

FDE 与 Harness 助力 AI 商业化落地

计算机

姓名	电话	邮箱	登记编号
杨林(分析师)	021-23183969	yanglin2@gtht.com	S0880525040027
刘天雅(研究助理)	010-83939837	liutianya@gtht.com	S0880126010022
朱瑶(分析师)	021-23187261	zhuyao@gtht.com	S0880526010002

本报告导读:

FDE 负责深入业务场景并沉淀行业经验, Harness 负责保障智能体稳定执行与能力复用, 两者协同有望推动企业 AI 实现更高效、更可复制的规模化部署。

投资要点:

- 投资建议:** 我们认为, 企业 AI 由 Demo 走向 Production, 产业价值有望进一步从单纯模型能力向“业务理解+工程交付+Agent Runtime+平台复用”迁移。建议重点关注具备行业场景 Know-how、FDE 交付能力及 Agent 平台沉淀能力的厂商。推荐标的: 海康威视、金山办公、金蝶国际、汉朔科技、迈富时、赛意信息、合合信息、中控技术、卓易信息、税友股份、泛微网络、科远智慧、深信服、滴普科技、范式智能、鼎捷数智、聚水潭、新大陆、新国都、科大讯飞、软通动力、中软国际。
- FDE 正在重构企业 AI 从产品到生产的交付模式。** FDE (Forward Deployed Engineer, 前线部署工程师) 并非传统售前或实施岗位, 而是重新组合业务理解、场景识别、原型验证、系统集成、生产部署和效果评估等能力, 连接客户业务现场与产品研发。AI 进入企业核心流程后, 需要适配差异化的数据、系统、权限、规则和组织流程, 深度定制需求反而提升; 与此同时, AI Coding 和 Agent 工具又显著降低定制开发成本, 使过去经济性较弱的“重交付”模式重新具备规模化可能。FDE 真正的规模化不在于持续增加人员, 而在于将现场经验沉淀为 Skill、连接器、行业模板和平台能力, 使单位客户的人力投入持续下降。
- Harness 成为 Agent 能力演进的关键基础设施, 模型竞争由单点能力走向系统工程。** Harness 本质上是包裹在基础模型之外的 Agent Runtime, 负责上下文管理、记忆管理、工具调用、状态维护、权限控制、结果验证、异常恢复以及长程任务管理, 将一次模型调用组织成能够持续运行的 Agent。随着 Agent 从单轮问答走向多步骤、跨系统和长时间任务执行, 错误累积、状态漂移、工具误用和安全风险随之增加, Harness 的重要性反而持续提升。未来 Harness 竞争重点预计将由“集成多少组件”转向以合理 Runtime 开销提升真实任务结果, “Cost per Successful Task”有望成为比单 Token 价格更接近 Agent 生产力的核心指标。
- FDE×Harness 形成 AI 商业化飞轮, 推动企业 AI 由“人力复制”走向“能力复制”。** FDE 解决“业务理解与落地”, Harness 解决“稳定执行与能力沉淀”, 两者结合后能够将企业现场隐性的业务知识、流程规则和工程经验持续编码为 Agent 可执行能力。FDE 进入客户现场获取高价值业务知识, Harness 将其转化为 Skill、Tool、Workflow、Ontology 和运行规则; 生产环境进一步产生真实任务、失败案例和用户反馈, 再反哺 FDE 与 Harness 迭代。随着共性经验不断沉淀, 新客户部署由从零开发转向配置和复用, 最终形成部署速度提升、边际交付成本下降、收入增长快于交付人力增长的商业化闭环。
- 风险提示:** 企业 AI 商业化及规模化部署进展不及预期风险; 下游客户 AI 预算、场景拓展及 ROI 兑现不及预期风险; 行业竞争加剧及技术路线快速变化风险。

评级:

增持

相关报告

计算机《计算机周观点第 54 期: FDE 推动 AI 从 Demo 走向生产, L3/L4 自动驾驶迈入强制标准时代》2026.08.09

计算机《计算机周观点第 53 期: DeepSeek V4-Flash 公测, 海外云厂商持续加码 AI 基建》2026.08.02

计算机《CUDA: 英伟达护城河的构筑、松动与国产突围》2026.07.25

计算机《计算机周观点第 52 期: Kimi 算力需求激增, WAIC 超节点集中亮相》2026.07.25

计算机《计算机机构持仓 1.2%, 为近十年最低值》2026.07.24

目录

1.	FDE × Harness 重构企业 AI 落地路径.....	3
1.1.	AI 商业化瓶颈迁移：从模型能力转向企业规模化部署	3
1.2.	FDE 崛起：重构企业 AI 从产品到生产的交付模式.....	4
1.3.	Harness 重塑 Agent 执行体系：成为 Agent 能力演进的关键基础设施.....	8
1.4.	FDE × Harness 形成 AI 商业化飞轮：将企业现场经验持续沉淀为 Agent 能力.....	11
2.	海外厂商实践.....	11
2.1.	Palantir：以 Ontology 与 AI FDE 实现规模化，现场知识加速转化为机器能力	11
2.2.	OpenAI：FDE 打通模型到业务，Harness 沉淀 Agent 执行能力	14
2.3.	Anthropic：以合作伙伴扩展 FDE 交付，Harness 动态适配模型能力边界.....	15
2.4.	Microsoft：FDE 联合行业伙伴推进组织变革，Harness 产品化为标准 Agent Runtime	16
2.5.	AWS：FDE 走向 Agentic-first，双层 Harness 推动交付能力规模复制.....	17
3.	国内厂商实践.....	18
3.1.	腾讯：FDE 上移至场景定义，云端 Harness 推动经验机器化.....	18
3.2.	阿里：FDE 补齐前线交付，Agent 平台承接规模化复制	19
3.3.	软通动力：FDE 打通咨询与实施，AI Factory 重构传统项目制交付.....	20
3.4.	金蝶国际：Harness 内生 ERP 底座，强调业务知识沉淀与 AIOS.....	20
3.5.	赛意信息：FDE 深入工业现场，平台化沉淀推动项目经验规模复制.....	21
3.6.	科大讯飞：从专家经验到机器能力，FDE 推动产业 AI 规模化复制	22
3.7.	中软国际：BootCamp 前置孵化，FDE 专业化分工推动 AI 交付规模复制	22
3.8.	易鑫集团：Harness 强化安全治理，推动 Agent 规模化进入金融生产环境.....	23
4.	投资建议.....	25
5.	风险提示.....	26

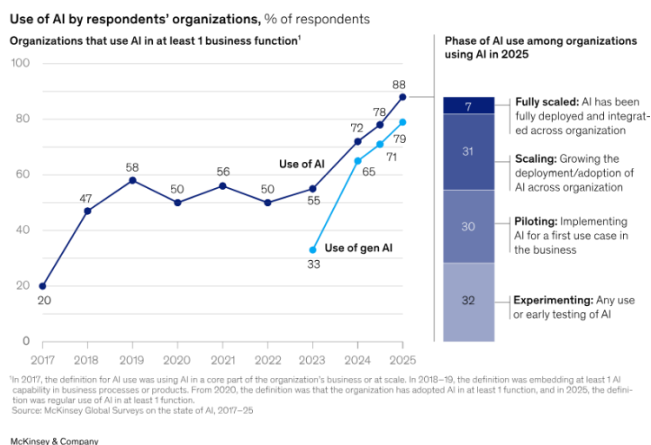
1. FDE×Harness 重构企业 AI 落地路径

1.1. AI 商业化瓶颈迁移：从模型能力转向企业规模化部署

AI 商业化的首要瓶颈已不再是“企业是否愿意使用 AI”，而是如何跨越从试点验证到规模化部署的鸿沟。随着大模型能力持续提升，AI 在企业端的渗透率已经较高，但应用深度与商业价值兑现明显滞后。McKinsey《The State of AI 2025》显示，接近九成受访企业已经在至少一个业务职能中常态化使用 AI，62%的企业已开始试验 AI Agent，但接近三分之二的企业仍未进入全企业规模化阶段，目前真正实现组织级 AI 规模部署的企业仅约三分之一。即使在收入超过 50 亿美元的大型企业中，进入规模化阶段的比例也不到一半。这意味着行业正在出现明显的“能力—采用”剪刀差：模型能够完成的任务快速增加，但企业把这些能力转化为稳定生产系统的速度相对较慢，AI 商业化的核心问题已经由“模型能不能做”逐步转向“如何让 AI 在企业内部真正跑起来并规模复制”。

图1：企业 AI 应用渗透率演变（2017 至 2025）

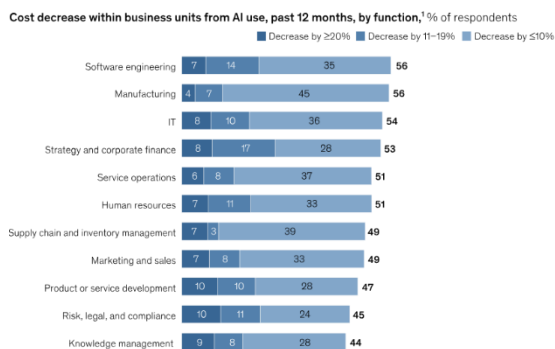
Reported use of AI in at least one business function continues to increase.



资料来源：McKinsey & Company

图3：AI 分职能成本节约效果

Respondents most commonly report cost benefits from AI activities in software engineering, manufacturing, and IT.

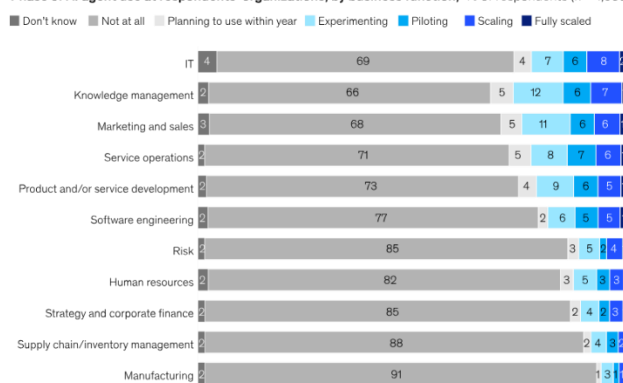


资料来源：McKinsey & Company

图2：各业务职能 AI Agent 落地情况

No more than 10 percent of respondents report scaling AI agents in any individual function.

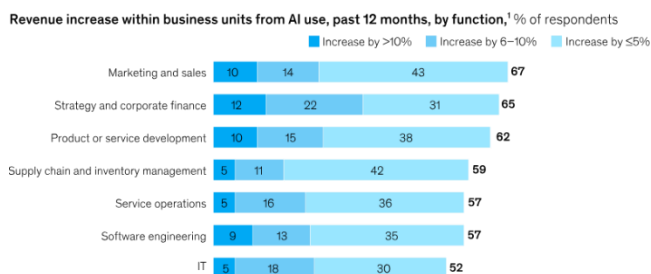
Phase of AI agent use at respondents' organizations, by business function,¹ % of respondents (n = 1,933)



资料来源：McKinsey & Company

图4：各业务职能 AI 营收增益效果

Respondents report the greatest revenue benefits from AI in marketing and sales, strategy and corporate finance, and product or service development.



