

OpenAI FDE研究报告



- FDE 是 OpenAI 将前沿模型落到复杂企业生产环境的工程组织方式。
- 关键对象是部署组织、生产系统、企业流程与可衡量业务结果。
- 部署能力正在成为企业 AI 从实验走向价值的关键壁垒。

研究问题与分析框架

从组织、工程、治理和价值四个维度理解 FDE

问题一

OpenAI 为什么需要专门的 Deployment Company 与 FDE 组织?



问题二

FDE 如何把 AI 系统部署到真实企业环境并沉淀为产品能力?



问题三

企业和人才如何围绕 FDE 构建能力、治理和投资决策?



分析链条

组织动因



工作机制



案例证据

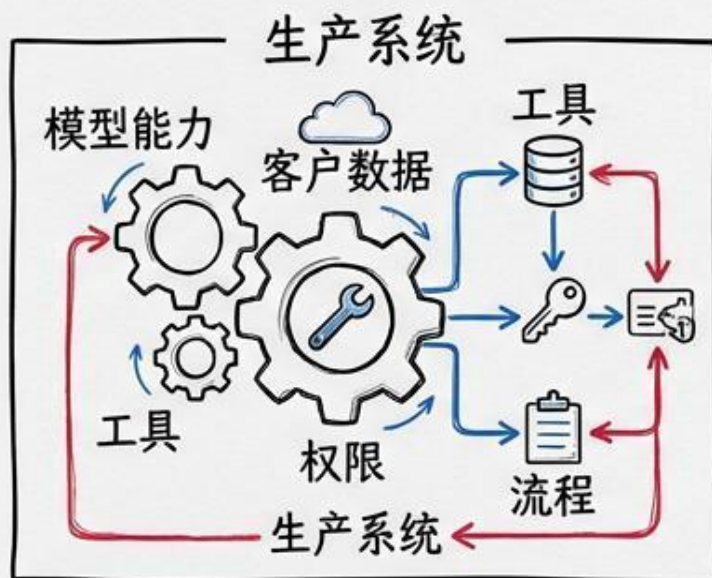


治理与人才启示



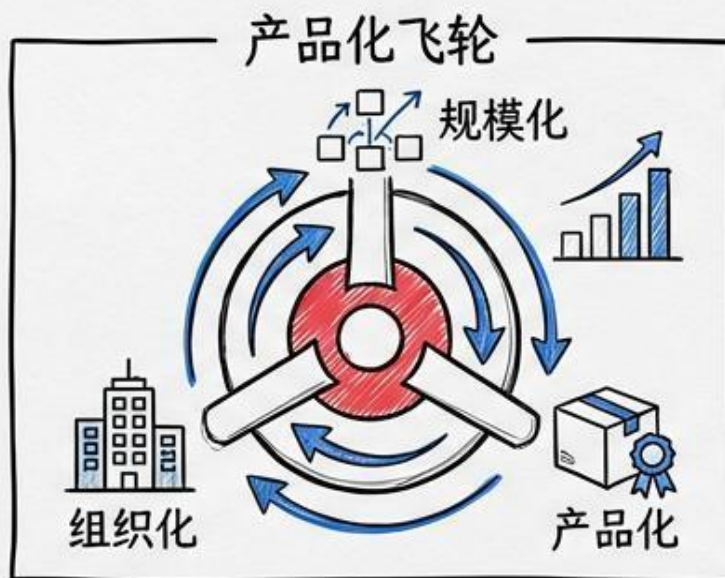
核心结论速览

FDE 是 OpenAI 的生产部署组织，不是单纯客户服务



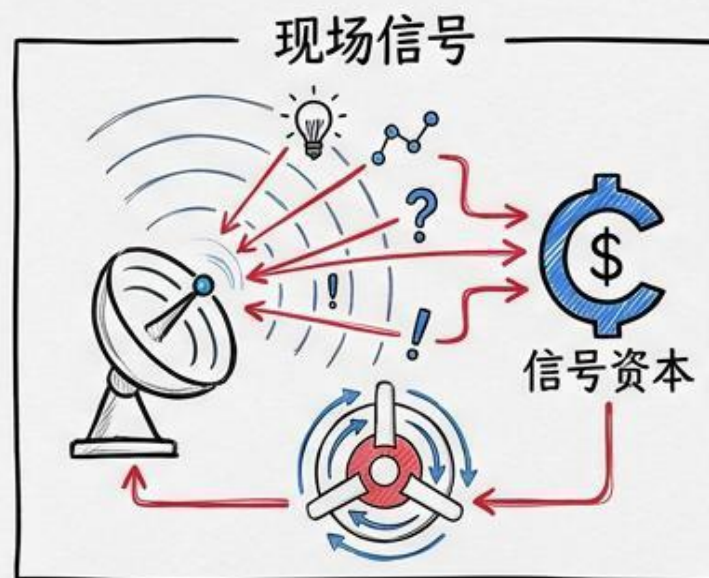
— FDE 把模型能力、客户数据、工具、权限和流程连接成生产系统。

Deployment Company 把现场部署组织化、规模化和产品化



— Deployment Company 把现场部署组织化、规模化和产品化。

FDE 的长期价值来自现场信号资本与部署飞轮



— FDE 的长期价值来自现场信号资本与部署飞轮。

@ 清新研究团队简介

- 领导学术研究团队近30人。指导大数据、AI、人形机器人等多个产业团队。
- 团队坚持：整体主义、实证主义、社会建构、进步主义。

沈阳：清华大学新闻学院/人工智能学院双聘教授、博导



视频号：@清新研究；公众号：@清新研究

六大研究方向：



1. AI大模型
理论与哲学



2. AI文艺



3. AI应用



4. 新媒体与
网络舆论



5. 大数据



6. XR应用

OpenAI Deployment Company 是什么

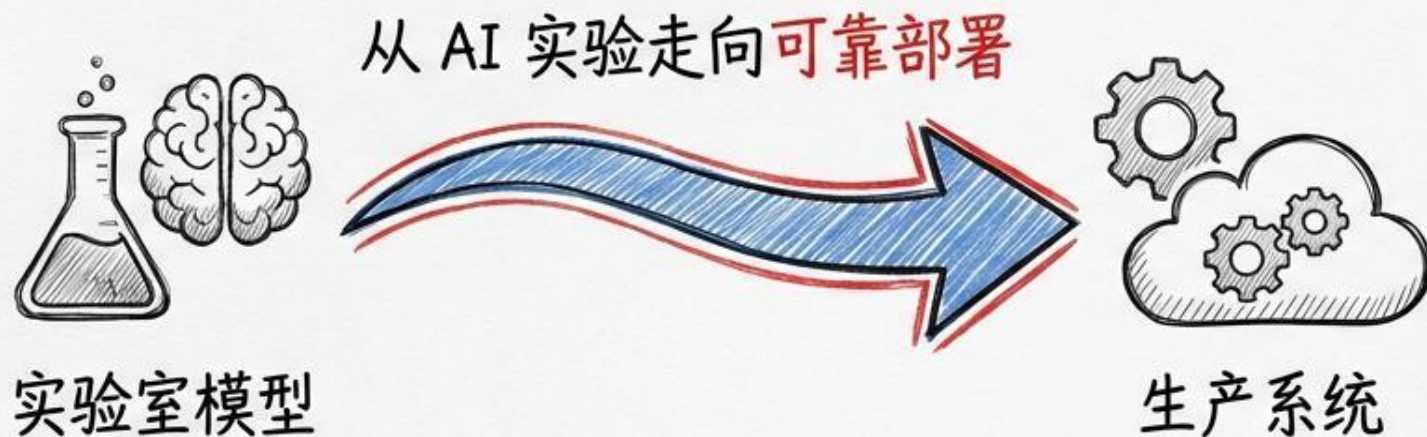
它是 OpenAI 帮助组织构建和部署 AI 系统的新组织形态



FDE 的官方定义

Forward Deployed Engineers

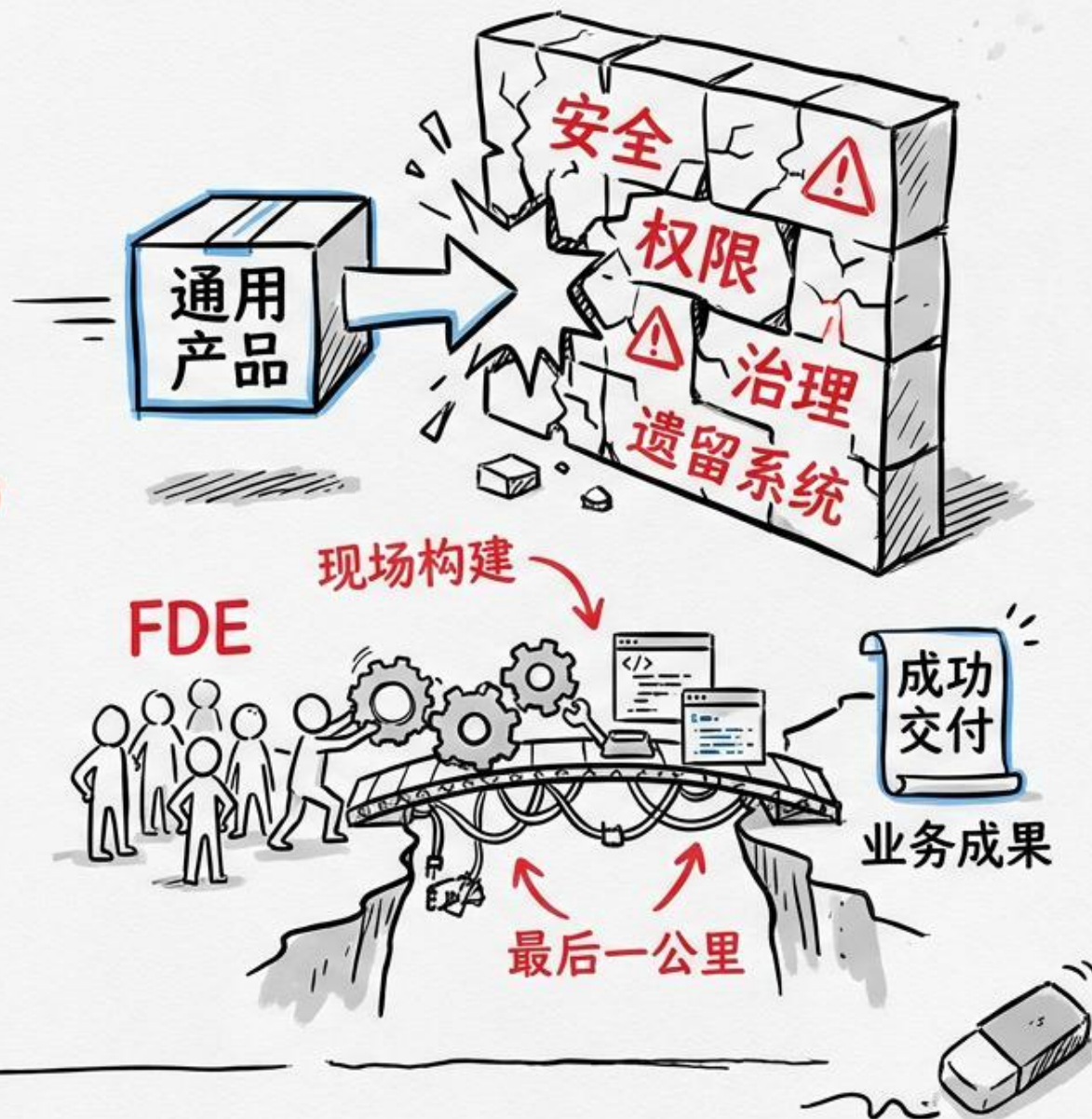
把 AI 带入复杂真实用例的生产环境



- OpenAI 明确写到 FDE 是把 AI 带入生产的方式。
- FDE 团队在真实企业复杂性中构建定制 AI 系统。
- 目标是从 AI 实验走向可靠部署。

为什么通用产品模式不够

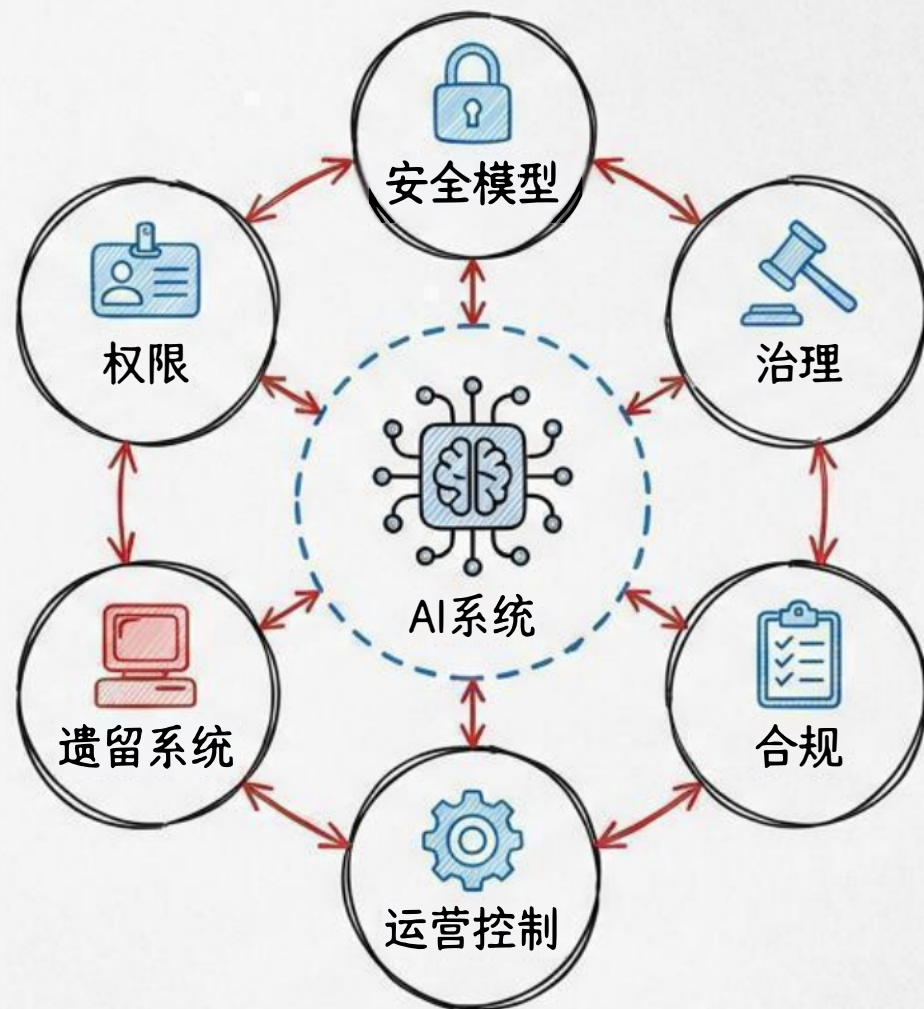
- 复杂企业环境里，安全、权限、治理和遗留系统都不是边缘问题。
- 通用产品往往无法直接吸收行业流程和组织责任。
- FDE通过现场构建解决最后一公里问题。



