阶段构建模型工程能力，根据 AI 应用的场景复杂度、数据准备数据、人才梯队准备人，逐步驾驭行业大模型的工程方法。

初阶（2024~2025年）模型应用与交互优化：涵盖 Prompt Engineering（提示词工程）、RAG（检索增强生成）、CoT Prompting（思维链提示词）和 Agent Engineering（Agent 工程）。这些技术是与现有 LLM 高效交互的基础。

- 适用场景：只依赖通用互联网知识通用生成场景，例如营销助手活动方案生成；依赖企业实时 / 事实业务知识问答，例如金融信贷知识问答

- 算力门槛：企业布局 AI 推理算力，但需要从分散平台，向统一平台演进，多模型协同调度，持续积累能力范式，通过平台支持数据飞轮，不断积累私有数据，形成数据资产积累。

中阶（2026~2027年）模型定制与优化：SFT（监督式微调）、模型蒸馏和 RL（强化学习），这些技术旨在定制化模型以适应企业垂直领域的特定任务，提升大模型企业级的推理能力，

参考文献

- [1] 孙佳, 杨华峰, 刘威辰. 面向政企行业的智能体平台技术架构和应用研究 [J]. 电信科学, 2025, 41 (11): 47-56.
- [2] 陈旭. 基于大语言模型的自主智能体技术综述 [J]. 中文信息学报, 2024, 38 (09): 21-32.
- [3] 吕元杰, 王程宇. 智能体落地中的数据飞轮与强化学习工程化实践 [J]. 计算机工程与应用, 2025, 61 (15): 132-141.
- [4] Xi Z, Wang L. The Rise and Potential of Large Language Model Based Agents: A Survey [R]. arXiv 预印本, 2023 (09): 2309.07864.
- [5] 罗明远. 面向企业持续优化的 Agent 闭环数据飞轮架构研究 [J]. 小型微型计算机系统, 2026, 47 (02): 389-395.

否则 Agent 就只是一个“有手没脑”机器人。

- 适用场景：依赖业务知识，特定生成要求的行业通用任务助手，例如工单生成、数据分析助手；涉及行业复杂业务流及应用系统的数据分析与审核业务，例如运维数据查询等。

- 算力门槛：企业需要构建“训推一体”算力集群，通过灵活算力调度，实现“朝推夕训”，最大化算力使用效率。

高阶（2027年以后）模型底座演进与持续学习：没有完备的行业基础知识就没有强大的推理模型，通过增量预训练和增量后训练，从底层持续更新和优化大模型的核心能力，多模型权重智能融合和防遗忘算法系统，抑制大模型通用能力与领域能力“跷跷板”效应，以适应不断变化的世界知识和业务需求，此阶段对企业大模型技术能力要求最高。

- 适用场景：涉及金融复杂业务知识及逻辑的行业金融专业 & 核心业务，例如信贷助手、投研分析，进一步提升专业金融问答及金融专业文档生成能力。

- 算力门槛：相对于增量预训练，增量后训练所需算力较小，通常在数十到数百颗 GPU 即可完成，但对数据标注的质量和效率要求极高，需要构建高效的人工反馈循环系统



五、结语

随着 AI 技术深度嵌入企业核心业务系统，企业四层架构都将迎来系统性重构。业务架构跳出固化流程模板，依托智能体生成个性化处置方案，工作模式从单人独立作业转变为人机协同；原有应用软件分为两类，新开发产品原生适配 AI 运行逻辑，存量系统改造升级后充当智能体的信息采集与执行端口；数据架构层面，企业存量业务数据转化为模型训练素材，依靠数据飞轮持续赋能 AI 迭代升级；技术架构则以训推一体化强化学习为核心，搭建云边端协同的新型 AI 基础设施，助力企业转型为可持续迭代进化的学习型组织。