资料来源：McKinsey & Company

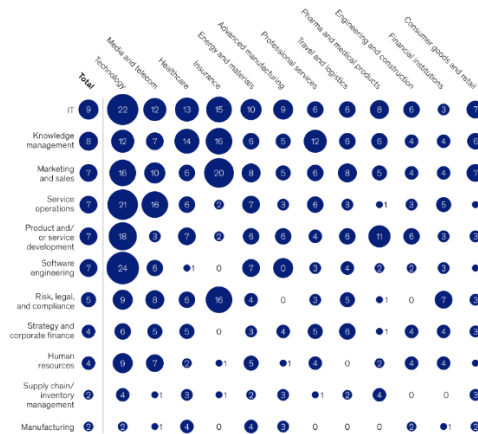
从试点走向规模化的关键制约，也正由单纯的技术能力转向业务流程、组织机制和数据治理等系统性问题。McKinsey 发现，在其分析的 25 项影响生成式 AI 价值兑现的因素中，重新设计业务工作流程是企业 EBIT 改善相关性最高的因素之一，但目前仅 21% 的生成式 AI 使用企业表示已经对至少部分工作流程进行了根本性重构。与此同时，能够有效嵌入业务流程、建立明

确的推广路线图、持续跟踪 KPI、建立用户反馈机制以及推动员工采用，均与更好的 AI 商业化效果显著相关。2025 年的进一步调查显示，AI 高绩效企业更加明显地体现出这一特征：其不仅更快规模化 AI，而且采用“重构业务”而非简单“增加 AI 工具”的思路，进行根本性工作流重构的概率接近普通企业的三倍。因此，AI 商业化并不是在原有软件和流程上简单增加一个模型入口，而是需要围绕 AI 重新组织数据、流程、权限、人机分工与绩效评价体系；企业组织自身的改造速度，正在成为限制 AI 价值释放的重要因素。

图5: 各行业业务线 AI Agent 成熟度

Use of AI agents is most often reported by respondents working in technology, media and telecommunications, and healthcare.

AI agent use that has reached the scaling phase,¹ by industry and business function, % of respondents



¹Includes respondents who answered "scaling" and "fully scaled". Question was asked only of respondents who reported regular use of AI in the respective functions and was released to reflect the total sample. In technology, n = 237; insurance, n = 90; healthcare, n = 250; media and telecommunications, n = 53; energy and utilities, n = 143; advanced manufacturing (includes advanced electronics, aerospace, automotive and assembly, and semiconductors), n = 188; professional services (includes legal services, management consulting, market research and product research), n = 226; consumer goods and retail, n = 76; travel, logistics, and transportation, n = 75; engineering, construction, and building materials, n = 77; banking and other financial institutions, n = 193; pharma and medical products, n = 76. Source: McKinsey Global Survey on the state of AI, 1,993 participants at all levels of the organization, June 26–July 29, 2025.

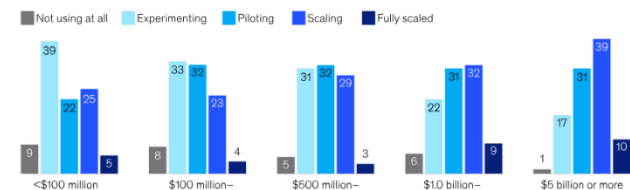
McKinsey & Company

资料来源：McKinsey & Company

图6: 企业规模与 AI 落地成熟度

Larger companies lead the way in scaling AI beyond pilots.

Phase of organization's use of AI, by company revenues,¹ % of respondents



Note: Figures may not sum to 100% because of rounding.
¹Respondents who said "don't know" are not shown, but represent <3% of the total.
 Source: McKinsey Global Survey on the state of AI, 1,993 participants at all levels of the organization, June 26–July 29, 2025.

McKinsey & Company

资料来源：McKinsey & Company

更深层的商业化瓶颈在于，AI 应用已经能够在单点场景产生效率和收入改善，但如何将局部效果转化为可持续、可衡量的企业级 ROI 仍未解决。McKinsey 数据显示，虽然 64% 的企业认为 AI 已经促进创新，许多企业也在软件工程、制造、IT 以及营销销售等具体场景看到成本下降或收入提升，但仅 39% 的受访者认为 AI 已经对企业整体 EBIT 产生影响，而且其中多数企业由 AI 贡献的 EBIT 比例仍低于 5%。随着 AI 进一步由 Copilot 走向能够调用工具、操作系统并持续执行任务的 Agent，生产环境中的可靠性、可控性和风险问题也更加突出：51% 的 AI 使用企业至少经历过一种负面后果，接近三分之一的受访者曾遭遇 AI 输出不准确带来的问题。我们认为，下一阶段 AI 商业化竞争的关键将逐渐从模型 Benchmark 转向真实业务结果，即能否以可接受的成本持续提高任务成功率、降低错误和人工介入，并最终形成可量化的收入增长、成本下降与投资回报；从“AI 可用”走向“AI 可靠、可控、可衡量”，将成为企业 AI 规模化落地需要跨越的下一道门槛。

1.2. FDE 崛起：重构企业 AI 从产品到生产的交付模式

1.2.1 FDE 连接业务与工程，成为企业 AI 规模化落地的关键角色

FDE (Forward Deployed Engineer, 前线部署工程师) 的核心并不是多了一种面容岗位，而是将过去分散在售前、解决方案、实施、产品和客户成功等环节的能力重新组合到客户一线。腾讯研究院将 FDE 定义为深入客户业务现场、将技术能力转化为可运行业务结果，并将一线经验持续反哺产品和平台能力的复合型角色。传统售前主要解决“客户为什么买”，实施团队解决“产品如何部署”，咨询公司更多解决“客户应该怎么做”，而 FDE 需要同时回答“什么问题值得用 AI 解决”“怎样把 AI 真正做出来”“上线后业务结果有没有改善”，并直接参与原型开发、系统集成、测试评估和产品迭代。其工作状态也并非“多写一点代码的售前”或“偶尔见客户的研发”，而是长期处于

客户业务与产品研发之间，在真实环境中完成问题定义和技术落地。

FDE 的核心产出应同时包含两类结果：对客户是业务结果，对厂商则是可复用的产品资产。FDE 与传统咨询、外包或驻场实施最关键的区分，在于其目标不仅是完成当期项目，更要让每一次项目成为产品学习过程。AWS 在介绍其 FDE 组织时明确将其与传统咨询区分：传统咨询往往进行评估、提出建议，并将单次部署作为独立项目，而 FDE 直接与客户业务、工程和安全团队共同建设生产系统，项目结束后既留下可运行的系统，也留下客户可以自行使用的工程能力、工作流与方法。腾讯研究院进一步给出了一个判断 FDE 的简洁标准：如果项目结束只留下一个客户定制系统，本质仍接近外包；如果带回了一些经验但无法复用，则仍是项目制；只有把现场经验进一步沉淀成 Skill、模板、测试集和产品能力，并显著降低下一个同类项目的交付成本，才构成真正可规模化的 FDE。

从能力结构看，FDE 实际上承担了“定义问题+解决问题”两种能力。腾讯研究院将其抽象为 Echo 和 Delta：Echo 负责理解客户的钱从哪里来、流程卡在哪里、AI 在哪个环节能够真正撬动价值，本质是判断“该做什么”；Delta 则负责快速搭建原型、接入系统、调试模型并在真实环境中迭代，解决“怎么做出来”。过去行业架构师可能懂业务但无法直接交付，工程团队能够开发但未必能够定义问题，产品团队又距离客户现场较远，FDE 的意义就在于重新连接这些被组织边界割裂的能力。这也决定了 FDE 最终比拼的并不是单纯编码能力，而是业务判断、工程能力和组织推进力的结合。

图 7: FDE 形成逻辑



资料来源：腾讯研究院，国泰海通证券研究

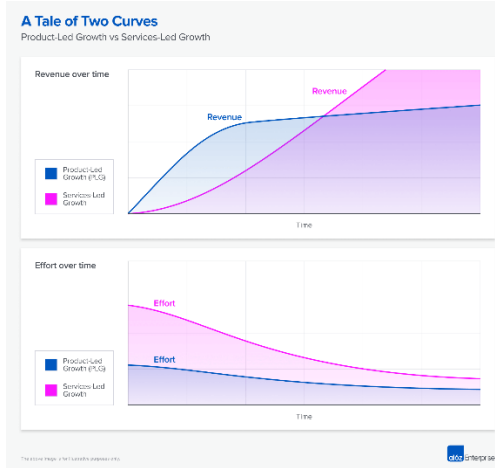
1.2.2 AI 重塑定制交付经济性，FDE 从特殊能力走向规模化模式

FDE 并不是 AI 时代才出现的新概念，真正变化的是 AI 改变了企业软件“标准化产品与定制化交付”之间的经济性。传统 SaaS 追求高度标准化，希望用统一产品服务尽可能多客户，以降低边际交付成本；但 Agent 真正进入企业后面对的是另一套逻辑——不同企业拥有不同的数据结构、CRM/ERP/MES 系统、审批规则、权限体系、行业术语及组织流程，客户购买的已经不是一个独立模型或 API，而是“一个能够把事情做完的业务流程”。因此，一个 Agent 从 Demo 走向生产通常需要完成数据连接、系统集成、权限治理、流程适配、异常处理、合规审计以及人工兜底等大量工作。

AI 本身又在降低过去“重交付模式”最昂贵的成本，使 FDE 从少数公司的特殊能力逐步变成更具普适性的商业模式。腾讯研究院将这一变化总结为三道成本门槛同时下降：行业知识蒸馏成本下降、定制开发成本下降、复合型人才供给成本下降。AI Coding 和 Agent 工具已经可以帮助工程师更快生成代码、接口、测试用例和文档，过去需要大量初级工程师完成的重复集成

工作开始自动化，因此 FDE 可以把更多时间放在需求澄清、架构判断、行业知识抽象、效果评测和组织推动等更高价值环节。AI 由此形成一个看似矛盾但非常重要的变化：企业 AI 需要更深的定制，而 AI 本身又让定制变得更便宜，过去经济上不成立的“高质量深度交付”开始重新具备规模化可能。这也解释了为什么 2026 年以来 OpenAI、AWS 等头部厂商开始主动扩大 FDE 组织，而不是单纯依赖传统 SI 完成最后一公里。OpenAI 成立独立 Deployment Company，明确提出基础模型只是工作的一部分，下一阶段企业 AI 的核心问题是如何将智能系统真正嵌入现实业务；其 FDE 将与业务负责人、运营人员和一线团队共同寻找高价值场景、重构 Workflow 并建设生产系统。AWS 则投入 10 亿美元建立 FDE 组织，将数千名工程师直接嵌入客户团队，目标是把 Agent 部署周期从“月”压缩至“天”。