真实企业环境意味着什么

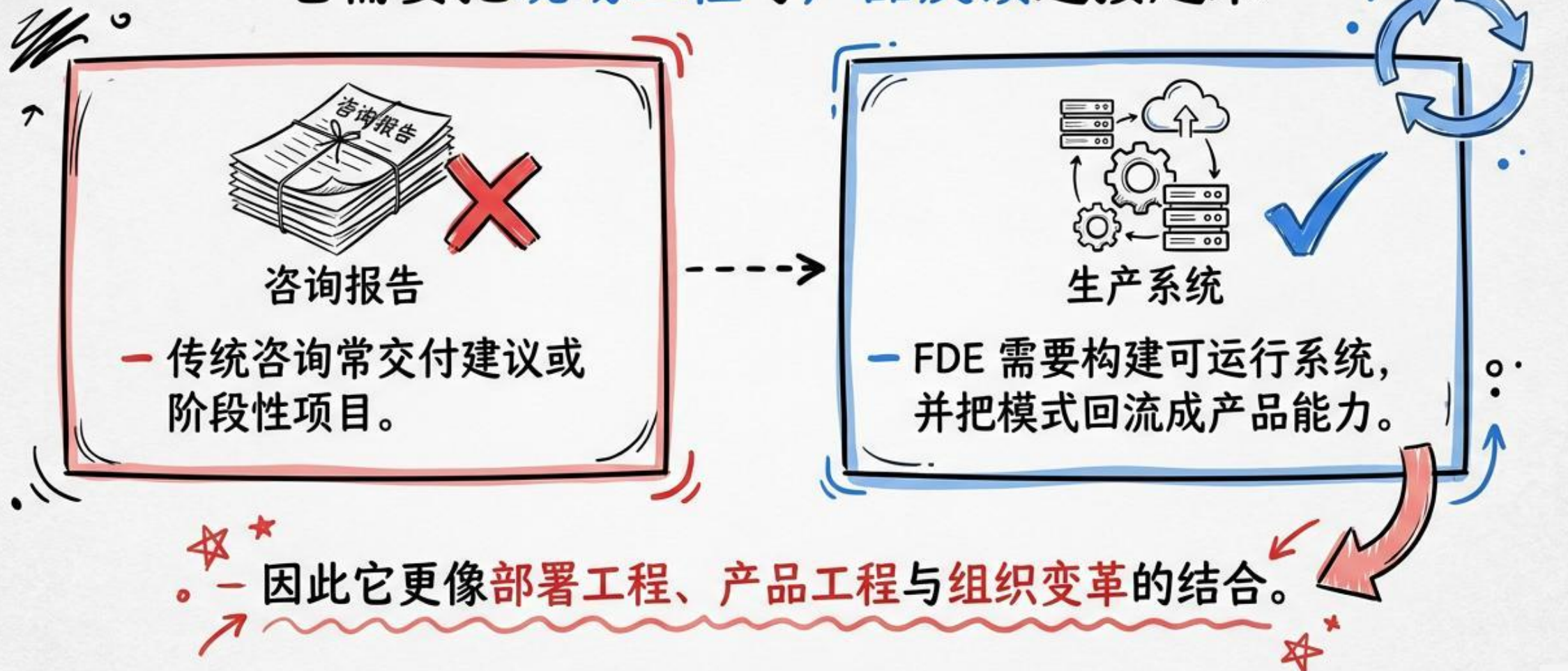
安全模型、权限、治理、合规、运营控制和遗留系统共同构成约束

- OpenAI 官方页面把这些约束列为红色核心而非边缘。
- 这要求 FDE 从架构第一天就处理企业边界。
- 部署难点不只是模型答案质量，而是系统能否可靠运行。



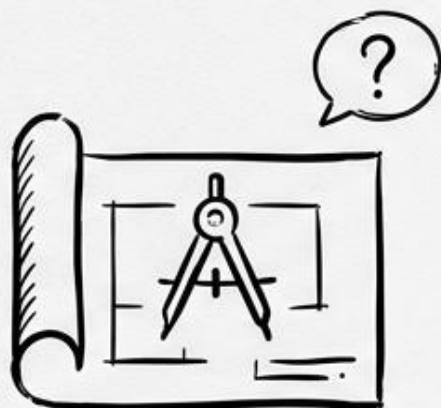
FDE 不是传统咨询外包

它需要把现场工程与产品反馈连接起来



FDE 不是单纯解决方案架构师

方案蓝图



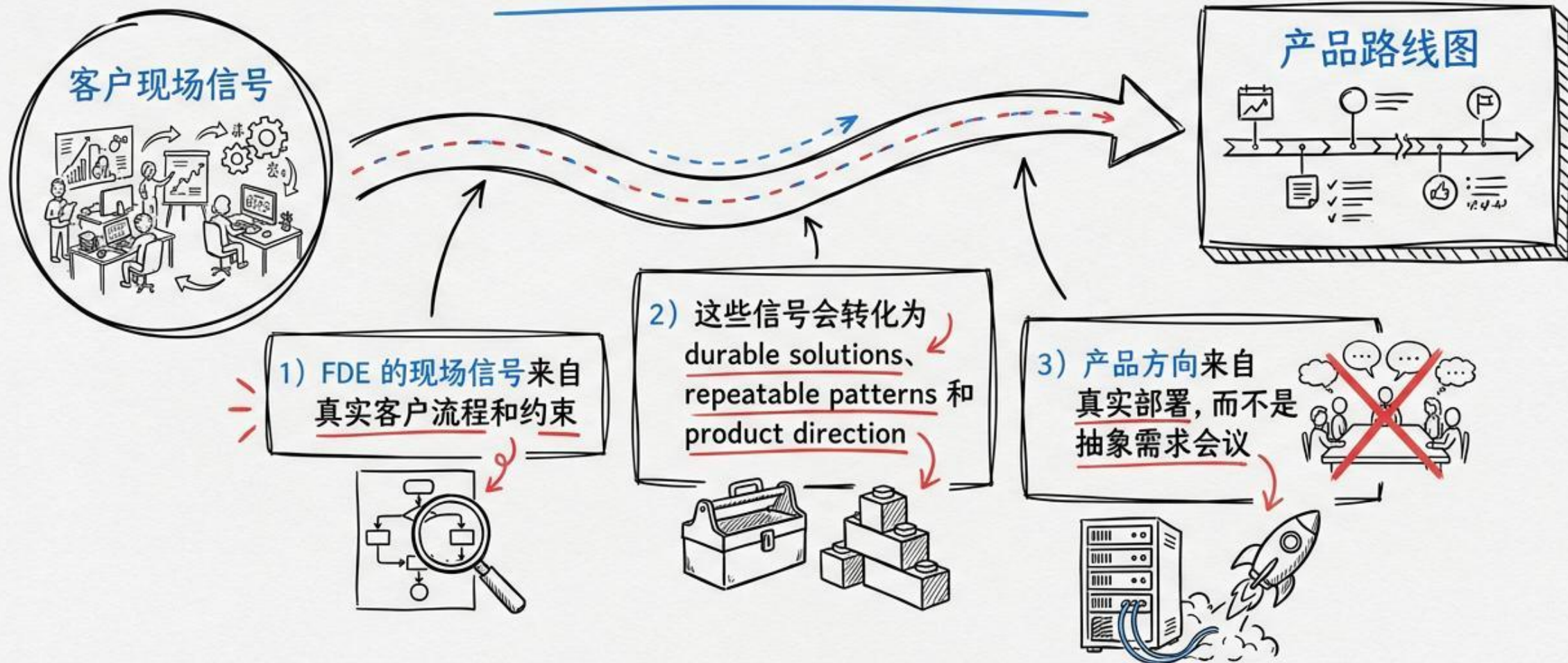
架构之后还要构建、上线、评测和迭代

- OpenAI FDE 岗位覆盖 discovery、scoping、system design、build 和 rollout。
- 这使 FDE 必须对 生产结果 负责，而非只对方案负责。
- 上线后的反馈会继续进入评测和产品方向。



FDE 也不是内部产品经理

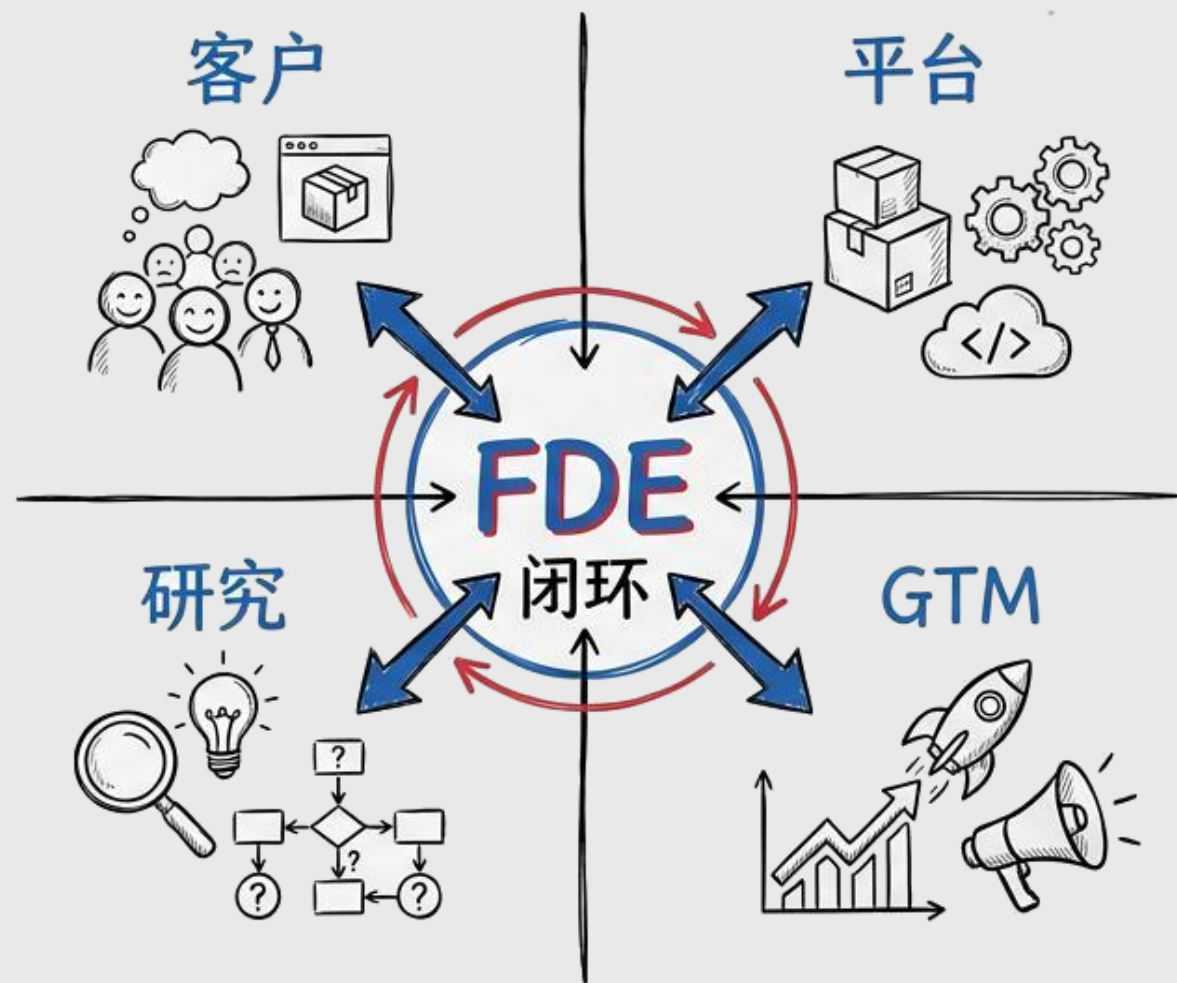
它在客户现场捕捉产品化机会



FDE 是现场工程组织

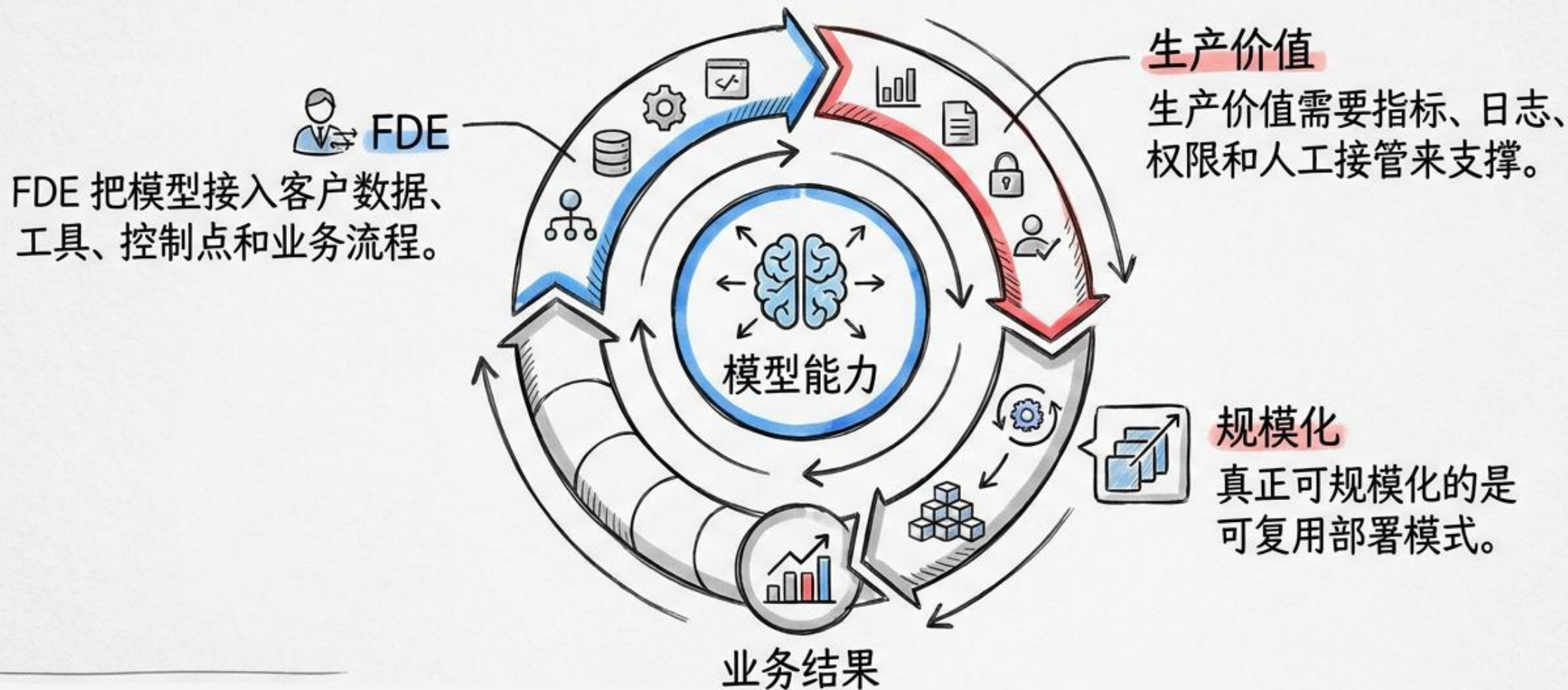
它把客户交付、平台开发和研究反馈放到同一闭环

- FDE 处在客户交付和核心平台开发交叉点。
- 它既处理端到端部署，又发现可复用平台能力。
- 这让现场部署成为产品演进的一部分。



从模型能力到生产价值

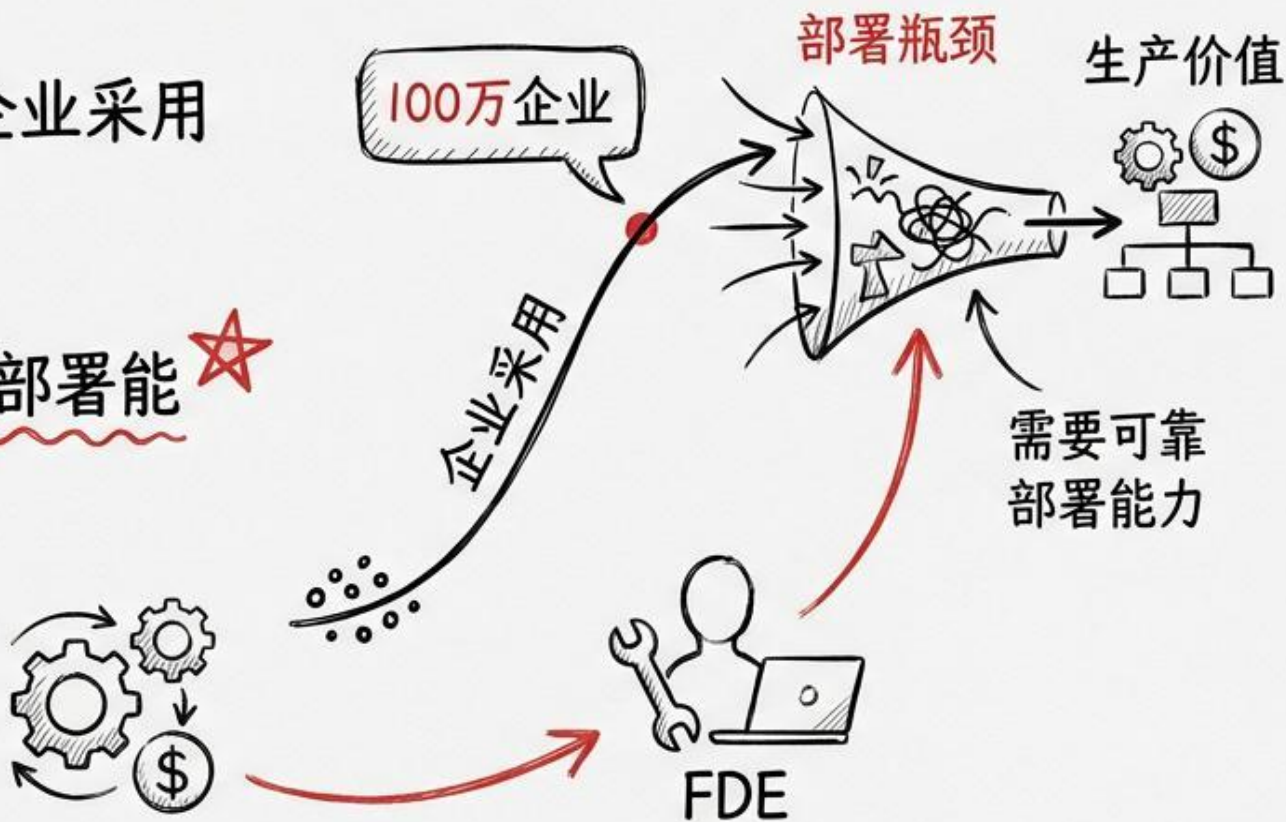
AI 的价值不在于能回答，而在于能进入业务流程



OpenAI 的企业采用背景

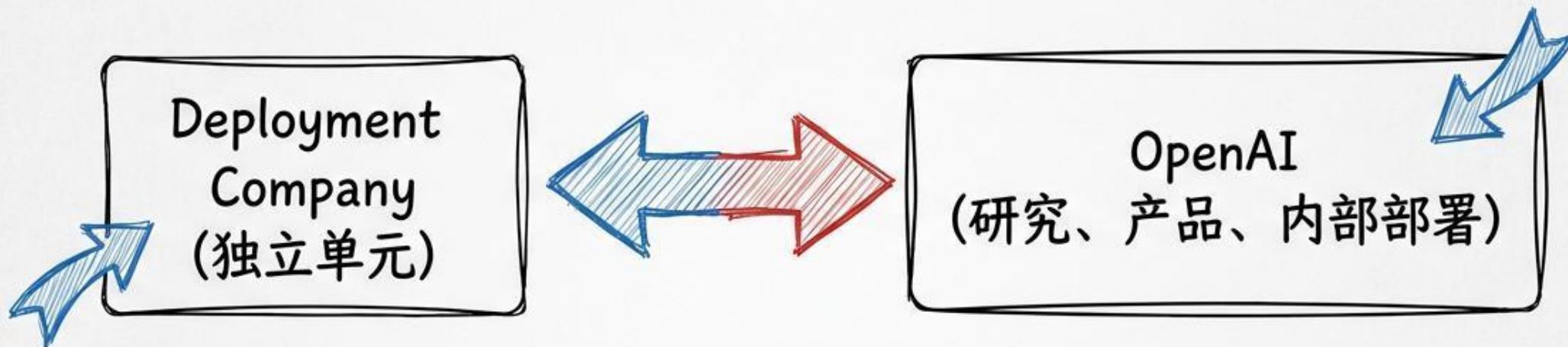
OpenAI 公告提到已有超过100万企业采用产品和 API

- 公告把 Deployment Company 放在企业采用快速扩张之后。
- 当试验足够多，真正稀缺的是可靠部署能力 ★
- FDE 是把采用量转化为生产价值的组织抓手。



新公司的结构信号

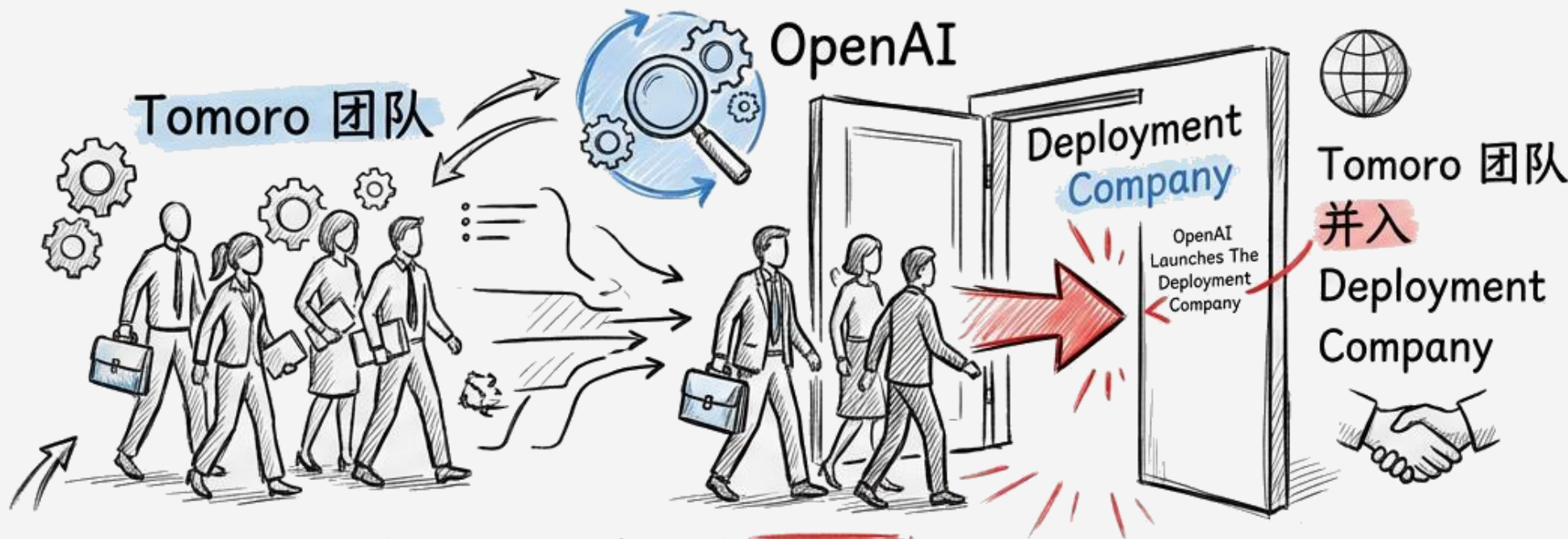
独立业务单元，但作为 OpenAI 的延伸



- 公告称 Deployment Company 是独立单元。
- 同时它会连接 OpenAI 的研究、产品和内部部署团队。
- 这种结构兼顾客户速度与前沿能力输入。

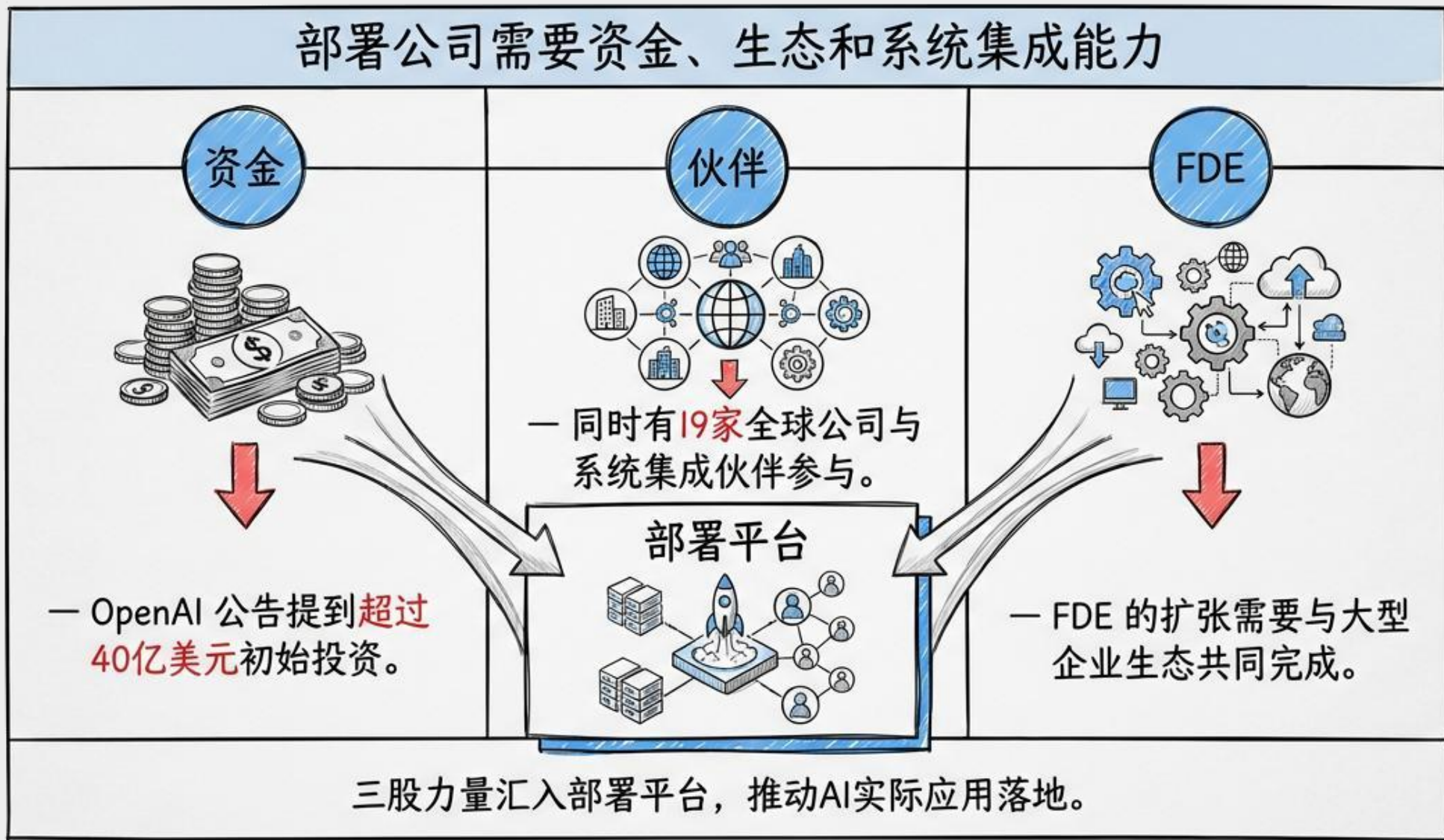
Tomoro 带来的部署能力

约150名 FDE 和 Deployment Specialists 成为启动资产



- OpenAI 公告披露已同意收购 Tomoro。
- Tomoro 会带来约150名有经验的 FDE 和部署专家。
- 这说明部署不是临时销售动作，而是 专门组织能力。