图8：产品驱动增长 (PLG) 与服务驱动增长 (SLG) 的收入对比



资料来源：a16z Andreessen Horowitz

1.2.3 从 Demo 到生产，FDE 完整交付链条

FDE 首先解决的不是“如何最快做出 Demo”，而是“选择什么问题值得进入生产”。腾讯研究院提出，从 Demo 到生产可以划分为 L0—L4 五个阶段：L0 首先进行场景识别，通过业务诊断、价值排序和 AI 可行性评估判断问题是否值得解决；L1 进入原型验证，以真实或接近真实的数据快速搭建可运行 Demo，让客户直观看到效果并暴露模型边界；L2 则进入小范围试点，把真实数据、异常情况和用户反馈纳入系统，同时建立测试集；L3 才是正式生产部署，需要补齐权限、审计、监控、回滚、SLA 以及员工培训等完整生产条件；最后 L4 将已经验证的方案跨部门、跨场景乃至跨客户复制。因此，Demo 并不是 AI 项目的终点，而应被视为提前暴露风险和积累真实反馈的起点。FDE 真正重要的能力并不是“帮客户装好一个 AI 产品”，而是从一开始就以 Production 为目标反向设计 Demo，把价值判断、技术实现和组织落地放进同一个闭环。

FDE 交付链条最终应形成“场景筛选—Demo 验证—真实环境试点—生产部署—业务指标验证—规模复制”的完整闭环。从 Demo 进入 Production 最容易失败的环节，往往也并非模型本身。腾讯研究院总结，大量项目会在 L1 到 L2 阶段夭折，典型原因包括试点用户没有被纳入考核、使用 AI 只是“可选项”；Demo 阶段绕开的数据质量问题进入真实环境后集中暴露；以及组织审批、数据权限和合规流程没有提前打通。一个生产级 Agent 还可能同时涉及知识检索、业务数据库、安全审核、意图识别和多个模型，金融、医疗、政务等行业进一步要求日志、审计、可解释和人工回溯。因此，FDE 的专业性并不是问题出现以后不断“救火”，而是在项目早期就识别数据、权限、组织和工程风险，并把它们纳入产品设计。最终，生产化是否成功需要由业务指标而非模型指标验证。腾讯研究院建议从真实聊天记录、历史工单和业务数据建立测试集，通过自动评测和 A/B Test 持续验证效果，并将 AI 组与人

工组在触达率、响应率、转化率、人效等指标上进行逐环节比较。其核心原则是“卖成果而不是卖软件”——客户真正付费的原因不是企业采用了某个模型，而是业务指标发生了可验证改善。

图9: FDE 项目推进五阶段



资料来源：腾讯研究院，国泰海通证券研究

1.2.4 从单点落地到客户扩张，FDE 天然适配 Land and Expand 模式

FDE 天然对应 Land and Expand，而非传统软件项目“一次性签大单、一次性交付”的商业模式。 AI 产品在部署初期存在较大的效果不确定性，客户很难一开始就把核心业务全面交给 Agent，因此更合理的方式是先选择一个高价值、可量化的场景 Land，用较短周期完成可运行系统并证明 ROI。腾讯研究院指出，首个项目甚至可以短期不盈利，关键是让第一批业务用户真正使用起来、形成业务数据并建立内部信任；当客户看到真实效果后，再由单一场景扩展至多个场景、由单个部门扩展至多个部门，从而逐步提升合同金额和续费率。因此，首单的价值并不只体现在当期收入，而在于获得客户业务入口、真实数据和后续扩张权。这也是为什么 FDE 模式需要比传统软件销售更关注“进入什么场景”。如果首个项目只是一个非核心、低频、无法形成明确业务指标的 Demo，即使技术效果不错，项目上线后的使用率也可能迅速下降；而如果切入 CEO 或业务负责人真正重视的核心问题，并能够通过转化率、人效、响应时长、合规效率等指标建立清晰基线，AI 系统就更容易从 IT 试点变为真实业务基础设施。腾讯研究院特别指出，FDE 的 Expand 难度通常高于 Land：首个项目上线后，如果 AI 没有嵌入日常 Workflow，几个月后新鲜感消退，使用率就可能下降，因此在 Land 阶段就应提前设计后续扩展路径，并从 IT 团队逐步深入真正产生价值的业务部门。

从商业模式角度看，FDE 前期的服务投入可以理解换取更深客户关系和更大生命周期价值的获客投资。对复杂企业工作流而言，过度追求纯 PLG 和高初始毛利率未必是最优策略；深度实施虽然短期压低毛利率，却可以帮助 AI 公司承接更大的合同、验证客户真实付费意愿，并逐步进入企业核心 System of Work。Salesforce、ServiceNow、Workday 等企业软件的发展也表明，平台型软件在早期往往需要较重的实施与支持，之后才通过产品成熟和生态体系获得更好的规模经济。但 Land and Expand 并不等于无限制增加人力服务。按业务结果收费虽然是一个潜在方向，但前提是业务结果能够建立清晰基线，否则 FDE 容易退化为传统“按人天收费”。健康的商业路径应当是：Land 阶段允许相对较重的人力投入换取标杆客户和业务知识；Expand 阶段则更多依靠已经形成的软件功能、Skill、连接器和平台能力扩大使用，最终使客户收入扩张速度快于对应 FDE 投入的增长速度。

1.2.5 FDE 规模化并非扩人，而是将项目经验持续沉淀为软件资产

FDE 模式最大的争议在于其是否能够规模化，而判断规模化是否成立的关

键，并不是 FDE 人数能否快速增加，而是单位客户需要投入的人力是否持续下降。如果新签一个客户就必须重新派出同样规模的工程团队、重新访谈需求、重新连接系统和重新编写流程，那么本质仍然是线性增长的项目制公司。腾讯研究院因此提出，FDE 的核心是“双向蒸馏”：第一层把客户员工脑中的业务经验、流程规则和隐性知识转化为 AI 能够调用的知识库、工作流、SOP 和规则系统；第二层则把重点客户项目中获得的行业知识、系统接口、测试方法、失败案例进一步沉淀成厂商自己的产品能力、行业模板和平台组件。只有当一次项目产生的知识能够降低下一次项目的成本，服务投入才真正转化成软件资产。如果公司只复制“派工程师驻场”这一表层模式，却缺少能够吸收现场经验的统一软件平台，最终很容易演变成高端咨询或系统集成业务，而不是软件公司。

FDE 规模化可以沿着“Skill/连接器→行业模板→本体/语义层→平台”的路径逐步演进，而不必从第一天就建设一个庞大的完整本体系统。腾讯研究院认为，对多数国内企业而言，更务实的起点是先沉淀高频 Skill 以及 CRM、ERP、知识库、审批系统等标准连接器，再与行业知识库结合形成可重复部署的方案；随着客户数量增加，进一步把共性的对象关系、流程规则、权限边界和异常处理方式沉淀为行业本体，使新客户越来越多的需求只需要“配置+少量定制”即可完成。FDE 最终的规模化形态很可能不是“拥有最多 FDE 的公司”，而是“少量核心 FDE+平台+AI 工具+生态伙伴+客户自助”的多层杠杆体系。衡量 FDE 模式能否成立的最终标准不是服务收入规模，而是边际交付成本能否持续下降：同类客户上线越来越快、FDE 人天越来越少、可复用资产占比越来越高，并最终实现客户收入增长快于交付人员增长。

图 10: ERP 行业 FDE 规模化复制



资料来源：腾讯研究院，国泰海通证券研究

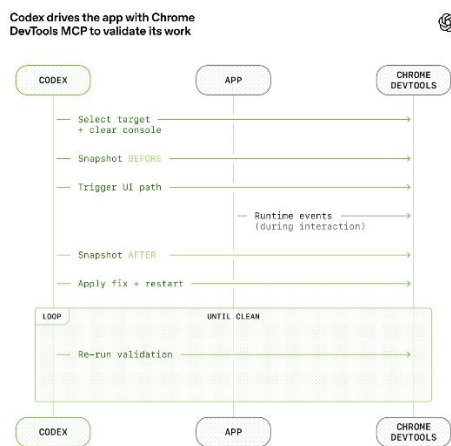
1.3. Harness 重塑 Agent 执行体系：成为 Agent 能力演进的关键基础设施

1.3.1 Harness 补齐 Agent 执行闭环，模型竞争走向系统工程

Harness 本质上是包裹在基础模型之外的一套 Agent 运行时系统，决定模型如何获取信息、调用工具、管理状态并持续完成任务。如果说大模型解决的是“理解和推理”，那么 Harness 解决的是“如何把一次模型调用组织成一个可以持续运行的 Agent”。Microsoft 将 Agent Harness 定义为将语言模型转化为 Agent 的“scaffolding”：单独的模型只能生成内容，而要实现工具调用、多步骤任务执行、记住此前状态并持续运行直至任务完成，需要在模型外围增加完整 Runtime。其最新 Agent Framework Harness 已经将 Function Invocation、History Persistence、Context Compaction、Todo 与 Plan/Execute 模式、File Memory、Skills、Tool Approval 及 Telemetry 等能力集成在统一运行时中。

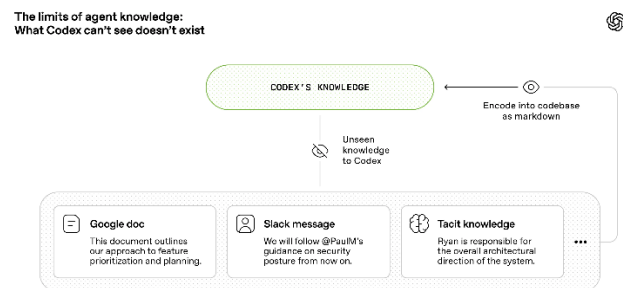
Agent 并不是“大模型+几个工具”的简单组合，而是“Model + Runtime”的系统工程。 Harness 处于模型与外部环境之间，一方面负责向模型提供正确的 Context、Memory 和任务状态，另一方面负责把模型生成的 Action 转换为真实的工具调用，并对执行结果进行验证、反馈和恢复。Harness 决定了模型“看到什么、能做什么、什么时候做，以及做错以后怎么办”。基础模型能力与最终 Agent 能力之间并不存在简单的一一对应关系。Harness-R1 论文明确指出，相同模型权重在不同 Harness 下可以表现出明显不同的 Agent 能力，因为 Context Construction、Memory、Skills、Tool Mediation、Action Validation 以及 Recovery 机制共同决定了模型与环境交互的方式。OpenAI 的 Codex 实践也体现出类似思路：其真正投入大量工程工作的并非持续修改模型，而是为 Agent 建立结构化知识库、执行计划、测试、CI、日志、Metrics 和反馈循环，使 Agent 能够获得自己执行任务所需的信息与工具。

图 11: Codex Chrome DevTools MCP 调试验证时序流程



资料来源：OpenAI

图 12: Codex 静态知识盲区：外部信息编码支撑运行时调试



资料来源：OpenAI

1.3.2 Agent 能力越强，Harness 的重要性反而越高

Agent 能力越强、能够承担的任务越复杂，对 Harness 的要求反而越高，核心原因在于任务链条拉长后，单次错误会逐步演变为系统性执行风险。简单 Chatbot 通常只需要完成一次“输入—生成—输出”，而 Agent 需要经历规划、检索、工具调用、环境反馈、再次决策等多个步骤，甚至持续数小时运行。任务执行时间越长、调用工具越多、涉及的系统权限越高，Context 丢失、工具误用、状态漂移、循环执行和错误累积的概率也越高。OpenAI 在 Codex 实践中已经观察到单个 Agent 可以连续执行六小时以上，而当代码生成吞吐快速提升后，人类 QA 能力反而成为新的瓶颈，因此其进一步将 UI、日志、Metrics 和 Trace 开放给 Agent，使 Agent 能够自行复现 Bug、验证修复并判断系统行为。

模型能力增强并不意味着 Harness 会消失，而是 Harness 需要随模型能力持续演化。Anthropic 在由 Opus 4.5 升级到 4.5 的实践中发现，更强模型已经可以原生承担部分过去必须由 Harness 完成的任务拆解和 Context 处理，因此可以删除部分 Sprint 结构、减少 Evaluator 调用；但在超出模型可靠能力边界的任务上，Planner 和 Evaluator 依然能够提供明显增益。Anthropic 因此提出一个重要原则：Harness 的每一个组件本质上对应着“模型当前无法可靠完成什么”的假设，随着模型进步，这些假设需要持续重新测试，而不是永久保留固定流程。模型本身逐渐承担基础规划、简单纠错和长 Context 处理后，Harness 的价值会进一步集中到跨系统执行、复杂工具治理、长程任务控制、自动验证、异常恢复和安全权限等更难标准化的环节。

1.3.3 Harness 的技术路线正在由“静态编排”走向“自进化运行时”

Harness 正在从“执行预先设计好的规则”转向“利用执行轨迹持续修改自身”。早期 Harness 更多表现为人为设计的静态 Workflow，即开发者预先规定 Prompt、任务步骤、工具选择和错误处理方式，本质上是利用固定规则对模型行为进行约束。这类方式在任务相对简单时有效，但随着 Agent 进入动态环境，其局限也开始显现：现实任务的失败模式往往难以提前穷举，同一种 Prompt 或 Reflection 机制并不一定适用于所有任务。Harness-R1 论文的实验显示，固定 Self-Refine 规则在三个 Benchmark 上均导致 Reward 下降，而 ReAct 等固定策略带来的改善也较为有限且不稳定；不同场景真正有效的干预位置也并不相同，例如 WebShop 更依赖 Pre-Action Mediation，而 ALFWorld 则更多依赖 Post-Feedback Recovery。这意味着未来 Harness 很难仅依靠人为提前写好一套固定流程，而需要根据 Agent 实际运行状态动态选择干预方式。

图 13: Runtime Harness 优化流水线

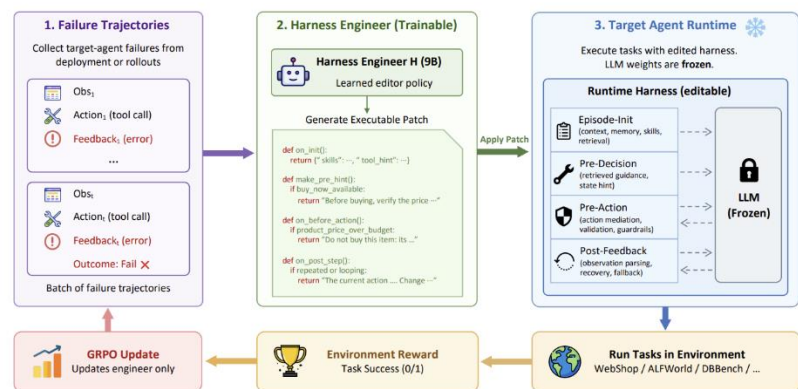


Figure 2 Overview of Harness-R1. Mined target failures become an evidence bundle (left); the harness engineer writes an executable patch that edits four lifecycle points in the runtime surrounding the frozen target (right); the patched target reruns the same tasks, and the resulting same-batch reward trains only the engineer via cold-start SFT and then GRPO (bottom loop). See the Method section for the loop, action space, and lifecycle hooks.

资料来源：Harness-R1: Learning to Edit Executable Runtime Harnesses from Agent Failure Trajectories

1.3.4 功能逐步标准化，Harness 竞争从“组件数量”走向“效果—成本”最优
Harness 优化的终点并非堆叠尽可能复杂的 Planner、Evaluator 和多 Agent 结构，而是在最少 Runtime 开销下获得最大的真实任务收益。随着 Planning、Memory、Tool Calling、Skills、Approval 和 Observability 等功能逐渐成为 Agent 框架的标准配置，未来 Harness 之间的竞争可能不再取决于“集成了多少功能”，而更多取决于这些组件能否以合理成本真正提升任务结果。Microsoft 目前已经能够将 Planning、Memory、Compaction、Tool Approval、Web Search 和 Telemetry 等组件直接作为 Harness 的默认能力提供，开发者只需要配置模型、Instructions 和 Tools 即可获得较完整的 Agent Runtime。当基础功能逐步标准化后，仅比较是否支持 Memory 或 Tools 的意义下降，评价重点自然会进一步转向最终结果——即 Agent 在真实任务中是否完成得更好、更快、更稳定。

Harness 问题从“架构设计”转化为一个经济学问题：为了提高一定幅度的任务成功率，究竟需要额外付出多少 Token、计算时间、Latency 和模型调用成本。Anthropic 的实验非常直观地展示了 Harness 的“效果—成本”权衡。在同样使用 Opus 4.5 完成应用开发的实验中，单 Agent 运行约 20 分钟、成本约 9 美元，而包含 Planner、Generator 和 Evaluator 的完整 Harness 运行约 6 小时、成本约 200 美元，成本超过 20 倍，但完整 Harness 生成的应用在完整度、可用性以及核心功能正确性上明显优于单 Agent 版本。Anthropic 后续并没有简单追求继续增加 Agent 数量，而是逐个删除 Harness 组件、验证性能是否下降，并指出 Evaluator 只有当任务处于模型“单独无法可靠完成”的能力边界之外时才值得支付额外成本。从商业化角度看，最终决定 Agent

价值的不是单次推理价格最低，也不是 Harness 组件最多，而是完成一个真实任务所需要的总成本以及最终成功概率；“Cost per Successful Task”有望成为比单 Token 价格更接近 Agent 实际生产力的核心指标。

图 14: Harness 运行模式资源开销

Harness	Duration	Cost
Solo	20 min	\$9
Full harness	6 hr	\$200

资料来源: Anthropic

图 15: Harness 各执行阶段时长与成本拆解

Agent & Phase	Duration	Cost
Planner	4.7 min	\$0.45
Build (Round 1)	2 hr 7 min	\$71.08
QA (Round 1)	8.8 min	\$3.24
Build (Round 2)	1 hr 2 min	\$36.89
QA (Round 2)	6.8 min	\$3.09
Build (Round 3)	10.9 min	\$5.88
QA (Round 3)	9.6 min	\$4.06
Total V2 Harness	3 hr 50 min	\$124.70

资料来源: Anthropic

1.4. FDE×Harness 形成 AI 商业化飞轮：将企业现场经验持续沉淀为 Agent 能力

FDE 与 Harness 分别解决 AI 商业化链条中的“业务理解”与“稳定执行”问题，两者结合后，企业现场经验可以进一步转化为机器可执行能力。FDE 负责回答“客户真实业务规则是什么”，Harness 负责回答“如何让 Agent 在执行过程中始终遵循这些规则”，两者共同完成从人的业务经验向机器执行能力的转换。真实生产环境本身能够持续产生新的数据和失败轨迹，使“现场经验—Harness 优化—业务结果验证”形成持续迭代的学习闭环。企业 Agent 上线以后，会不断产生工具调用错误、状态丢失、流程异常、边界案例和人工接管等真实反馈，这些反馈并非单纯的项目运维信息，而可以成为下一轮系统优化的数据资产。FDE 恰好处于连接这些机器反馈与真实业务结果的位置，可以判断某次失败究竟来自模型能力、数据质量、权限设置、业务规则缺失还是流程设计错误，并将新的行业知识重新转化成测试集、Skill、连接器及规则。Agent 每一次上线都不仅产生业务价值，也产生新的训练和工程反馈，使 FDE 与 Harness 形成“部署—运行—发现失败—抽象经验—修改系统—重新验证”的持续学习循环。

FDE 与 Harness 结合的最终价值在于把原本依赖高强度人工交付的 AI 项目，逐步转化为可复制、边际成本下降的平台化商业模式。早期进入一个新行业时，FDE 需要投入较多人力完成业务访谈、系统适配和规则梳理，Harness 也需要针对客户环境建立大量专属 Context、工具和控制逻辑；但随着同类客户数量增加，共性的业务对象、流程规则、异常处理方式和系统接口可以不断沉淀为行业模板、Skill、Connector 以及标准 Harness 组件，新项目便不再需要从零开发。AWS 进一步将这一思路明确为可复制的“Delivery Harness”：一次 FDE 项目中形成的领域 Ontology、评测框架、MCP Server 和 Agent Operations 能力，应能够被后续客户和合作伙伴重复使用。因此，我们认为 FDE×Harness 构成了一套 AI 商业化的正向飞轮：FDE 进入客户现场获取高价值业务知识→Harness 将知识编码为 Agent 可执行能力→生产运行产生真实任务和失败数据→FDE 重新解释并抽象这些反馈→Harness 和平台持续迭代→下一客户部署更快、成本更低。最终竞争的核心不再是谁能投入更多工程师，而是谁能够最快把现场经验转化为机器能力，并通过持续复用实现收入增长快于交付人力增长。