超过40亿美元初始投资与19家伙伴

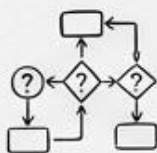


部署从诊断开始

不是先卖许可证，而是先识别高价值问题



— 合作开始于诊断客户最重要、最棘手的问题。

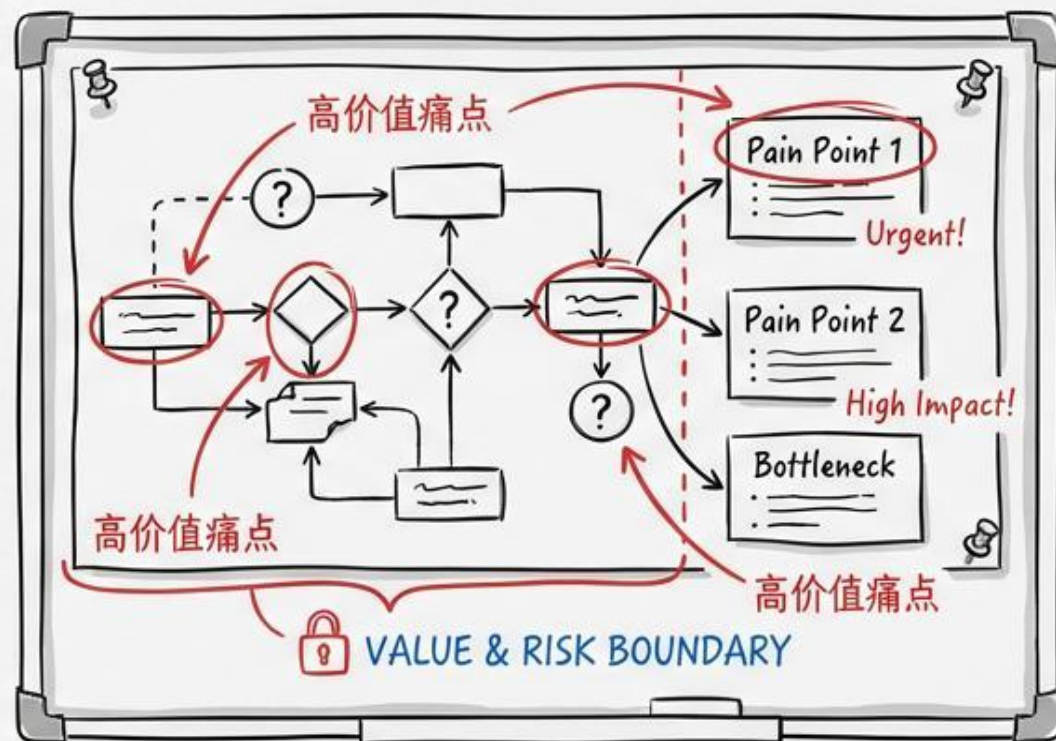


— 诊断阶段决定后续工作流是否值得投入。



— 好的 FDE 会先定义业务价值和风险边界。

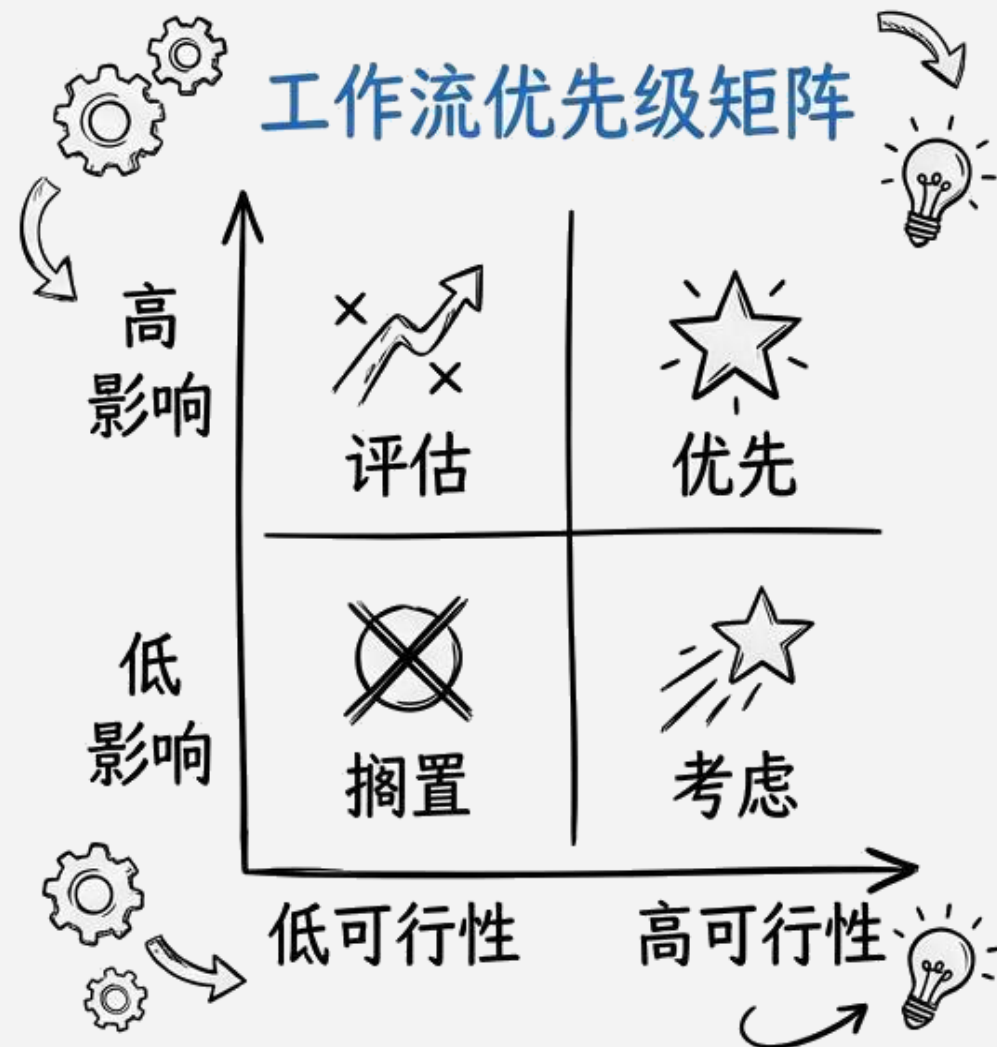
诊断白板上标注高价值痛点



优先工作流选择

FDE 不应同时改造所有流程

- 公告强调选择优先工作流。
- 优先级应由影响、可行性、风险和复用潜力共同决定。
- 小范围高质量闭环比大范围松散试点更重要。
- 聚焦产生最大价值。



设计、构建、测试、部署

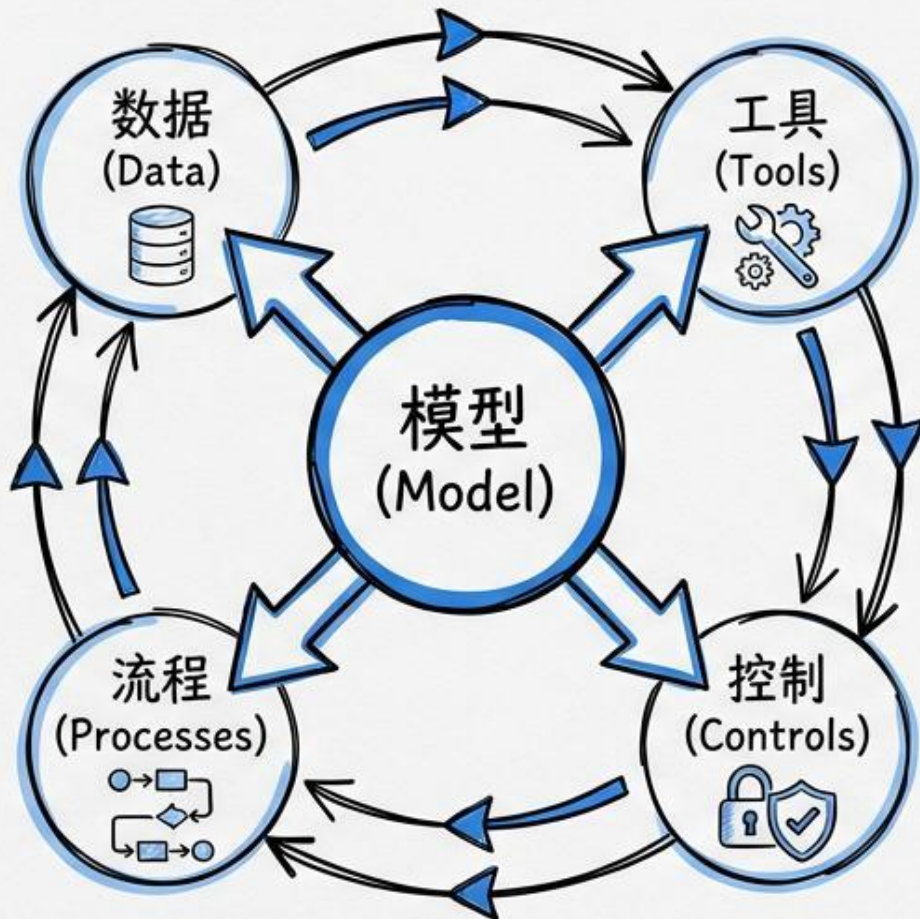
生产系统需要完整工程链条



- FDE 负责设计、构建、测试和部署。
- 测试必须覆盖真实样本、边界条件和失败模式。
- 部署必须包括监控、回滚、接管和评估。🚩

连接模型到数据、工具、控制和流程

这是 FDE 与普通模型接入的分水岭

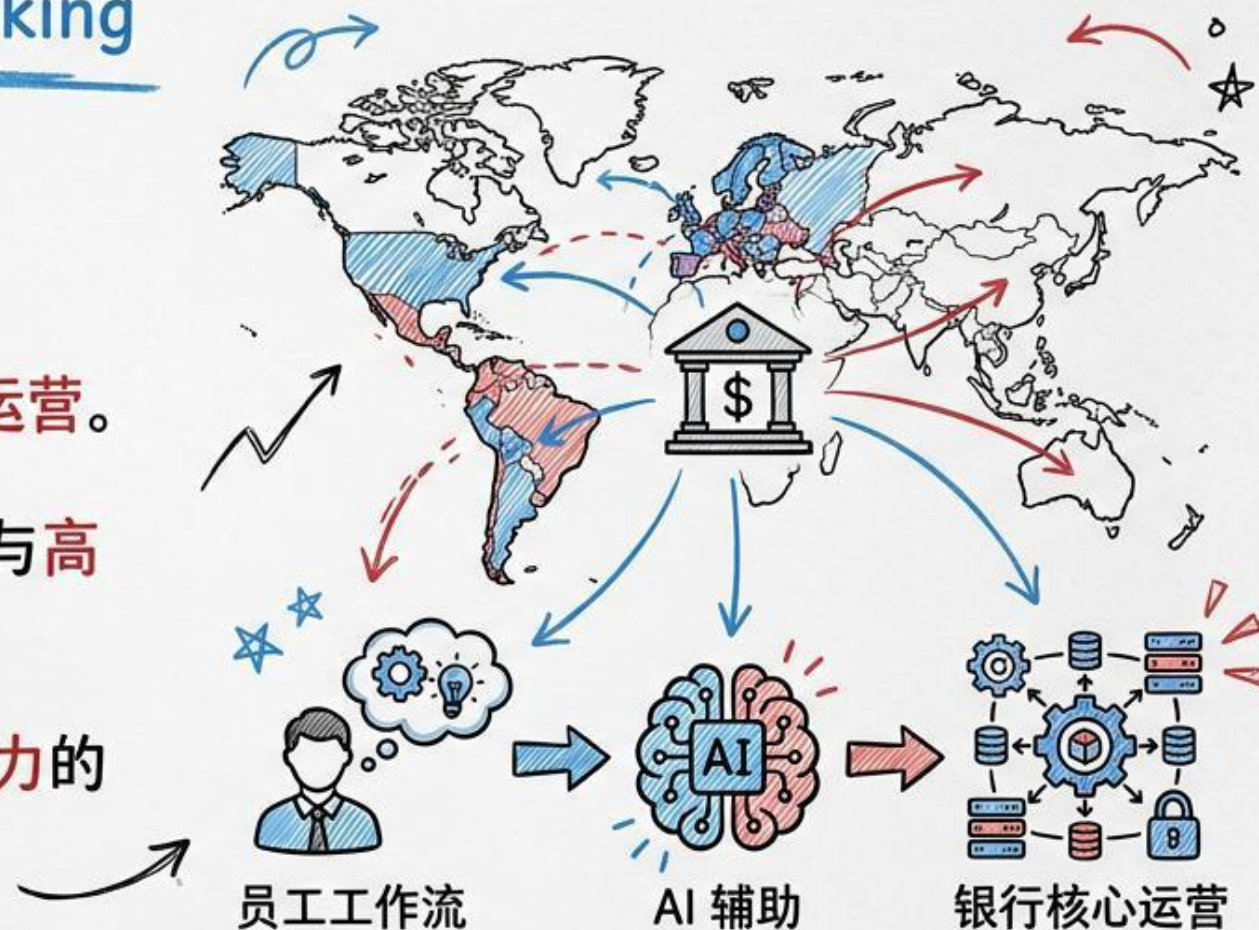


- 公告明确生产系统会连接 OpenAI 模型 与 客户数据、工具、控制和业务流程。
- 连接点越多，权限和责任 越重要。
- FDE 的核心是 控制下的系统集成。

BBVA: 银行级规模化部署

从员工生产力到 AI-native banking

- ✓ OpenAI 官方页面显示
BBVA 合作扩展到120,000名员工。
- ✓ 覆盖25个国家，AI 被嵌入银行核心运营。
- ✓ 金融场景说明 FDE 必须面对高监管与高信任要求。
- ✓ 规模化部署展示了对企业级 FDE 能力的极高要求。



John Deere: 农业推荐系统

FDE 进入季节性、专业性强的真实场景

- OpenAI 与 John Deere 合作部署种植季 AI 推荐。
- 团队与领域专家审查真实样本并建立定制评测。
- 官方页面称该系统帮助农民最多减少70%的化学品使用



案例说明 FDE 的价值标准

价值要能被真实业务结果衡量



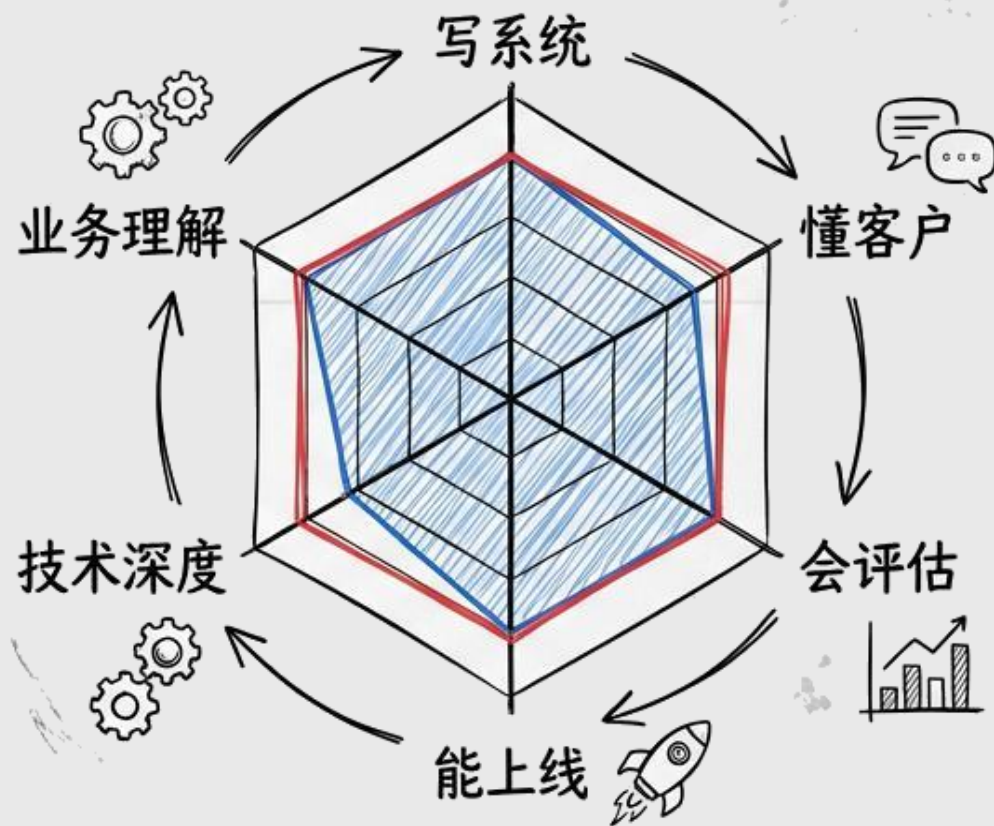
- 二者共同说明 FDE 不是演示，而是生产场景影响。
- 可衡量结果是 FDE 从试验走向部署的核心证据。
- BBVA 展示组织级规模，John Deere 展示专业场景结果。

FDE 岗位画像

端到端生产部署，而不是只做售前

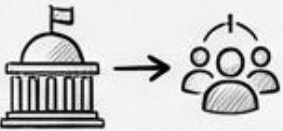
- OpenAI FDE 岗位描述强调复杂端到端部署
- 范围包括 discovery、scoping、system design、build、rollout
- 这要求 FDE 能写系统、懂客户、会评估、能上线

FDE能力雷达图



Gov FDE: 公共部门与 mission impact

政府客户要求更强的治理、合规和现场协作

-  Gov FDE 会嵌入政府与公共部门客户。
-  职责包含从原型到生产的技术交付。
-  它还要把现场模式和最佳实践反馈给产品与 GTM 团队。



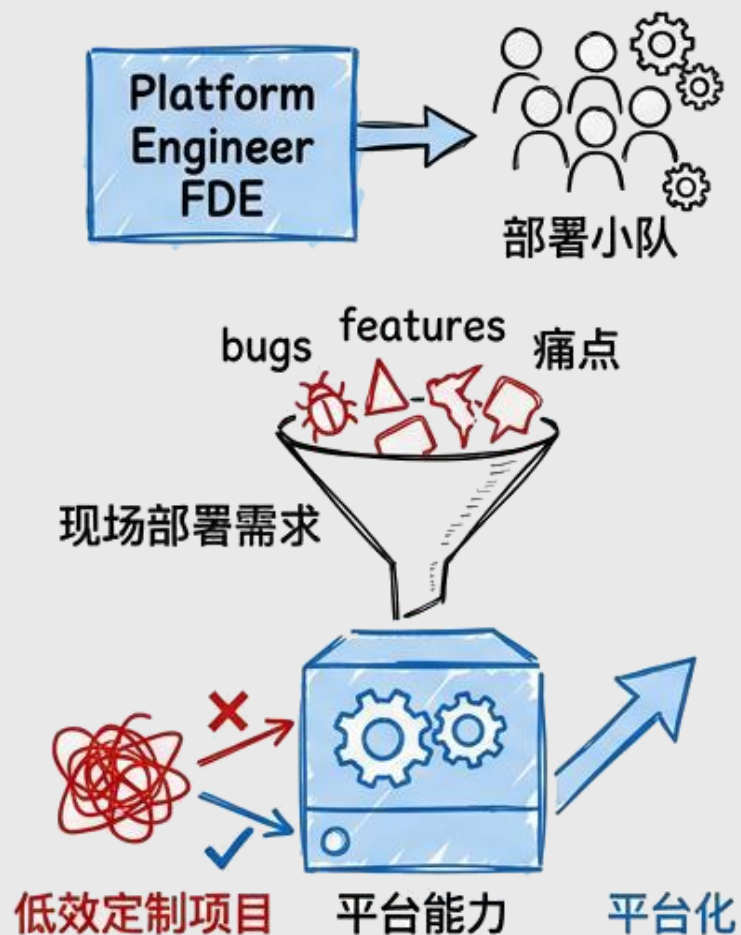
Strategy & Operations FDE

部署组织也需要组合管理与执行节奏



Platform Engineer FDE

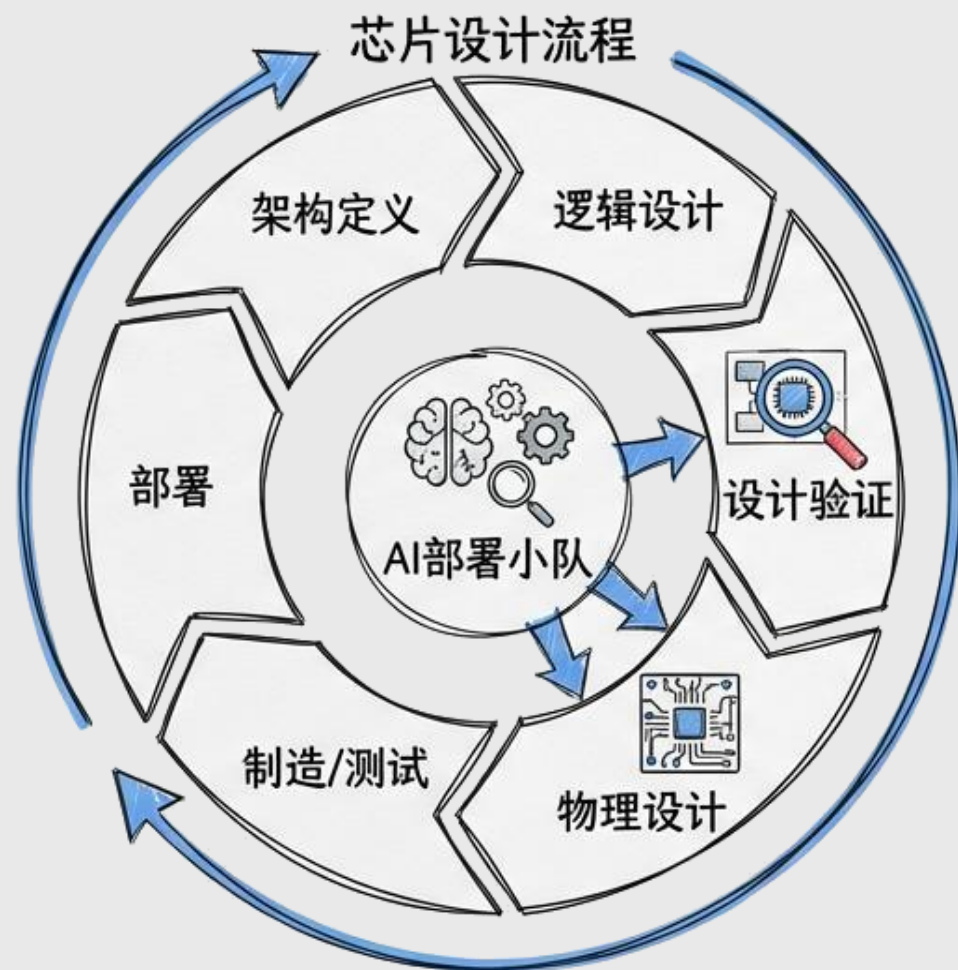
把现场部署放大为平台能力



- Platform Engineer FDE 嵌入部署小队。
- 目标是把真实部署中发现的问题转化为平台能力。
- 平台化是防止 FDE 变成低效定制项目的关键。

半导体 FDE：垂直行业正在出现

设计验证 与物理设计岗位显示 FDE 走向专业 workflow



- OpenAI 已出现面向半导体设计验证和物理设计的 **FDE 岗位**。
- 这些岗位服务芯片设计生命周期中的 **专业任务**。
- 垂直 FDE 将成为 **高价值行业部署** 的关键形态。

本

【



【

⚙️

💡

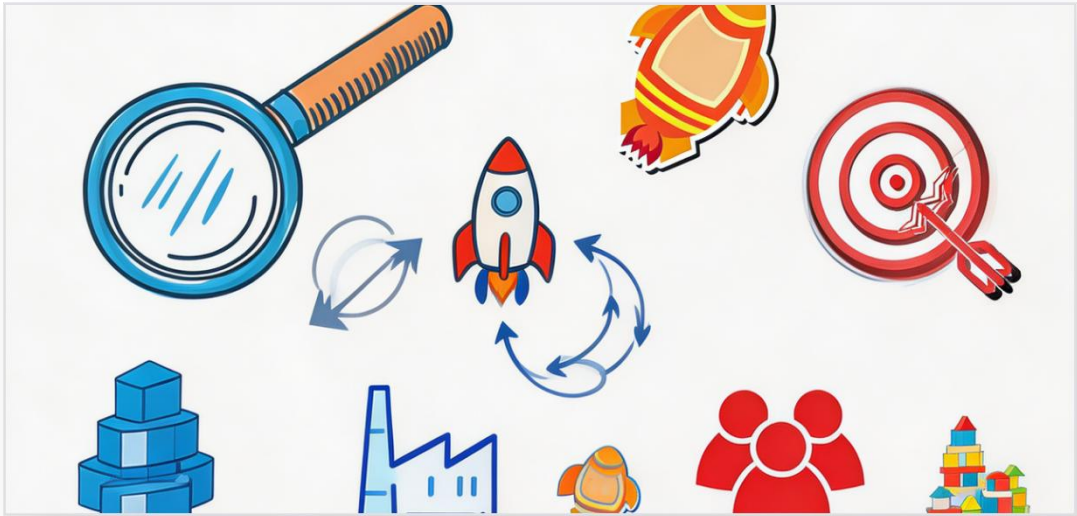
🛡️

六

0

0

0



0

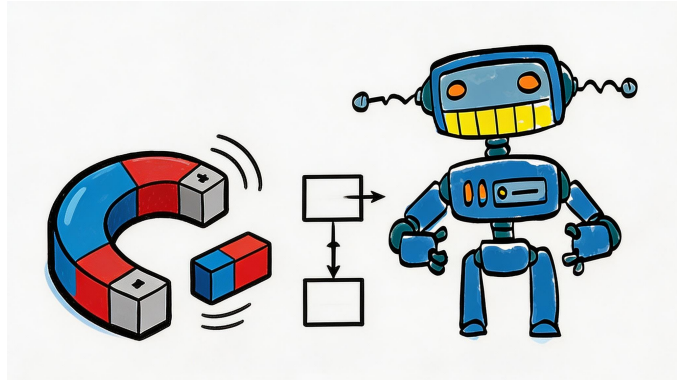
0

0



Z

大



A

过

现

FDE 能力栈总览

工程、产品、研究、行业和变革能力叠加

- FDE 需要生产级软件工程能力。
- 还需要理解客户流程、风险和指标。
- 最稀缺的是把现场问题抽象成可复用产品能力的的能力。



生产级全栈构建

FDE 必须能把原型推到可运行系统

OpenAI 岗位提到 production-grade frontend/backend code 能力。★

前端、后端、API、权限和日志都服务于生产闭环。

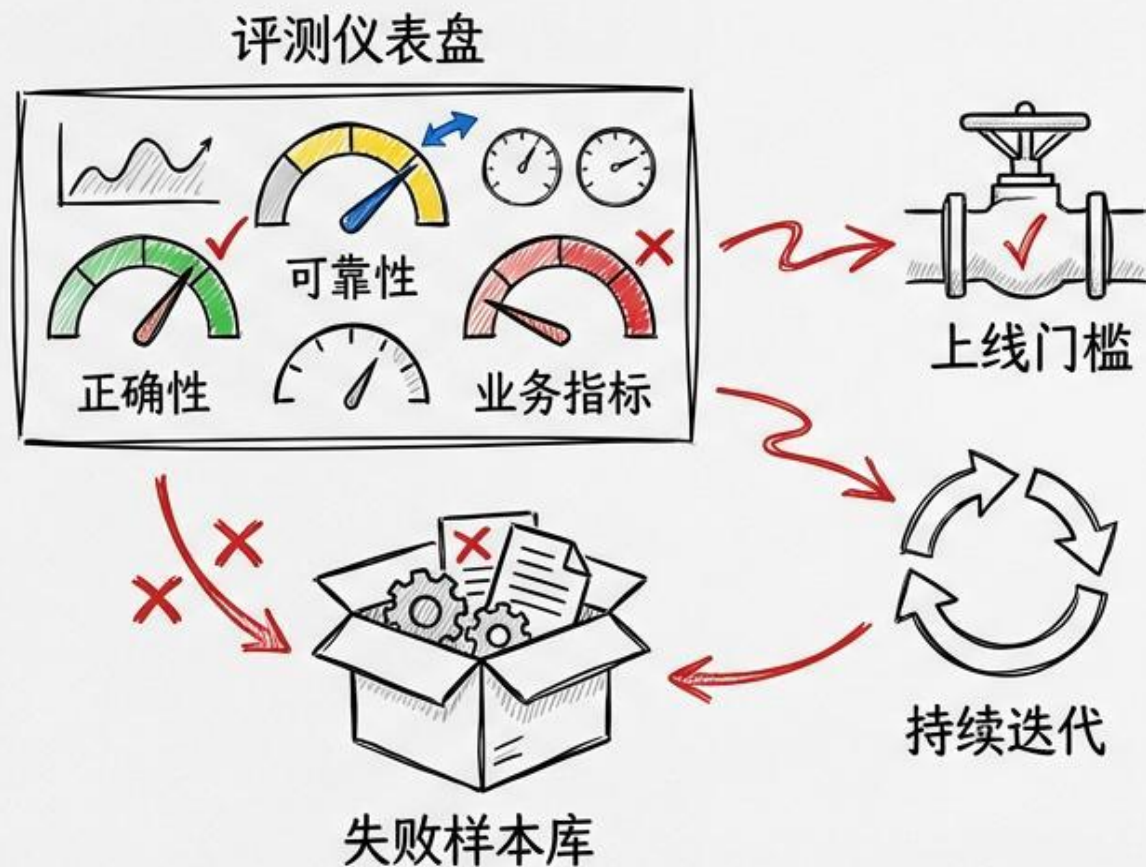
原型只是起点，稳定运行才能交付。→ ✓



从原型代码到生产服务的迁移

评测工程是 FDE 的安全阀

没有评测，部署只是在赌运气



- FDE 页面提到**定制评测系统**用于测量**准确性**。
- 评测应覆盖**正确性**、**可靠性**、边界样本和业务指标。
- 评测结果应进入**持续迭代**和**上线门槛**。

安全与权限是第一层工程约束

AI 系统接入企业工具后，权限风险会放大

- 官方定义把安全模型和权限列为**核心约束**。
- FDE 必须处理**身份、访问范围、工具白名单**和**最小权限**。
- 权限设计决定系统能否进入**核心流程**。

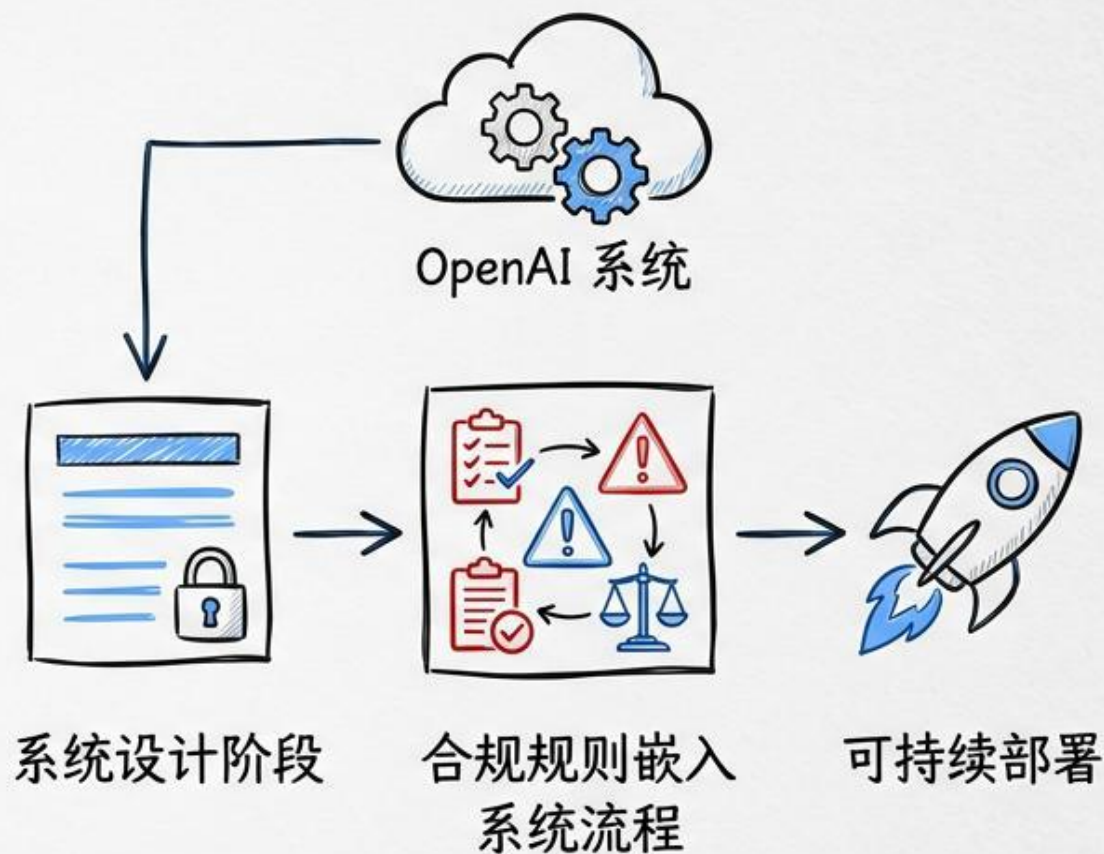
权限网关与工具白名单



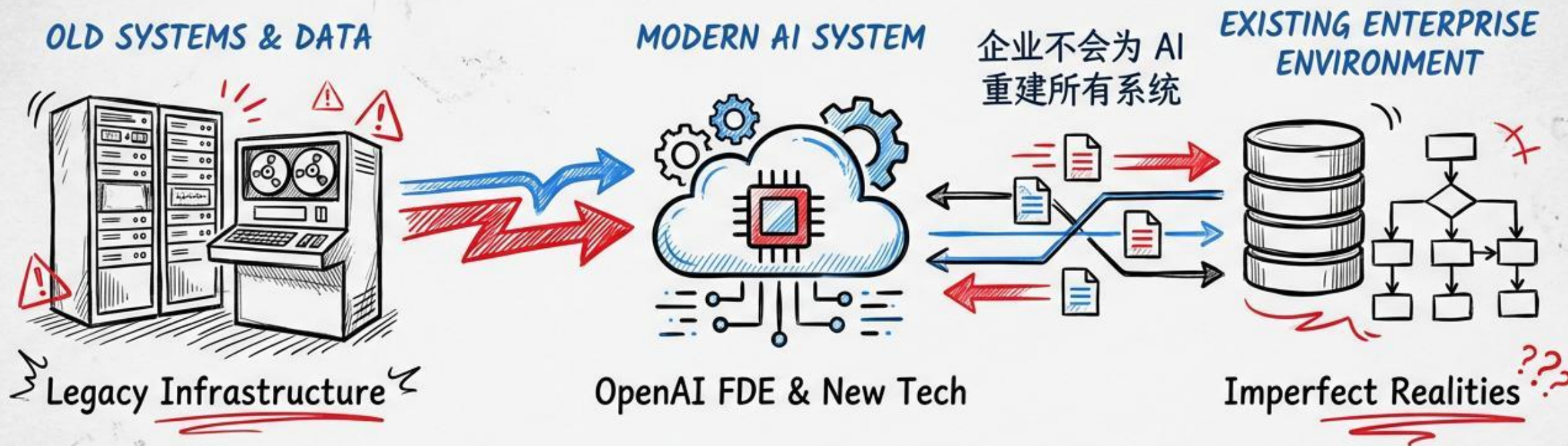
治理与合规不能后置

FDE 要把规则写进系统，而不是写在报告里

- 治理和合规是 OpenAI 官方定义中的**核心约束**。
- 高风险行业需要**审批、日志、审计和责任归属**。
- 合规设计**越早进入**，部署越**可持续**。



遗留基础设施是现实入口



- OpenAI官方定义：把 legacy infrastructure 作为 核心约束。
- FDE 任务：要连接 旧系统、旧数据和旧流程。
- 真正部署能力：体现在与 *不完美环境* 共存。

运营控制决定能不能日常运行

➔ 上线之后才是真正的开始



— 运营控制包括监控、告警、回滚、例外处理和人工接管。

— 没有运营控制，AI 系统很难被关键流程接受。

— FDE 必须把运行期设计作为架构的一部分。



FDE 与领域专家共创

真实问题必须由现场专家解释

— OpenAI 官方页面强调与领域专家一起理解问题。

专家提供业务知识

— 提供业务语言、例外情况和价值判断。

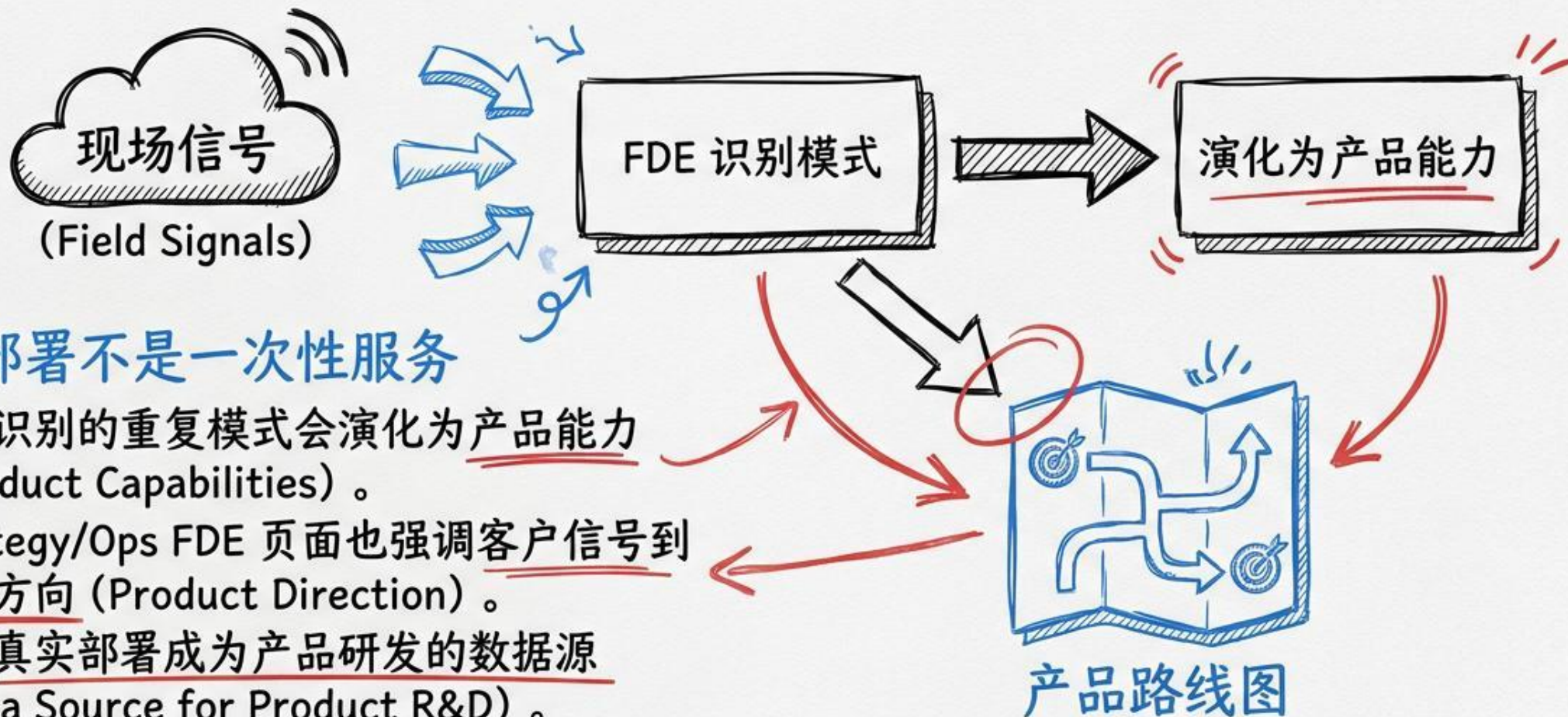


围绕流程图协作

FDE 转化知识

— 把这些知识转成系统设计和评测样本。

客户信号进入产品路线



现场部署不是一次性服务

- FDE 识别的重复模式会演化为产品能力 (Product Capabilities)。
- Strategy/Ops FDE 页面也强调客户信号到产品方向 (Product Direction)。
- 这让真实部署成为产品研发的数据源 (Data Source for Product R&D)。