2. 海外厂商实践

2.1. Palantir: 以 Ontology 与 AI FDE 实现规模化，现场知识加速转化为机器能力

Palantir 的 FDE 模式并非单纯依靠工程师深入客户现场，而是建立在一套能够持续吸收前线经验的平台底座之上，这是其区别于传统咨询和项目制交付的核心。FDE 首先进入客户真实业务环境，理解其数据结构、工作流程、业务对象以及决策约束，再将具体问题抽象并映射至 Ontology。根据 Palantir 官方定义，Ontology 并非传统意义上的数据模型，而是企业的一层“Operational Layer”，其核心不仅描述企业中有哪些对象和关系，还进一步描述能够执行哪些 Action、调用哪些 Function 以及受到何种权限和治理规则约束；Palantir 最新架构文件进一步将 Ontology 描述为其体系的核心，用 Data、Logic、Action 和 Security 四个维度对企业决策过程进行统一建模。FDE 并不是每进入一家客户就重新开发一套独立系统，而是不断将现场验证有效的业务对象、逻辑和流程写回平台，使同类问题逐渐由“客户定制”转变为“平台标准能力”。因此，FDE 负责把现实世界中模糊、隐性的业务知识抽象出来，Ontology 则负责把这些知识转变为机器可以理解和操作的业务语义，这是 Palantir 重交付能够最终走向产品化的第一步。

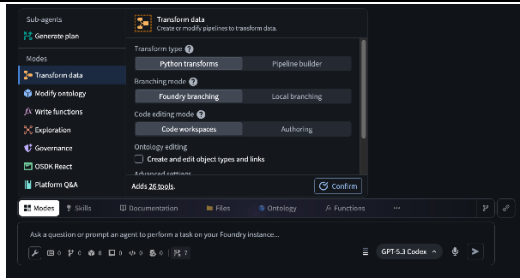
图 16: Palantir 三大产品与 FDE 的关系



资料来源：腾讯研究院，国泰海通证券研究

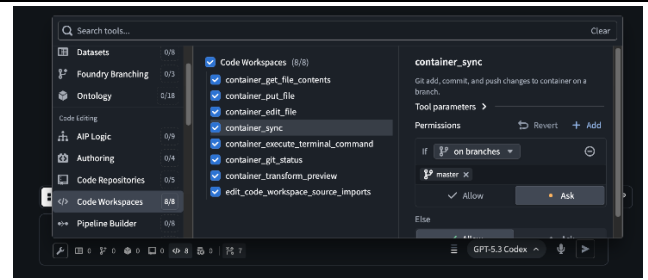
Palantir 正在进一步把过去由人类 FDE 完成的部分工程工作直接 Agent 化，AI FDE 实际上已经呈现出较完整的 Harness 特征。Palantir 官方将 AI FDE 定义为一个能够通过自然语言直接操作 Foundry 的交互式 Agent，其可以理解用户意图和上下文，并调用 Foundry 原生工具完成数据管道构建与修改、数据连接、Ontology 编辑、Function 开发、权限治理乃至前端应用建设；在执行过程中，还能够运行 Function Preview 验证函数行为、检查 CI 结果验证代码，并默认通过 Global Branch 或 Code Repository PR 提交修改供人工 Review。从本文 Harness 框架看，这意味着 Palantir 已经不仅提供“模型+工具”，而是在模型外围建立了一层完整的运行控制系统：AI FDE 的 Modes 决定当前任务类型并动态加载相应文档和工具，Capabilities 则控制 Agent 可以调用的具体能力，同时 Agent 能够主动生成执行计划、管理 Context、动态启停 Capabilities，并在任务变化时切换 Mode。Palantir 虽然并未在官网将这一整套体系统一命名为 Harness，但其本质已经与 Harness 高度一致——通过 Context Engineering、Tool Selection、Planning、Permission、Validation 和 Execution Control，把基础模型转化为能够安全操作企业系统的生产级 Agent。

图 17: Palantir AI FDE 模式与能力配置界面



资料来源: Palantir

图 18: Palantir AI-FDE Tool configuration 工具配置界面



资料来源: Palantir

AIP 进一步把这套运行机制从单个 AI FDE 扩展为整个企业 Agent 平台，使 Palantir 形成“FDE 现场学习—Ontology 知识抽象—AIP 运行执行—真实反馈—平台再迭代”的闭环。Palantir 官方将 AIP 定位为连接 AI 与企业 Data 和 Operations 的平台，可以在 Ontology 之上构建生产级 Workflow、Agent 和 Function；其最新 AIP 架构进一步将能力划分为安全模型接入、End-to-end Observability、Context Engineering、Ontology、Vector/Compute/Tool Services 等多个模块。尤其重要的是，Palantir 对 Agent 运行过程提供端到端可观测能力，包括监控进入 Ontology 的数据流、记录人类及 AI Agent 采取的 Action、追踪 Workflow 中的链式执行，同时监控 Token 消耗等资源指标。Ontology 又可以把 Workflow 中的反馈重新纳入 Continuous Learning Loop，使企业逐步由 AI 辅助决策走向流程自动化。因此，FDE 不是把知识“交给模型”之后就结束，而是通过 AIP 让这些知识进入真实生产运行，再通过 Agent 实际执行产生的新数据、异常和用户反馈重新更新 Ontology、Workflow 和 Evaluation 体系，从而形成持续学习。

图 19: Ontology 三层结构



资料来源: 腾讯研究院，国泰海通证券研究

Palantir 将“人的经验复利”和“机器的能力复利”结合起来。早期进入一个客户时，人类 FDE 承担大量问题定义、现场理解和业务抽象工作；随着共性知识进入 Ontology 和平台，后续 FDE 可以更多复用既有数据对象、Action、Function、权限结构和 Workflow；而 AIFDE 又进一步自动化数据转换、代码编写、Ontology 编辑和配置等原本需要人工完成的工程工作。AI FDE 的出现使人类 FDE 可以进一步聚焦于更高壁垒的客户理解、问题抽象和复杂业务落地。Palantir 官网进一步证明，这种自动化已经不只是“AI 帮工程师写代码”，而是在向能够自主调用 Foundry 工具、修改 Ontology、运行验证并接受权限治理的 Agent 演进。Palantir 正在形成更完整的 FDE×Harness 商业化飞轮：人类 FDE 发现和抽象业务知识→Ontology 将知识结构化→AIP/AI FDE 将知识转化为 Agent 可执行能力→生产运行产生真实

Feedback 与 Evaluation→有效经验继续写回 Ontology 和平台→下一客户所需的人工投入进一步下降。其真正的竞争壁垒不是 FDE 人数，而是将每一次现场经验转变为下一次机器能力的速度。

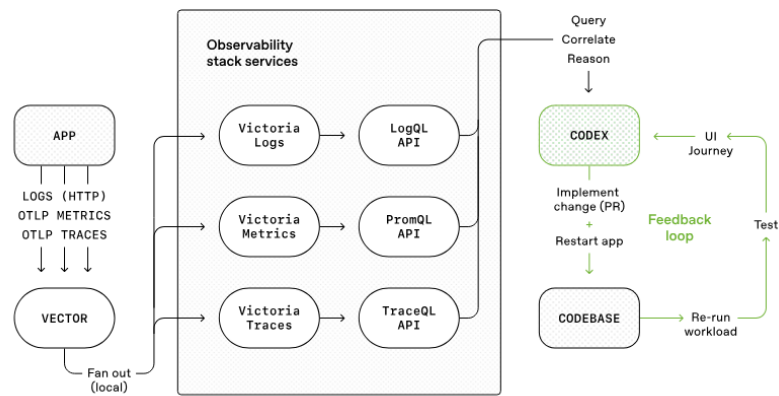
2.2. OpenAI: FDE 打通模型到业务, Harness 沉淀 Agent 执行能力

OpenAI 正在通过 FDE 把企业 AI 商业化重心从“提供模型能力”进一步延伸至“重构业务流程并交付可衡量结果”。2026 年 5 月，OpenAI 成立独立的 Deployment Company，并计划收购企业 AI 咨询与工程公司 Tomoro，后者将为公司带来约 150 名具有企业部署经验的 FDE 和 Deployment Specialists；与此同时，OpenAI 联合 19 家投资机构、咨询公司和系统集成商建立合作体系，并投入超过 40 亿美元用于扩大企业 AI 部署能力。OpenAI 指出，目前已有超过 100 万家企业使用其产品和 API，但随着模型能力增强，下一阶段企业 AI 的核心已经不只是模型本身，而是企业能否将智能能力真正嵌入关键业务流程。典型 FDE 项目首先从企业经营诊断开始，与管理层和业务团队筛选少数高价值 Workflow，随后 FDE 深入客户组织内部完成系统的设计、开发、测试与部署，将 OpenAI 模型连接至企业的数据、工具、控制体系和业务流程，并以日常业务中的可靠运行和可衡量结果作为最终目标。因此，OpenAI 的 FDE 实践本质上是在补齐“模型到企业生产系统”的中间环节：FDE 并非传统售前或单纯实施，而是与业务负责人、技术团队、运营及一线员工共同重新设计 Workflow，将通用模型能力转化为具体企业的生产能力；同时再通过咨询和系统集成伙伴将经过验证的部署模式向更多客户复制，从而形成“头部客户深度共创—方法论沉淀—生态规模复制”的商业化路径。

OpenAI 内部的 Codex 实践说明，FDE 发现和定义出的业务流程若要进一步由 Agent 稳定执行，还需要一套能够把人的经验持续编码进运行环境的 Harness。OpenAI 在“Harness Engineering”实践中，用 Codex 构建了一款内部真实使用的软件产品，应用代码、测试、CI、文档、Observability 和内部工具均由 Agent 生成；三名工程师在五个月左右驱动 Codex 合并约 1,500 个 PR，OpenAI 估算整体开发时间约为人工编码方式的十分之一。其核心经验并不是简单使用更强模型，而是“Humans steer, Agents execute”：人类工程师的工作从直接写代码转向设计 Environment、Scaffolding 和 Feedback Loop，把业务意图与工程经验逐步转化为 Agent 能够读取和执行的结构化文档、Skills、工具、架构约束、测试规则、日志、Metrics 和 Trace。当 Agent 失败时，团队也不只是重新 Prompt，而是追问“系统缺少了什么能力”，再把缺失的 Tool、Guardrail、Documentation 或 Validation 机制编码回运行环境；随着 Testing、Review、Feedback Handling 和 Recovery 逐渐进入 Harness，Codex 已经能够从单个 Prompt 出发，自主完成 Bug 复现、代码修改、测试验证、PR 提交、Build Failure 修复乃至 Merge，单次任务可连续运行六小时以上。

图 20: Codex 可观测栈与反馈闭环

Giving Codex a full observability stack



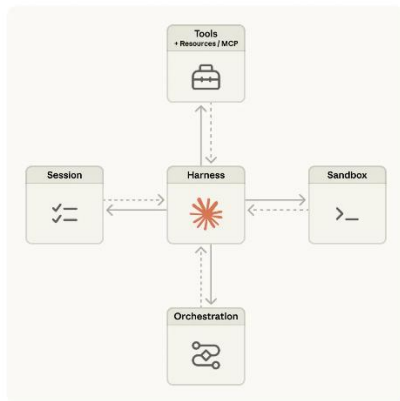
资料来源: OpenAI

2.3. Anthropic: 以合作伙伴扩展 FDE 交付, Harness 动态适配模型能力边界

Anthropic 在 FDE 侧的实践,正在从“自建前线部署团队”进一步走向“FDE+合作伙伴”的规模化交付体系,其核心目标是把 Claude 从模型/API 嵌入企业真实生产流程。Anthropic 已经设置 Forward Deployed Engineer, Applied AI 岗位,要求 FDE 嵌入战略客户推动 Claude 生产部署,其典型交付物并非传统软件安装,而是直接面向 Agent 系统的 MCP Server、Sub-Agent 和 Agent Skill,同时要求 FDE 识别能够跨客户复用的部署模式;这意味着 Anthropic 的 FDE 天然兼具“企业交付”和“Agent 基础设施建设”两种属性。Anthropic 正在通过生态体系进一步放大这一模式。2026 年 6 月,Anthropic 与 DXC Technology 宣布多年合作,DXC 计划培训数万名 Claude 认证的 FDE,将工程师直接嵌入银行、航空、保险、制造和政府等客户组织,把 Claude 接入这些企业已经运行多年的核心系统;在对外推广前,DXC 还先在自身约 11.5 万名员工的组织中使用 Claude,并与 Claude 协同完成其 AI 原生托管服务编排平台 DXC OASIS 超过 95% 的代码。因此,Anthropic 的 FDE 路径已经不只是“模型厂商亲自服务几个大客户”,而是在形成“Anthropic 提供模型和 Agent 能力—核心 FDE 攻坚复杂客户—大型 IT 服务伙伴复制交付”的多层体系;与此同时,FDE 交付 MCP Server、Agent Skill 等可复用资产,也意味着一次客户现场实践有机会进一步沉淀成后续 Agent 可以直接调用的机器能力,从而降低新客户部署成本。

在 Harness 侧,Anthropic 则是目前头部模型厂商中对“模型之外的 Agent 运行时”探索最系统的公司之一,其核心思路是:模型能力决定 Agent 能力上限,而 Harness 负责把这一能力转化为长期、连续、可验证的任务执行。Anthropic 在 2025 年底的长程 Agent 实验中发现,即使 Claude Opus 4.5 已经具有较强 Coding 能力,直接让模型跨多个 Context Window 持续构建复杂应用仍会出现过早结束、状态丢失和任务衔接困难,因此构建了 Initializer Agent+Coding Agent 的 Harness,由前者拆解任务、初始化环境,后者逐步执行,并通过结构化 Artifact 在多个 Session 之间传递 Context。Anthropic 随后在 Managed Agents 中进一步提出,Harness 本质上编码的是“当前模型还不能可靠完成什么”的假设,而这些假设会随模型升级而过时,因此 Agent Runtime 需要能够持续演进,而不能被固化为一套静态 Workflow。结合其 FDE 实践来看,Anthropic 实际上正在形成一条较清晰的 FDE×Harness 路径:FDE 进入客户现场识别业务流程并建设 MCP、Skills 和 Sub-Agent 等可执行资产,Harness 负责 Planning、Context、Tool Use、Evaluation 和长程任务控制,再由真实生产运行反馈不断调整 Agent 运行机制;由此,人的行业知识可以逐步从一次性交付经验转化为可重复执行的 Agent 能力。

图 21: Claude Managed Agents 组件虚拟化架构



资料来源: Anthropic

图 22: Agent 大脑-执行体解耦及接口规范

Component	Interface (pseudocode)	Satisfied by
Session	<pre>getSession(session_id) -> {session, Event[]} getEvents(session_id) -> PendingEvent[] // not yet processed emitEvent(id, event)</pre>	Any append-only log that can be consumed in order from any event point and accepts idempotent appends—Postgres, SQLite, in-memory array, etc.
Orchestration	<pre>wake(session_id) -> void</pre>	Any scheduler that can call a function with an ID and retry on failure—a cron job, a queue consumer, a while-loop, etc.
Harness	<pre>yield Effect<T> -> Effect<Result></pre>	Any loop that yields effects and appends progress to the Session.
Sandbox	<pre>provision({resources}) execute(name, input) -> string</pre>	Any executor that can be configured once and called many times as a tool—a local process, a remote container, etc.
Resources	<pre>{[source_ref, mount_path]}</pre>	Any durable store the container can fetch from by reference—filesystem, GCS, s3, git remote, S3.
Tools	<pre>{name, description, input_schema}</pre>	Any capability describable as a name and an input shape—MCP server, custom tool, etc.

资料来源: Anthropic

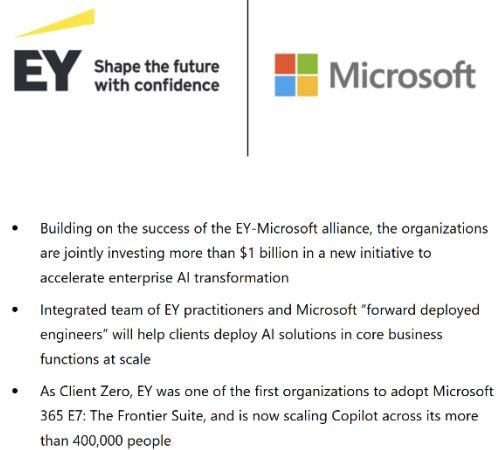
2.4. Microsoft: FDE 联合行业伙伴推进组织变革, Harness 产品化为标准 Agent Runtime

Microsoft 在 FDE 侧的实践,正在形成“微软工程能力+行业咨询能力+客户组织变革”的联合交付模式,核心目标是帮助企业从 AI 试点跨越到核心业务流程的规模化部署。2026 年 5 月, Microsoft 与 EY 宣布未来五年联合投入超过 10 亿美元推进企业 AI 转型,其中一个关键机制就是由 Microsoft Forward Deployed Engineers (FDE) 与 EY 行业专家组成联合团队,面向金融、工业、能源、消费、政府、医疗等行业,将 Agentic AI 直接部署到 Finance、Tax、Risk、HR、Supply Chain 等核心职能。Microsoft 将这一方法称为 AI-native Hypervelocity Engineering: FDE 提供模型、Agent 平台和工程实现能力, EY 则提供行业知识、流程设计和 Change Management, 双方共同围绕客户最高价值场景设计和开发行业 AI 解决方案,并通过 Shared Governance、统一商业模式和共同结果责任推动项目从 Pilot 进入生产。EY 同时承担“Client Zero”的角色,先在自身组织中大规模验证 Microsoft 技术: Copilot 早期覆盖 15 万名员工并带来约 15% 的生产率提升,后续向超过 40 万员工扩展;其基于 Power Platform 和 Copilot Studio 重构部分财务流程后, Lead Time 缩短 95%、运营成本下降超过 37%。因此, Microsoft 的 FDE 并非传统意义上的软件实施团队,而更接近连接产品研发、行业 Know-how 和客户组织的前线工程组织:先进入高价值业务流程完成深度共创,再借助 EY 等大型服务伙伴将经过验证的场景、Workflow 和方法论规模复制,这与前文“FDE 由单项目交付转向平台+生态杠杆”的发展方向高度一致。

在 Harness 侧, Microsoft 则已经把“如何让模型稳定做事”直接产品化为标准 Agent Runtime, 并且这一基础设施与 FDE 的前线交付逻辑天然互补。2026 年 7 月, Microsoft 正式发布 Agent Framework Harness, 并明确将 Harness 定义为“把语言模型变成真正 Agent 的 Scaffolding”: 模型本身只能生成内容,而要让其调用工具、持续处理多步骤任务、保留历史状态并运行直到任务完成,需要在模型外围建立完整 Runtime。Microsoft 的 Harness 已经将 Function Invocation、History Persistence、Context Compaction、Todo 与 Plan/Execute、File Memory、Skills、Web Search、Tool Approval 和 OpenTelemetry 等能力集成为默认组件,开发者只需要提供模型、Instructions 和 Tools,即可获得较完整的生产级 Agent 运行环境;其中 Approval 负责约束敏感操作, Compaction 解决长任务 Context 溢出, Memory 维持跨轮次状态, Skills 则把行业知识包装成可以按需发现和调用的可复用能力。Microsoft Agent Framework 本身也已在 2026 年将 AutoGen 与 Semantic Kernel 能力收敛到统一的开源 SDK 和 Runtime,并在 Harness 之外继续扩展 Hosted Agents、CodeAct 和多 Agent Workflow。如果将 FDE 与 Harness 放在同一商业化链

条中理解，Microsoft 实际上形成了较清晰的两层结构：FDE 进入客户现场识别高价值流程、获取行业规则并推动组织变革；Harness 则把这些规则进一步转化为 Instructions、Skills、Tools、Approval 和 Workflow，让 Agent 能够稳定重复执行。由此，前线一次性的项目经验有机会逐步被编码为标准 Runtime 能力，再通过 Microsoft 平台和合作伙伴复制到更多客户，这正是“现场经验→机器能力→规模化部署”的 FDE×Harness 闭环。

图 23: EY-Microsoft 全球 AI 合作



资料来源：Microsoft

2.5. AWS: FDE 走向 Agentic-first, 双层 Harness 推动交付能力规模复制

AWS 在 FDE 侧已经从“工程师深入客户现场”进一步发展为“Agentic-first 的前线部署体系”，目标是把企业 AI 从实验阶段直接推进到生产系统。2026 年 6 月，AWS 宣布投入 10 亿美元成立独立的 Forward Deployed Engineering 组织，计划将数千名 AI 工程师直接嵌入客户团队，与客户的业务、工程和安全部门共同建设运行在真实数据、治理规则和业务流程之上的 Agentic AI 系统。AWS 特别强调，其 FDE 与传统咨询不同：项目不是以评估、建议或 Billable Hours 为核心，而是围绕共同业务目标建设生产系统，并在项目结束时让客户具备独立运行和继续迭代的能力；客户不仅获得 Agent 系统，还会获得 Knowledge Graph、Runbook、架构文档、Workflow 以及经过培训的内部工程人员。更值得关注的是，AWS 提出 AI-Driven Development Lifecycle，即 FDE 本身也利用 Purpose-built Agents 完成开发，让 Agent 参与需求实现、工程执行和部署，人类工程师负责验证和指导，从而希望将传统以“月”为单位的 AI 项目压缩至“天”。AWS 还在客户自身账户内部构建 Semantic Layer，将企业分散数据与 Metadata 组织成受治理、可版本管理的 Knowledge Graph，使 Agent 可以基于企业业务语义进行推理，而不是依赖可能随人员流动而消失的隐性经验。因此，AWS 的 FDE 已经不是简单增加高水平工程人力，而是试图建立“人类 FDE 理解业务—Agent 提高工程执行效率—企业知识沉淀为语义层和机器资产”的新型交付模式，使每次项目都能为下一次部署积累可复用能力。