Build, Prove, Generalize

从定制系统到可复用能力

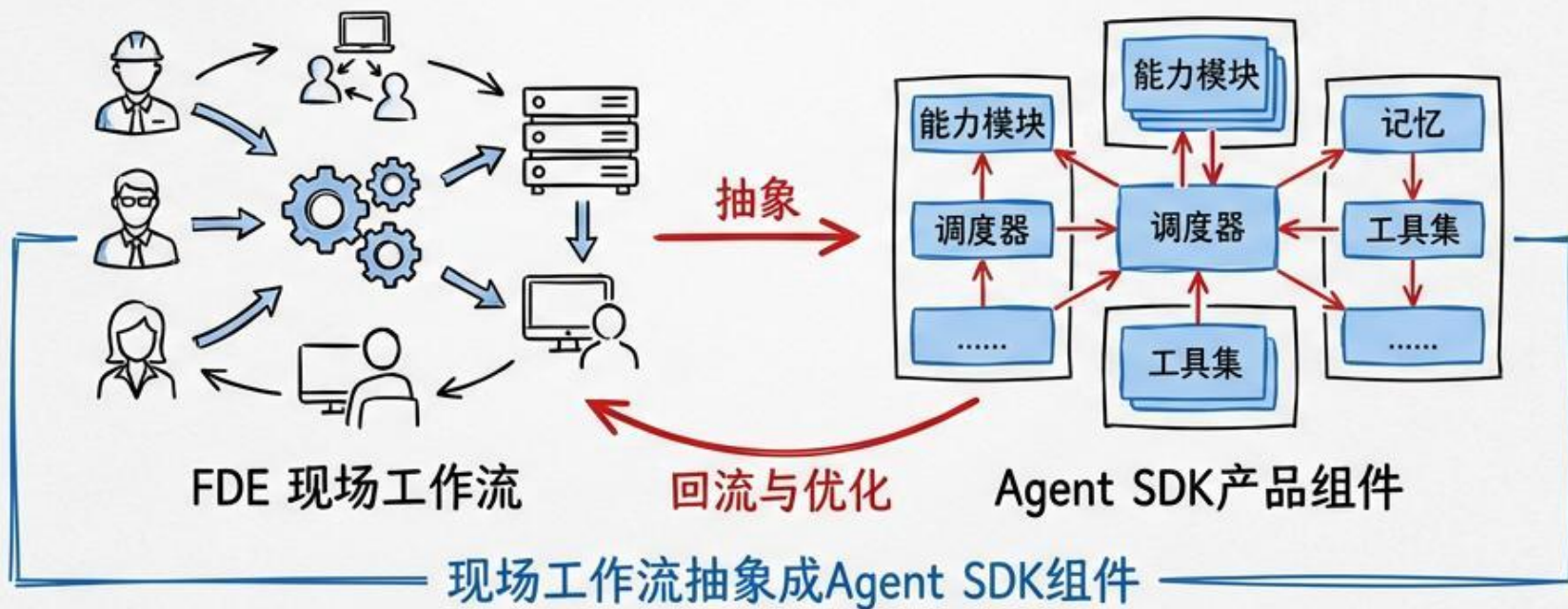
- OpenAI 页面把部署到产品的循环概括为 build, prove, generalize。
- 先在真实场景构建，再证明价值，最后抽象成产品能力。
- 这正是 FDE 与传统项目交付的核心差异。



Agent SDK 与部署产品化

FDE 模式会回流到 Agent SDK 等产品能力

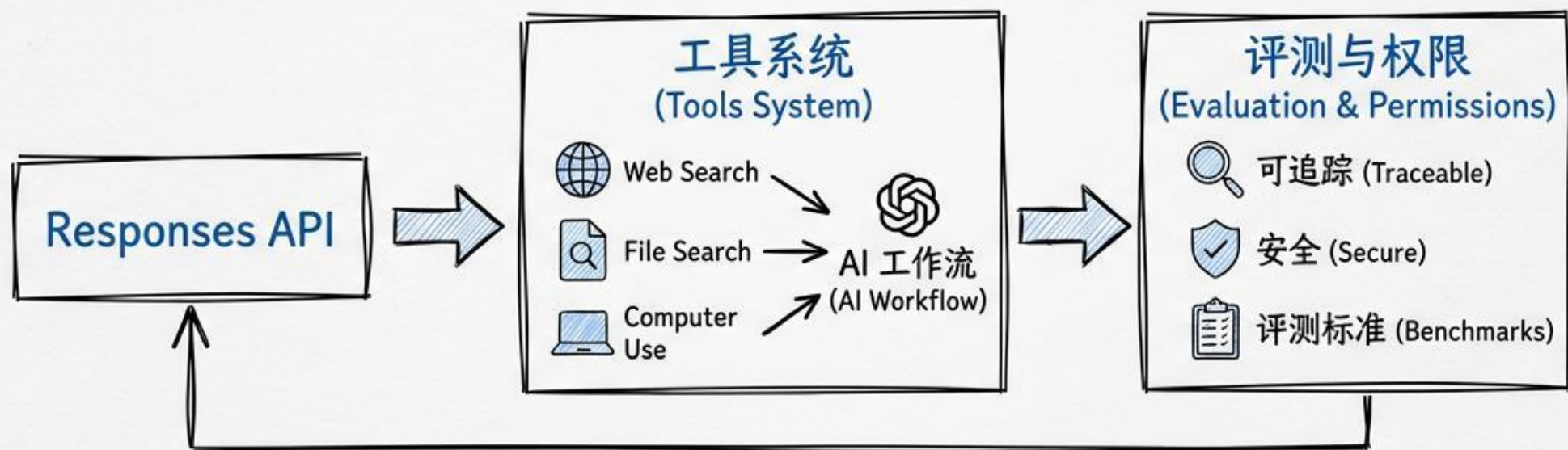
- 官方页面指出重复模式会进入 Agent SDK 等产品能力。
- FDE 现场经验能帮助平台知道哪些抽象最有用。
- Agent 平台不是凭空设计，而是从部署中生长。



Responses API 与工具系统

FDE 需要可编排、可追踪、可接工具模型接口

- OpenAI 将 Responses API 作为构建 agentic 应用的重要原语。
- Web search、file search、computer use 等工具让 AI 进入 workflow。
- 工具越强，评测和权限越重要。

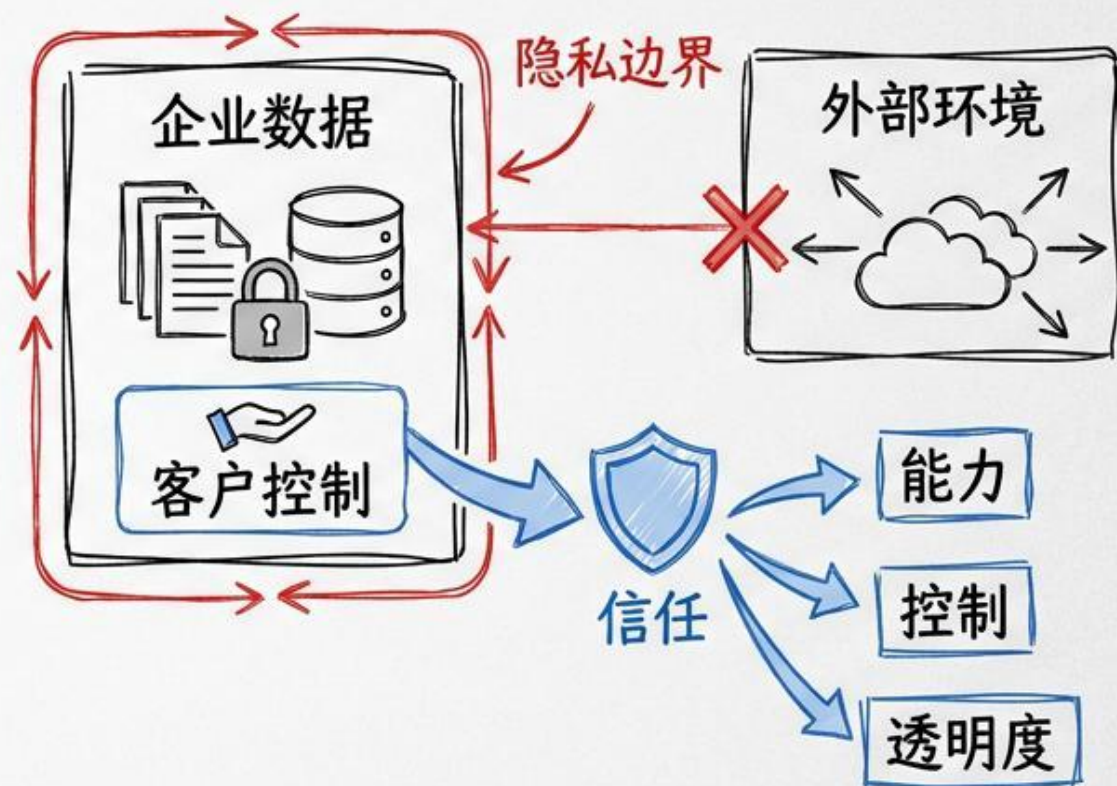


企业数据与隐私边界

部署必须尊重客户对业务数据的控制

- OpenAI 企业隐私页面强调**业务数据控制**和默认**不用于训练**。
- FDE 需要把**数据边界**写进架构与合同。
- 客户信任来自**能力**、**控制**和**透明度**共同作用。

数据边界和隐私控制示意图



OpenAI FedRAMP 场景

政府部署需要单独看功能边界与授权环境



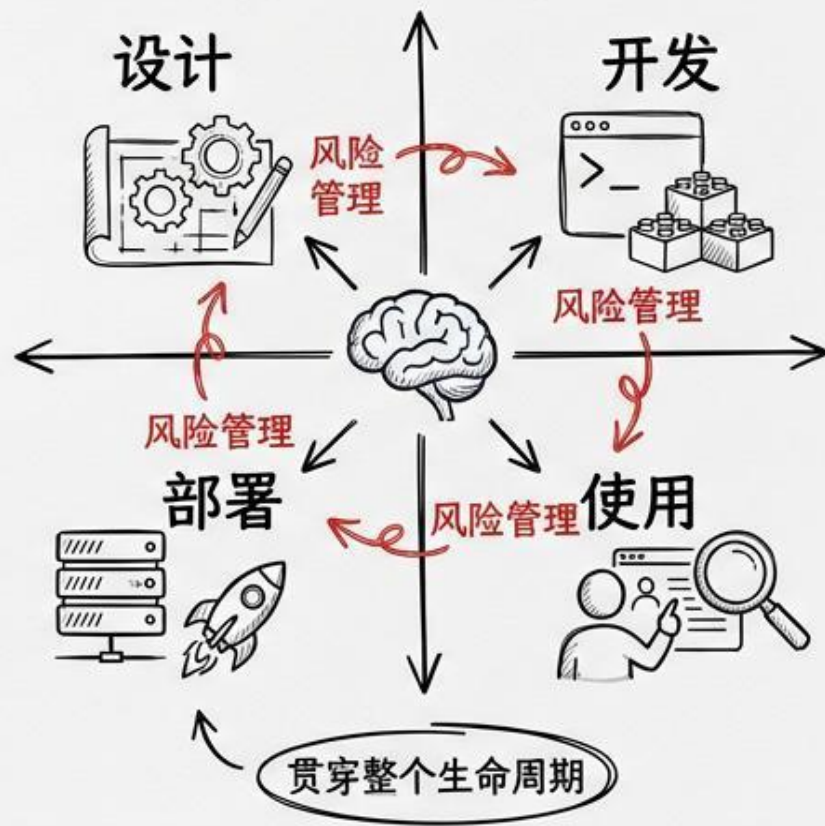
- OpenAI FedRAMP 页面说明 FedRAMP 环境功能与商业环境不同。
- API 需要使用指定 `gov.api.openai.com` 端点。
- FedRAMP Marketplace 列出 OpenAI 产品为 FedRAMP Certified。

NIST AI RMF: 可信 AI 的管理框架

FDE 应把风险管理贯穿设计、开发、部署和使用

- ✓ — NIST AI RMF 1.0 是自愿、跨行业、用例无关的 AI 风险管理框架。
- ✓ — 它帮助组织管理 AI 风险并促进 可信、负责任的 AI。
- ✓ — FDE 可以把它作为 高风险部署的治理语言。

AI风险管理框架四象限



NIST Generative AI Profile

生成式 AI 需要专门风险画像

- NIST 发布 Generative AI Profile 作为 AI RMF 的伴随文件。
- 它帮助组织识别生成式 AI 特有风险。
- FDE 在生成式应用部署中应结合该 profile 做风险清单。



生成式AI风险雷达

NIST SSDF 与 SP 800-218A

AI 部署首先仍是安全软件开发问题

安全软件开发生命周期



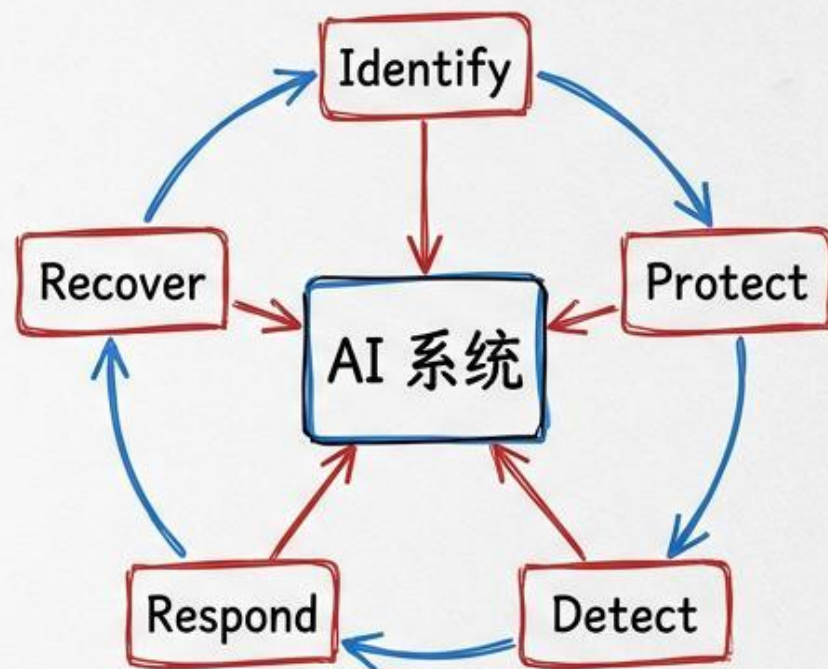
AI 模型实践

- NIST SSDF 提供安全软件开发实践。
- SP 800-218A 扩展到生成式 AI 与双用途基础模型。
- FDE 的生产构建应遵循安全开发、供应链和模型特定实践。