图 24: AWS FDE 三种交付模式



资料来源：AWS，国泰海通证券研究

AWS 更进一步的特点在于，它已经同时把 Harness 产品化和 FDE 交付方法 Harness 化，使“FDE×Harness”形成了较为明确的商业化闭环。在技术侧，AWS 于 2026 年 6 月正式推出 Amazon Bedrock AgentCore Harness，并明确提出“Agent 不只是一个模型”：如果模型是大脑，Harness 就是使其能够完成工作的“身体”，负责 Orchestration Loop、Tool Execution、Context Window 管理、跨轮次 State Persistence、Failure Recovery 和 Session Isolation；开发者只需配置 Model、Tools、Skills 和 Instructions，即可运行 Production-grade Agent，而且 Harness 与模型解耦，可以更换模型而无需重建 Agent 逻辑，并与 AgentCore 的 Identity、Memory、Observability 和 Gateway 安全治理体系打通。在交付侧，AWS 又将同样的思想引入 Partner-Led FDE，提出每一次 FDE 项目都应建立一套可复用的“Delivery Harness”，其中沉淀 Domain Ontology、Evaluation Framework、MCP Server、Agent Operations Tooling 以及记录架构选择、评测标准和行业模式的 Context Graph；随着项目推进，AWS 工程师逐步退出，合作伙伴依靠已经积累的 Delivery Harness 独立完成后续同类客户部署。这里实际上形成了两层 Harness：AgentCore Harness 解决单个 Agent“如何稳定运行”，Delivery Harness 解决一次 FDE 项目的经验“如何被下一次项目复用”。AWS 已经较清晰地展示了 FDE 与 Harness 结合后的规模化路径——FDE 进入现场获取业务知识和真实问题，Delivery Harness 将行业知识、评测与工程模式资产化，AgentCore Harness 则将这些知识转化为 Agent 可稳定执行的 Runtime 能力；随着项目数量增加，现场经验不断沉淀为机器能力，使部署速度提高、对 AWS 核心工程师的依赖下降，并最终实现从‘人力复制’向‘能力复制’迁移。

3. 国内厂商实践

3.1. 腾讯：FDE 上移至场景定义，云端 Harness 推动经验机器化

腾讯正在以“智能体开发平台+AI 开发工具+客户陪跑”构建更具本土云厂商特征的 FDE 体系，FDE 的角色也由技术搭建者逐步上移至场景发现、需求翻译和组织推动。腾讯云智能体开发平台 ADP 的演进经历了从 RAG 知识问答、Workflow 编排、Multi-Agent 协作到自然语言构建的过程，随着应用搭建门槛不断降低，企业 AI 落地的瓶颈反而由“能不能做出来”转向“到底应该做什么、如何拆成 AI 可执行任务”。因此，FDE 在项目中需要先理解业务问题并完成价值判断，再把业务需求拆解成 Agent 能够执行的流程，随后利用 ADP、CodeBuddy/WorkBuddy 等工具快速搭建原型，并在真实业务中围绕 Prompt、知识库、评测、Pipeline 持续调优，最终将有效经验沉淀为 Skill、连接器和行业模板。CodeBuddy/WorkBuddy 客户成功团队则进一步覆盖从管理层认知拉齐、业务流程梳理到工程实施和效果评测的完整 FDE 链条，部分头部客户已经在工具采购之外单独为 FDE 陪跑服务付费，显示企业付费对象正由单一 AI 工具进一步扩展到“帮助 AI 真正进入业务流程”的部署能力。从商业模式看，腾讯云正在尝试将 FDE 从少数专家的个人

能力转化为“平台+工具+伙伴”三层体系：原厂 FDE 负责标杆项目与前线学习，平台吸收共性经验，生态伙伴负责后续复制，从而在深度交付与规模化之间寻找平衡。

在 Harness 侧，腾讯云正在把模型之外的 Agent 运行能力进一步产品化，ADP 已明确引入“云端 Agent Harness”，使 FDE 在客户现场形成的业务知识能够被编码为机器可重复执行的 Runtime 能力。腾讯云最新 ADP 产品体系将云端 Agent Harness 作为核心能力之一，支持 7×24 小时中心化运行、云端 Sandbox 自主编码、企业级 Skills 调用和长时任务执行；在任务执行层，Agent 可以负责开放式意图理解和自主规划，Workflow 承接规则严谨、稳定可控的业务流程，两者可以双向调用，同时通过动态 Agentic RAG、连接器和 Skills 接入企业知识、业务系统与外部工具。生产治理方面，平台进一步加入内容审核、细粒度 RBAC 权限、操作审计和运行可观测能力，使 Agent 调用过程能够被约束、记录和持续优化。由此，腾讯的 FDE 与 Harness 正在形成较清晰的协同关系：FDE 从客户现场识别高价值场景，将隐性的业务经验、流程规则和系统约束抽象为 Prompt、Skill、Workflow、连接器和评测标准；Agent Harness 再负责将这些资产组织进 Context、工具调用、沙箱执行、权限治理和长程任务运行，使一次性的专家经验转化为可以稳定重复执行的机器能力。这一机制也与腾讯研究院提出的“双向蒸馏”一致：一方面将客户员工脑中的经验和流程转化为 AI 可调用的知识库、规则、工作流和 SOP，另一方面将重点项目中的行业知识、接口、测试方法和失败案例沉淀为平台组件和行业模板。因此，腾讯的潜在规模化路径并非持续增加 FDE 人力，而是形成“FDE 发现和抽象业务问题—Harness 将经验机器化—平台沉淀 Skill/Workflow/连接器—伙伴和客户重复调用”的商业化飞轮，使前线交付逐渐由“人力复制”转向“能力复制”。

图 25：腾讯云 ADP：云端 Agent Harness 企业智能体平台



资料来源：腾讯云

3.2. 阿里：FDE 补齐前线交付，Agent 平台承接规模化复制

阿里云正在从传统“模型/API+解决方案”模式向“前线部署+平台化 Agent 能力”延伸，FDE 逐步成为连接大模型能力与企业真实业务流程的重要组织形态。截至 2026 年 6—7 月，阿里云已开始以“前沿部署工程师 FDE”作为正式岗位招聘，反映其企业 AI 交付正在从传统解决方案架构、实施交付进一步向兼具业务理解、Agent 开发和生产部署能力的复合型角色演进。与

别负责明确 Agent 任务与执行顺序、维持跨会话 Memory、校验执行结果、限制模型操作边界以及管理任务运行全流程。基于这一思路，金蝶于 2026 年推出面向企业核心 ERP 业务的灵基 AIOS，在 OS 内核中集成完整 Harness 体系，并向上提供多 Agent 编排、Skill Hub、MCP 网关、记忆服务和 AI 效果评估等能力，使财务、供应链、人力等业务 Agent 能够直接连接企业数据与业务流程。其 Agent 引擎采用“主 Agent+子 Agent”双循环架构，由主 Agent 完成业务 Context 组装、模型调用和结果校验，复杂任务则动态调度子 Agent 协同执行；同时通过全链路可观测、Release Gate 质量门禁、离线评测、灰度发布以及 Trust Shield 安全体系，对 Agent 从开发到生产运行进行持续治理。金蝶此前的苍穹 Agent 平台也已具备自主规划、多工具调用、多 Agent 协同、日志追踪、评测分析、权限控制和 MCP 连接等能力，并将企业 30 余年管理经验沉淀为业务模板与工具。因此，金蝶的 Harness 实践本质上是把企业长期积累的业务知识、流程规则、权限体系和管理经验从模型外部固化为可持续复用的运行时资产，使底层模型可以持续替换升级，而企业自身的业务 Context、Skill、Workflow、评测和治理能力不断沉淀，进而形成比单一模型更具长期价值的企业 AI 基础设施。

图 27：金蝶灵基高管智能体工作台



资料来源：金蝶国际

3.5. 赛意信息：FDE 深入工业现场，平台化沉淀推动项目经验规模复制

赛意信息将 FDE 定位为全栈工业 AI 战略的核心交付体系，通过“嵌入业务现场、交付可衡量结果”推动 AI 从 Demo 走向生产可用。公司 FDE 团队全程进入客户业务一线，围绕真实经营痛点完成需求对齐、业务建模、快速验证、工程落地和能力沉淀，形成“客户价值提升+平台能力积累”的双向闭环：对外以降本增效、流程优化等业务指标验证项目价值，对内则将项目中形成的业务规则、智能体模板、开发技巧和接口能力持续沉淀为标准化产品能力。该体系目前已在十余家大型集团企业完成规模化验证，并依托 SDD 规范开发平台、自研善谋 GPT 平台和统一 DevOps 底座形成完整工程支撑，交付周期最高压缩 36%，人力投入平均降低 16% 以上。其核心差异在于不以一次性项目验收为终点，而是要求每次现场共创都能够反哺平台，使同类项目逐步由定制交付转向能力复用，从而形成交付效率和方案成熟度的持续提升。

赛意信息 FDE 形成“AI 辅助系统搭建+数字员工场景落地”两条主线，并以平台化方式承接现场经验。一方面，SDD 规范驱动开发通过标准化规格说明、项目知识库、统一开发规则和专属上下文管理，将自然语言需求转化为机器可无歧义执行的工程规范，解决企业 AI Coding 中返工率高、架构不一

致和难以生产上线的问题；另一方面，善谋 GPT 平台围绕智能体架构、MCP 外部系统接入、效果评估和 RAG 知识库工程化，持续沉淀财务审核、招投标风控、智能问答、物料编码校验等数字员工能力。例如，在围串标检测项目中，客户原本的一次性系统对接需求被进一步抽象为中台通用 MCP 标准接口，后续客户可直接复用，从而形成“业务场景迭代—平台能力升级—交付效率提升”的正向循环；三套自研平台再通过统一 DevOps 底座完成代码资产和智能体资产的一体化管理。因此，赛意信息 FDE 的本质不是单纯加强驻场交付，而是通过现场驱动工程，把项目中的业务知识和工程经验持续转化为可复制的平台资产，推动工业 AI 交付从“做一个项目”向“沉淀一套能力、复制一类场景”演进。