NIST CSF 2.0

FDE 需要把网络安全结果嵌入部署目标

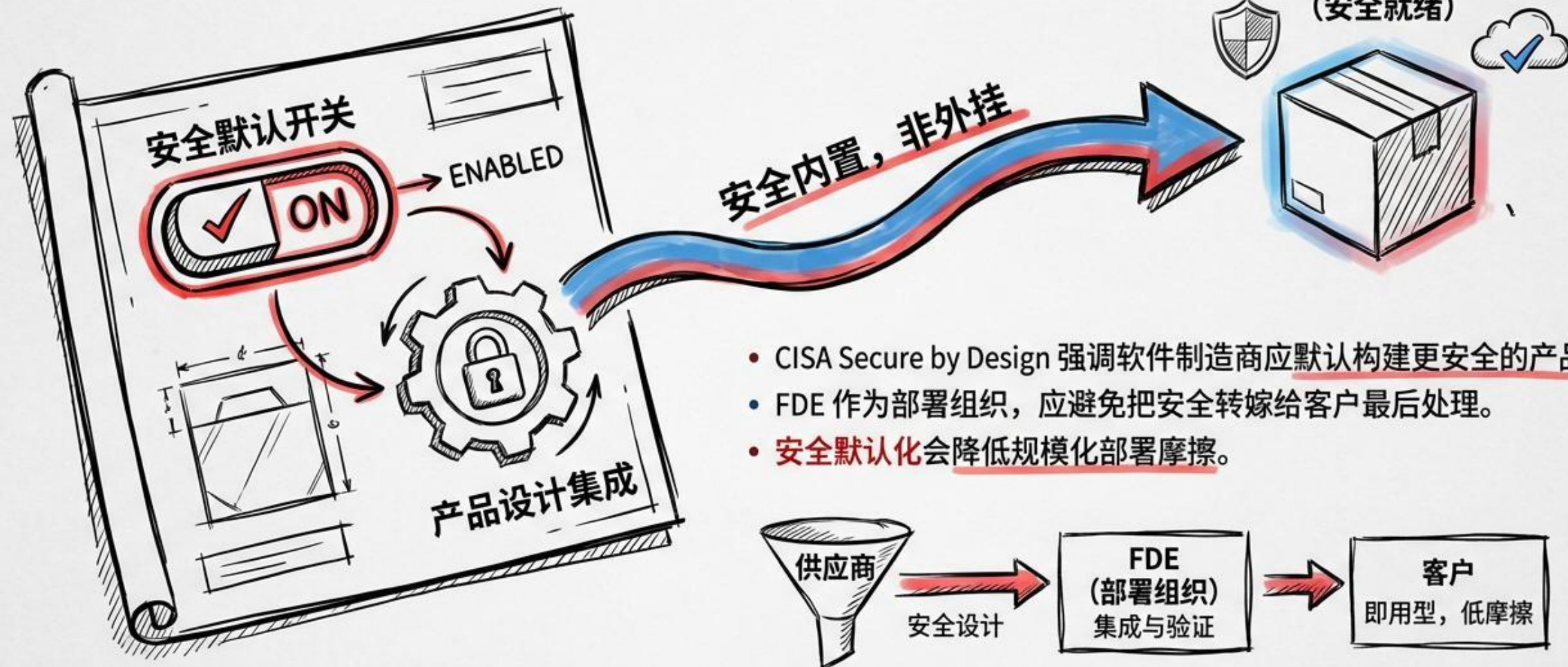
- NIST CSF 2.0 用于管理网络安全风险。
- 它提供任何组织都可使用的成果分类。
- FDE 应把 Identify, Protect, Detect, Respond, Recover 等能力映射到 AI 系统。



网络安全框架与AI部署映射

CISA Secure by Design

供应商应把安全作为默认产品属性



Section 508 与 USWDS

公共服务场景请考虑可访问性

— Section 508 要求联邦机构 ICT 提供等效访问。

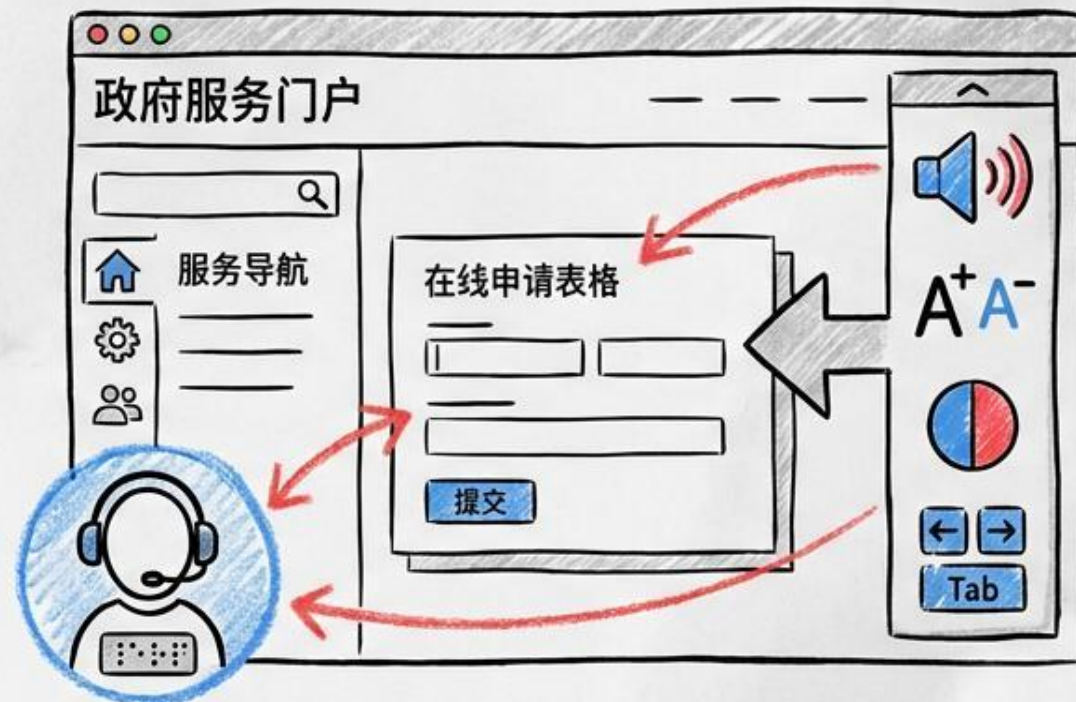
- 联邦机构必须确保其电子和信息技术（ICT）可供残障人士访问和使用，实现平等参与。

— USWDS 提供建设可访问、移动友好政府网站的设计系统。

- USWDS 是一个开源设计系统，提供组件、样式和指南，帮助构建一致且无障碍的政府数字服务。

— Gov FDE 若构建面向公众或公务员界面，需把无障碍当作工程要求。

- 政府前端开发（Gov FDE）在创建用户界面时，应从一开始就将无障碍标准纳入工程规范和测试流程。



无障碍用户

无障碍界面与政府服务门户集成示意

BLS 没有 FDE 独立职业口径

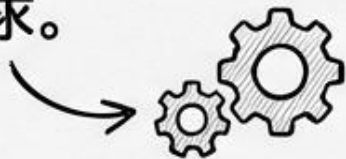
本报告使用相邻职业作为技能市场代理

FDE 现状

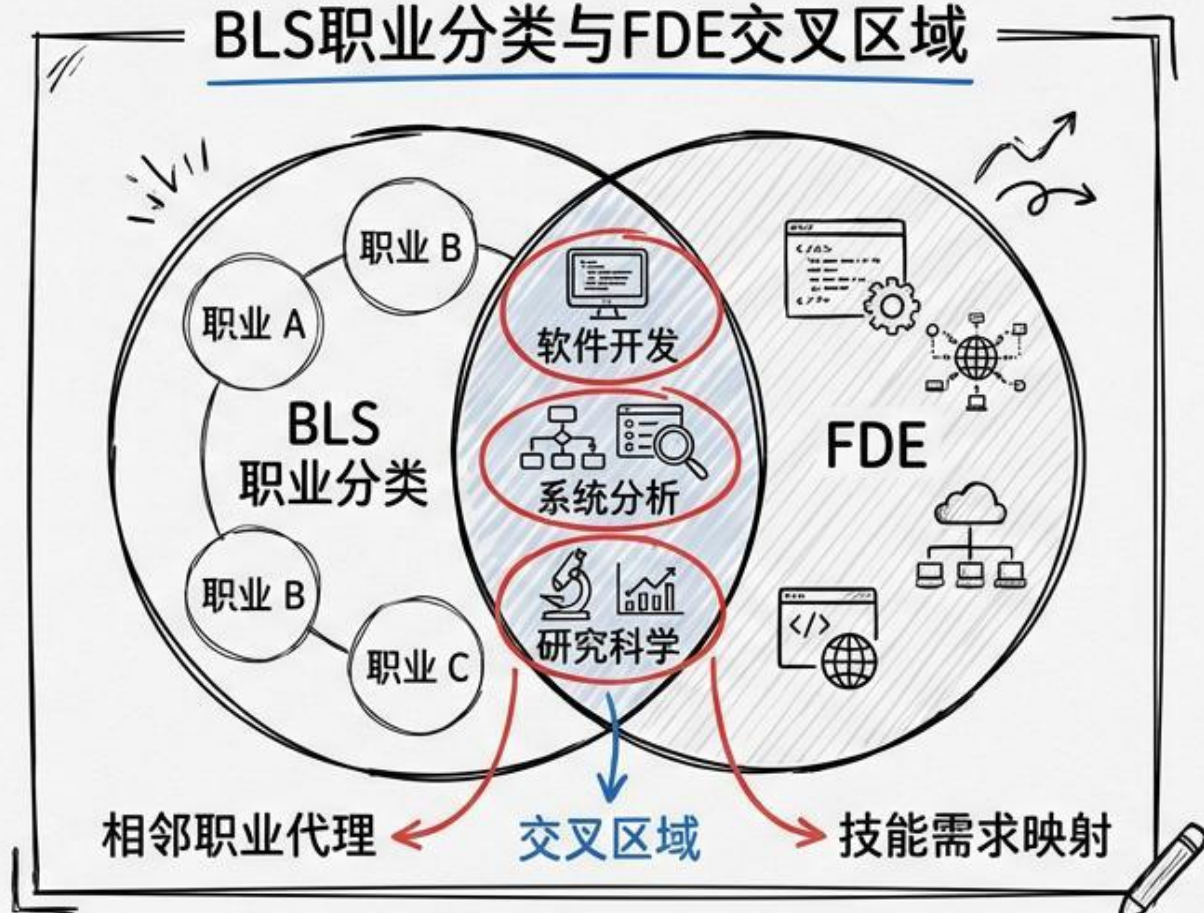
✕ Forward Deployed Engineer 不是 BLS 独立职业分类。