图28: 赛意信息 FDE 交付体系



资料来源：赛意信息

图29: FDE 双主线交付闭环



资料来源：赛意信息

3.6. 科大讯飞：从专家经验到机器能力，FDE 推动产业 AI 规模化复制

科大讯飞正在通过 FDE（前置部署工程师）模式，将大模型能力由单点技术验证进一步推进至复杂产业场景的生产落地。其核心做法是由同时理解技术与业务的团队深入客户一线，与业务专家共同定义问题、梳理规则、还原流程，并把现场形成的行业 Know-how 持续沉淀为平台能力。以国家能源集团智能招采项目为例，科大讯飞自 2022 年起围绕“编、报、评、核”等招采环节持续共创，将专家长期积累的业务知识、评审规则和判断逻辑转化为 AI 可理解、可执行的系统能力，并依托星火大模型、星辰 Agent 底座及招采智能体平台，围绕供应商、标的物、评标办法、技术评分要素、采购文件、合同和风险规则等核心对象建设 12 类标准化知识库，形成统一的业务语义体系。在真实运行中，专家对 AI 结果的复核、规则修正和异常案例处理又会持续反哺模型和平台，形成“业务即标注，标注即业务”的迭代机制；目前采购项目智能评审系统已累计稳定运行超过 21 万个项目，智能评审准确率达 97%，非招标采购询价通知单智能评审无人化率超过 89%，每年节约人工评审成本超亿元。因此，科大讯飞 FDE 模式的核心并非增加驻场人力，而是形成“深度共创—平台沉淀—价值验证—持续复用”的闭环，把一次项目中的业务理解、数据资产和工程经验沉淀为可复制的平台能力，并进一步向更多招采场景及能源、电力、基建等行业扩展。

3.7. 中软国际：BootCamp 前置孵化，FDE 专业化分工推动 AI 交付规模复制

中软国际正在以“AI BootCamp+FDE 专业化团队”构建企业 AI 从场景孵化到生产落地的全生命周期交付体系，推动传统项目实施向前线共创和能力沉淀转变。公司将 FDE 交付划分为项目负责人、Echo 和 Delta 三类角色：项目负责人负责整体统筹，Echo 深入客户业务一线完成需求挖掘、场景梳理和优先级定义，Delta 则承担 AI Coding、Agent 开发以及 Harness/Loop 工程，并进一步完成系统集成、生产部署、性能调优和持续迭代，从而形成“战

略—需求—技术”的完整闭环；前端再通过 AI BootCamp 在 5—7 天内完成从业务痛点识别、ROI 预评估到 MVP 开发和原型验证，降低企业 AI 项目早期试错成本。更重要的是，中软国际并不将 FDE 定位为单纯驻场实施，而是要求将前线项目中形成的业务语义、工程方法和可复用资产持续沉淀至 allmeta 企业智能操作系统：ECore 负责整合企业异构数据，Ontology 层将行业知识和业务规则转化为系统可理解、可执行的语义模型，Agentic 层则负责智能体编排与自主执行，使项目经验进一步转化为 MCP Server、子 Agent、Skill 及标准化部署模式。同时，公司基于 SmartCode 建立 FDE Delta 人才梯队，通过实操训练、诊断 Agent 和认证定级持续提升前线工程能力，并通过 Skill 和场景模板的“使用—贡献—优化—复用”机制，将个人最佳实践沉淀为组织资产。因此，中软国际 FDE 模式的核心并非依靠更多人力复制项目，而是通过“BootCamp 快速验证—FDE 现场共创—Ontology/Agent 平台沉淀—标准资产复用”形成交付飞轮，使一次性客户经验逐步转化为可跨项目、跨行业复用的企业 AI 能力。

图30: 中软国际 AI BootCamp 与 FDE 交付体系



资料来源：中软国际

3.8. 易鑫集团: Harness 强化安全治理, 推动 Agent 规模化进入金融生产环境

易鑫集团将 Harness 定位为金融垂直领域 Agent 规模化进入生产环境的核心治理基础设施，并由“聚焦模型”进一步转向“模型+Harness”的体系化 AI 路线。围绕汽车金融业务流程长、决策节点多、强合规和高风险的特点，易鑫在 2026 年形成自研 Harness 治理体系，与 Agentic 大模型共同构成企业级 Agent 运行的“双轮”：模型负责理解、推理和任务执行，Harness 则负责

对执行过程进行约束、监控和兜底。在实际业务中，Harness 支持 Agent 与真人在同一订单流中实时无缝切换，当模型出现幻觉、违规承诺等异常行为时，可在毫秒级触发熔断及治理策略并切换至人工链路；同时建立覆盖数据接入、模型调用和业务执行的全流程数据关联体系，实现 Agent 行为的可审计、可追溯和持续更新。在此基础上，易鑫已将 Agentic AI 嵌入智能呼叫、预审、风控、客服、资管和质检等汽车金融全业务流程，截至 2026 年 5 月底 AI 平台累计调用超过 1.25 亿次，智能体自主交付业务占比达到 65%，业务转化率提升超过 20%，整体运营效率提升 100% 以上。因此，易鑫的 Harness 实践并非简单增加模型外部的安全模块，而是将人机协同、执行边界、异常熔断、数据治理和运行审计统一纳入 Agent Runtime，使大模型从“能够完成任务”进一步升级为“能够在金融生产环境中安全、可控、持续地完成任务”。

4. 投资建议

我们认为，企业 AI 由 Demo 走向 Production，产业价值有望进一步从单纯模型能力向“业务理解+工程交付+Agent Runtime+平台复用”迁移。建议重点关注两类厂商：一是具备较强行业 Know-how 和 FDE 能力、能够深入客户核心业务流程并持续沉淀可复用资产的交付型厂商；二是拥有 Harness、Agent 平台、Ontology 或企业级 AIOS 等基础设施，能够将行业经验进一步固化为 Runtime 能力的产品型厂商。同时关注阿里云、腾讯云等平台型厂商在企业 AI 交付与 Agent 基础设施体系上的持续进展。推荐标的：海康威视、金山办公、金蝶国际、汉朔科技、迈富时、赛意信息、合合信息、中控技术、卓易信息、税友股份、泛微网络、科远智慧、深信服、滴普科技、范式智能、鼎捷数智、聚水潭、新大陆、新国都、科大讯飞、软通动力、中软国际。

表1: 相关公司估值表

股票代码	股票简称	收盘价 (元)	EPS (元)			PE			评级
			2025A	2026E	2027E	2025A	2026E	2027E	
002415.SZ	海康威视	36.47	1.55	1.89	2.16	23.59	19.27	16.89	增持
688111.SH	金山办公	260.48	3.97	5.54	5.67	65.61	47.03	45.98	增持
0268.HK	金蝶国际	8.21	0.03	0.11	0.20	313.64	73.71	42.05	增持
301275.SZ	汉朔科技	44.18	1.10	1.44	1.94	40.16	30.68	22.79	增持
2556.HK	迈富时	46.10	0.35	1.02	1.54	132.83	45.07	29.92	增持
300687.SZ	赛意信息	28.34	-0.26	0.29	0.44	-	97.15	64.16	增持
688615.SH	合合信息	120.98	3.24	2.94	3.72	37.34	41.15	32.56	增持
688777.SH	中控技术	99.19	0.56	0.95	1.30	177.13	104.83	76.41	增持
688258.SH	卓易信息	109.30	0.66	1.46	2.37	165.61	75.12	46.05	增持
603171.SH	税友股份	46.91	0.34	0.63	0.90	137.97	74.78	52.07	增持
603039.SH	泛微网络	45.76	1.14	1.20	1.43	40.14	38.09	32.09	增持
002380.SZ	科远智慧	29.43	0.58	1.72	2.16	50.74	17.08	13.60	增持
300454.SZ	深信服	121.50	0.93	1.29	1.71	130.65	93.87	71.17	增持
1384.HK	滴普科技	36.36	-2.86	0.03	0.21	-	1212.00	172.16	增持
6682.HK	范式智能	28.74	-0.05	0.20	0.66	-	145.67	43.84	增持
300378.SZ	鼎捷数智	36.60	0.61	0.79	0.96	60.00	46.62	38.32	增持
6687.HK	聚水潭	14.38	-3.79	0.80	1.22	-	17.90	11.77	增持
000997.SZ	新大陆	18.10	1.01	1.15	1.47	18.00	15.72	12.29	增持
300130.SZ	新国都	19.30	0.83	1.00	1.27	23.25	19.36	15.20	增持
002230.SZ	科大讯飞	42.59	0.36	0.51	0.66	118.31	83.99	64.33	增持
301236.SZ	软通动力	40.07	0.22	0.34	0.47	182.14	119.15	85.97	增持
0354.HK	中国软件 国际	3.52	0.12	0.21	0.27	29.87	16.55	13.04	增持

资料来源：iFind，国泰海通证券研究

注：收盘价单位为当地货币，为 2026/8/12 数据，EPS 来自 iFind 一致预期，其他数据来自国泰海通证券研究

5. 风险提示

企业 AI 商业化及规模化部署进展不及预期风险。企业 AI 从试点验证走向生产环境仍需解决数据治理、系统集成、权限管理、组织协同及效果评估等问题，若相关环节推进速度低于预期，可能影响 AI 应用渗透率提升及相关厂商收入兑现节奏。

下游客户 AI 预算、场景拓展及 ROI 兑现不及预期风险。企业客户 AI 投入仍受宏观环境、IT 预算及内部投资回报要求影响，若客户预算增长放缓、高价值场景拓展不足，或 AI 项目难以形成可量化的降本增效与收入提升，可能影响后续扩单、续费及商业化进程。

行业竞争加剧及技术路线快速变化风险。当前大模型、Agent 平台、Harness 及企业 AI 交付体系仍处于快速演进阶段，国内外厂商竞争持续加剧。若模型能力、Agent 架构、工具调用或企业 AI 基础设施出现较大技术迭代，可能对现有产品路线、研发投入及市场竞争格局产生影响。

本公司是本报告所述金山办公(688111),中控技术(688777)的做市券商。本报告系本公司分析师根据金山办公(688111),中控技术(688777)公开信息所做的独立判断。

本公司具有中国证监会核准的证券投资咨询业务资格

分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

免责声明

本报告仅供国泰海通证券股份有限公司（以下简称“本公司”）授权客户使用。本公司不会因接收人收到本报告而视其为本公司的当然客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌。过往表现不应作为日后的表现依据。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。本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此，投资者应注意，在法律许可的情况下，本公司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或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

市场有风险，投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“国泰海通证券研究”，且不得对本报告进行任何有悖原意的引用、删节和修改。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

评级说明

		评级	说明
投资建议的比较标准 投资评级分为股票评级和行业评级。 以报告发布后的 12 个月内的市场表现为比较标准，报告发布日后的 12 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数涨跌幅为基准。	股票投资评级	增持	相对沪深 300 指数涨幅 10%以上
		谨慎增持	相对沪深 300 指数涨幅介于 5%~10%之间
		中性	相对沪深 300 指数涨幅介于-5%~5%
		减持	相对沪深 300 指数下跌 5%以上
		无评级	对于个股未来市场表现与基准指数相比无明确观点
	行业投资评级	增持	明显强于沪深 300 指数
		中性	基本与沪深 300 指数持平
		减持	明显弱于沪深 300 指数

国泰海通证券研究所

地址 上海市黄浦区中山南路 888 号

邮编 200011

电话 (021) 38676666