? 因此不能直接声称 FDE 市场规模。

➡ 可用 软件开发、系统分析、研究科学 等相邻职业观察技能需求。

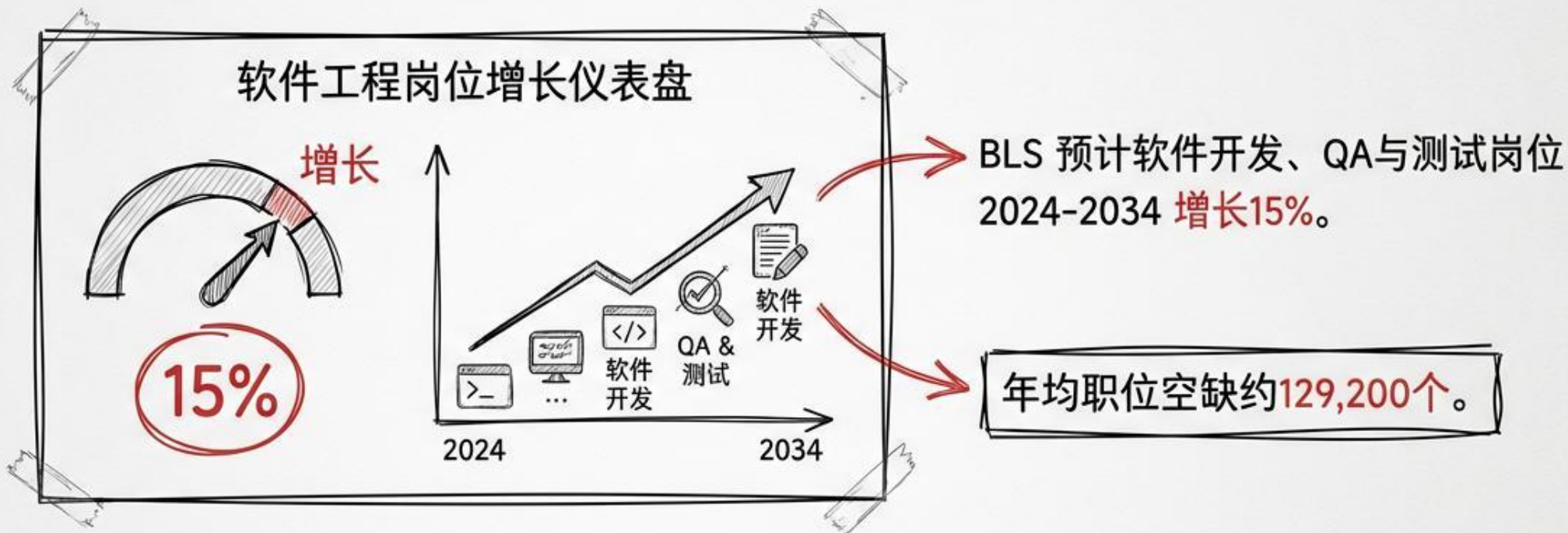


BLS职业分类与FDE交叉区域

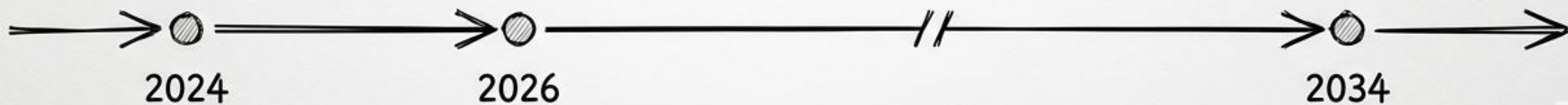


软件开发相关岗位增长 15%

FDE 的工程底盘来自生产软件能力

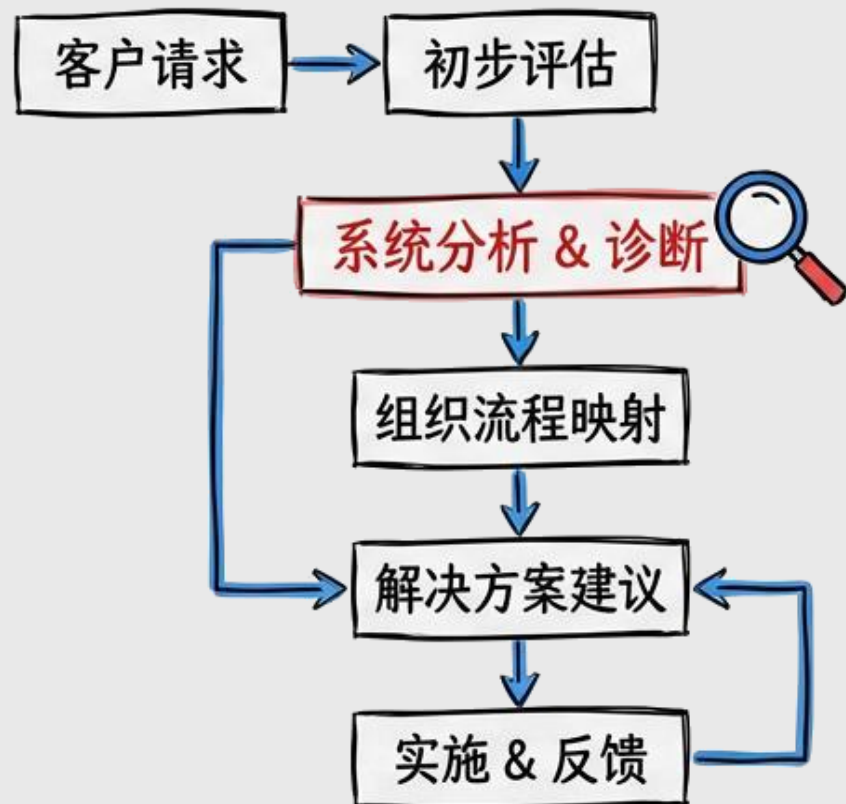


这支撑 FDE 对生产级软件工程能力的长期需求。

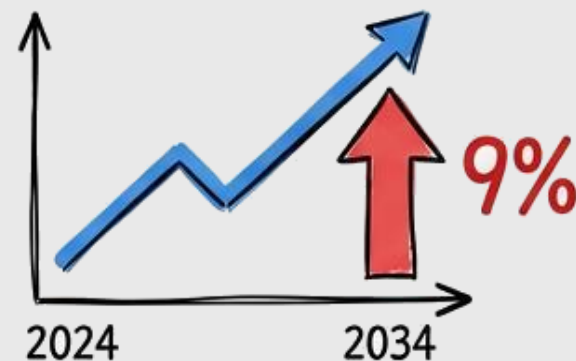


计算机系统分析师增长 9%

FDE 需要系统分析与组织流程理解



系统分析能力是
FDE 诊断客户流程
的重要组成。



BLS 预计计算机系统分析师
2024-2034 增长9%。



年均职位空缺约34,200个。

网页开发与数字设计增长 7%

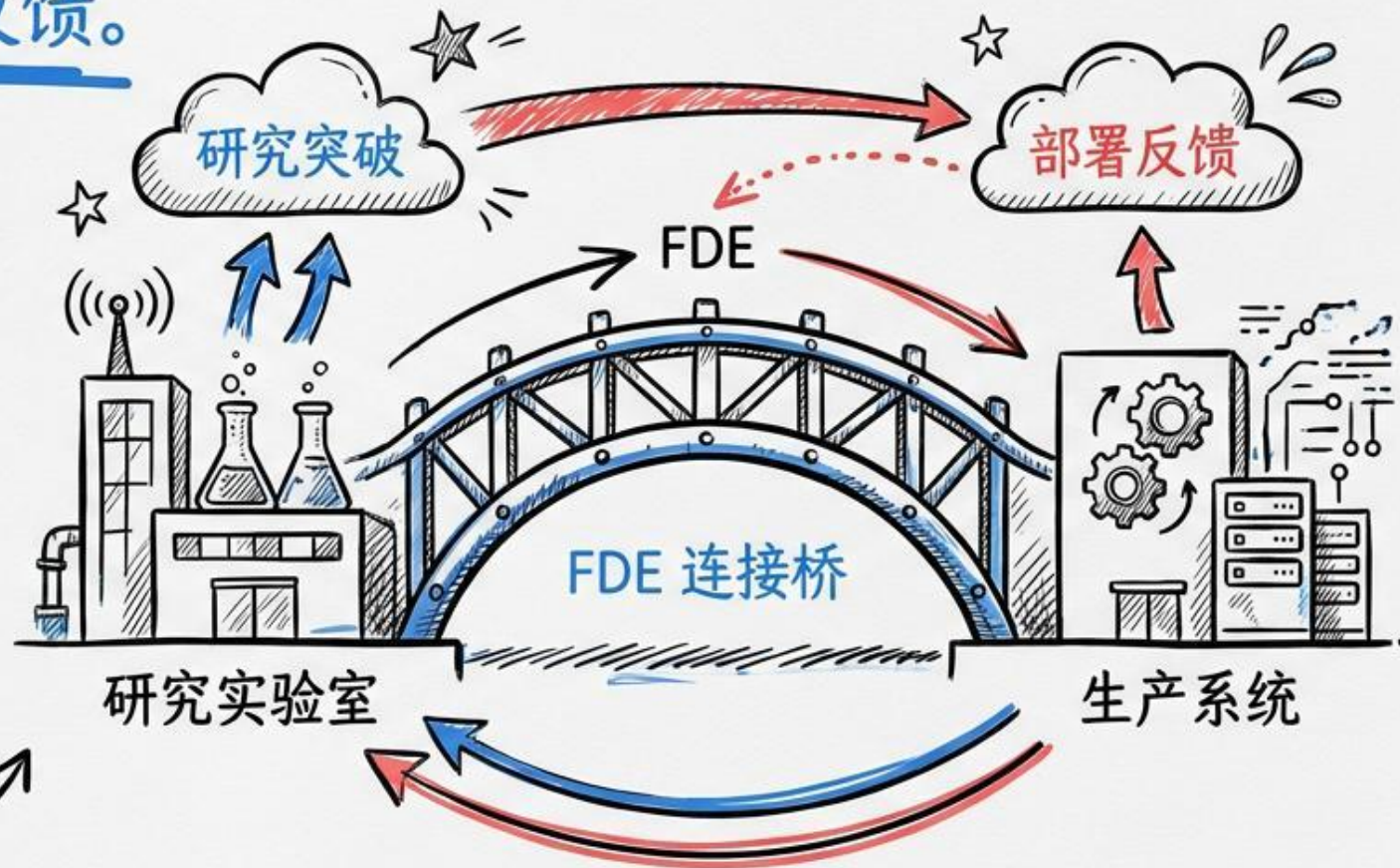


- ✓ BLS 预计网页开发与数字设计 2024-2034 增长 7%。
- ✓ 年均职位空缺约14,500个。
- ✓ 虽然 FDE 不是前端部署工程师, 但生产系统仍需要清晰界面。

计算机与信息研究科学家增长 20%。

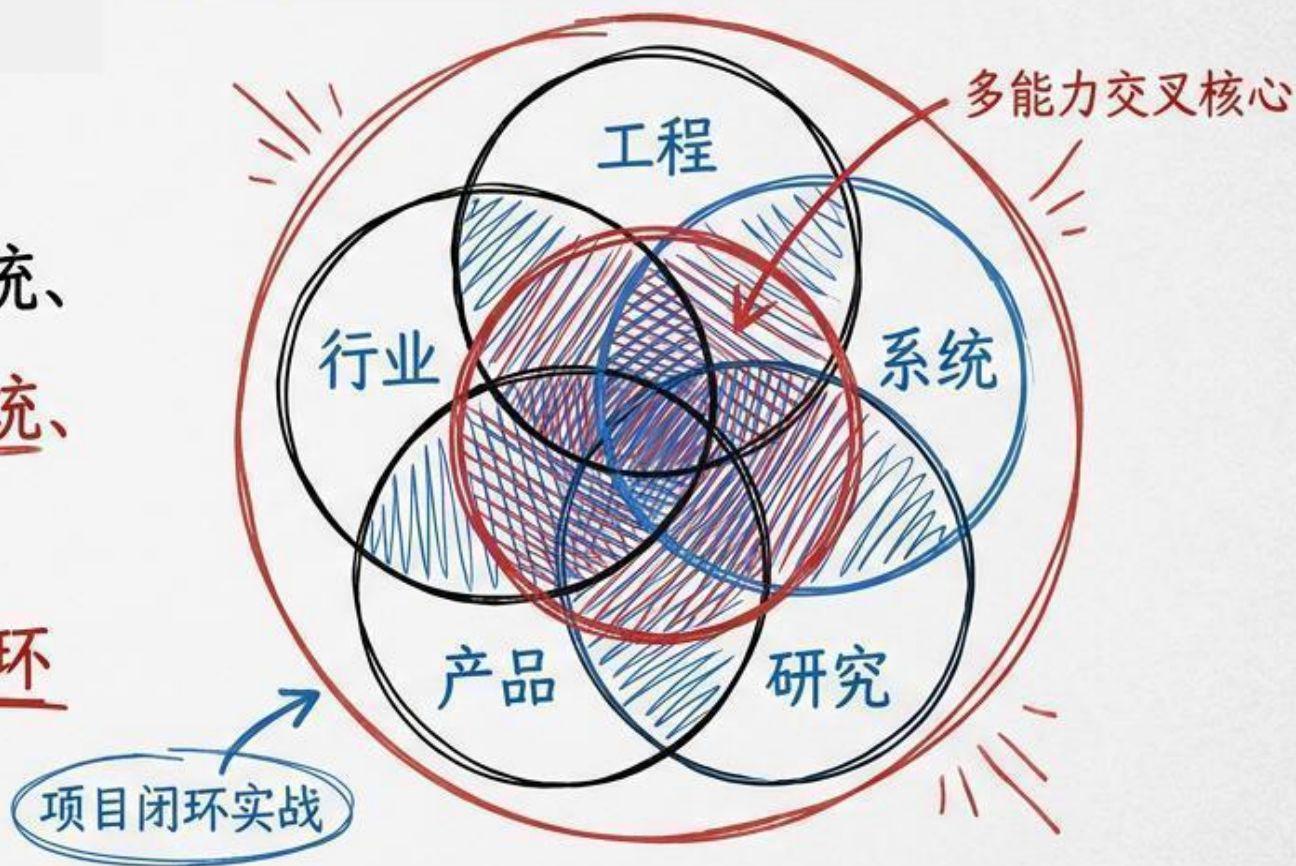
FDE 连接研究能力和部署反馈。

- 1) BLS 预计计算机与信息研究科学家 2024-2034 增长 20%。
- 2) 年均职位空缺约 3,200 个。
- 3) FDE 需要把研究突破翻译成生产系统。



FDE 是多职业交叉能力

- FDE 不应被单一岗位口径限制。
- 它更像多能力组合：工程、系统、
- 它更像多能力组合：工程、系统、研究、产品和行业。
- 企业培养 FDE 人才要从项目闭环而非单课培训开始。

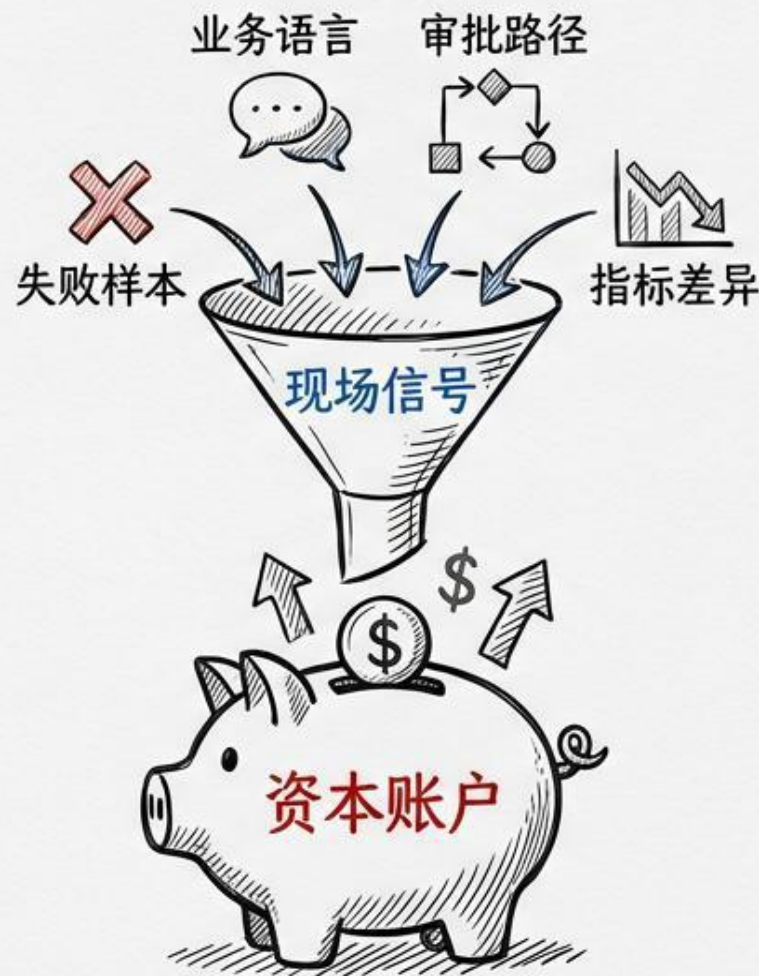




原创概念一：现场信号资本

真实部署信号是 OpenAI FDE 的战略资产

- **现场信号**包括失败样本、业务语言、审批路径和指标差异。
- 这些信号可转化为**评测集**、产品需求和部署 playbook。
- 谁能更快吸收现场信号，谁就更接近**真实产品市场**。



原创概念二：约束内生化工程

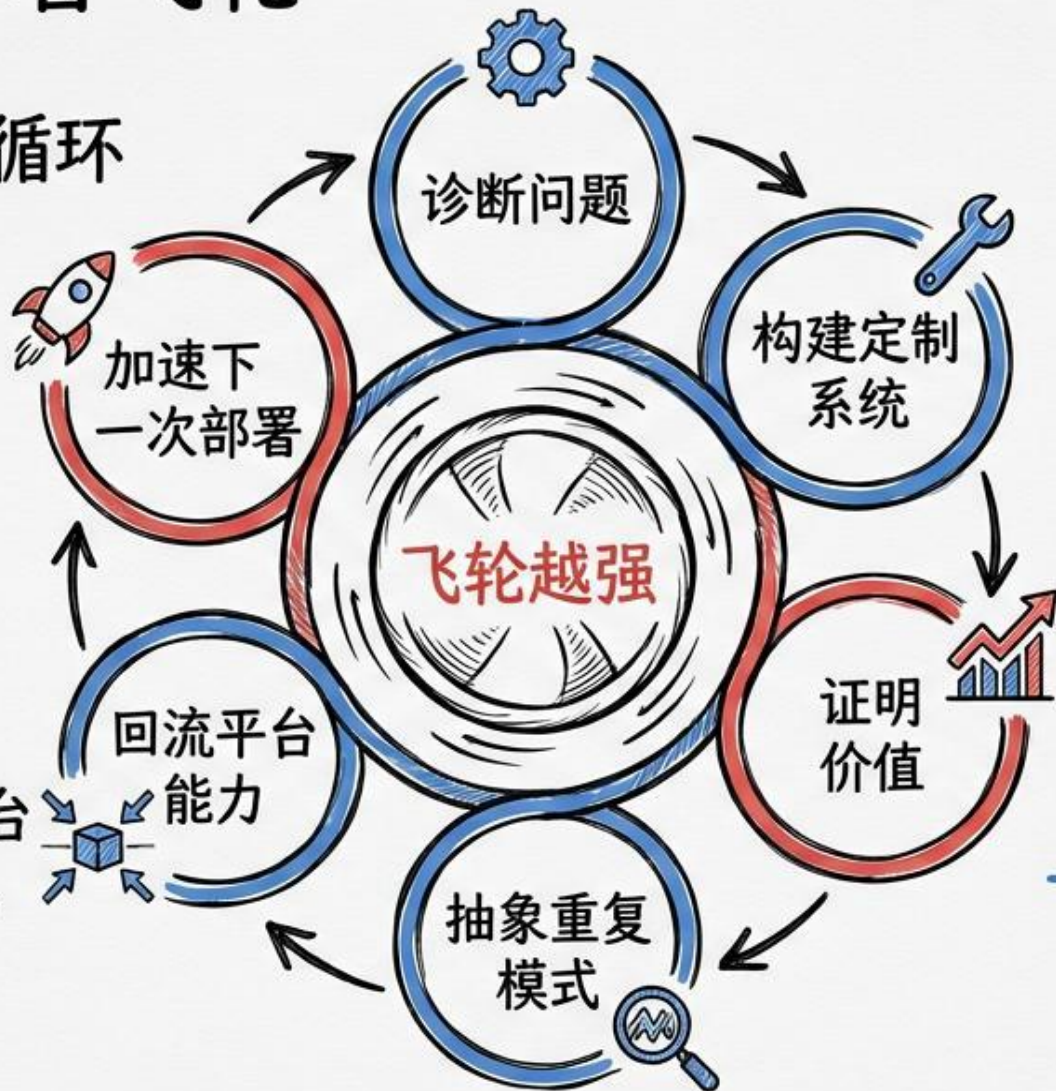
• 把企业约束作为系统设计输入

- 安全、权限、治理、合规和遗留系统必须从第一天进入架构。
- 这不是降低创新速度，而是减少上线阻力。✓
- 复杂行业越需要约束内生化。



原创概念三：部署飞轮

从诊断到产品能力的循环



— 抽象重复模式，回流平台能力，加速下一次部署。

— 诊断问题，构建定制系统，证明价值。

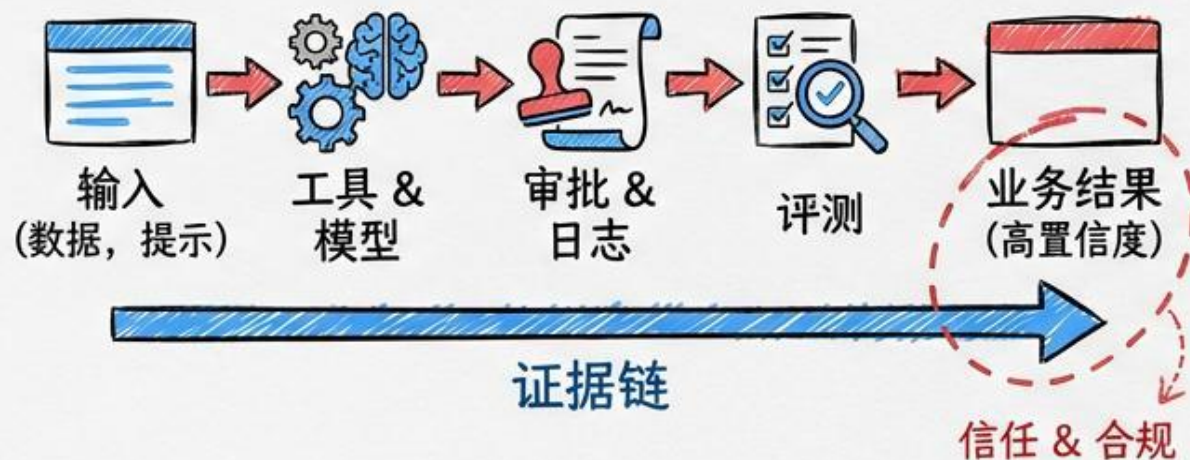
— 飞轮越强，FDE 越不只是项目团队。

原创概念四：可审计价值链

AI 价值必须能被追踪和复盘

- ✓ 价值链连接数据、提示、工具、模型、审批、日志、评测和结果。
- ✓ 任何断点都会削弱客户信任。
- ✓ 可审计价值链是高风险行业部署的前提。

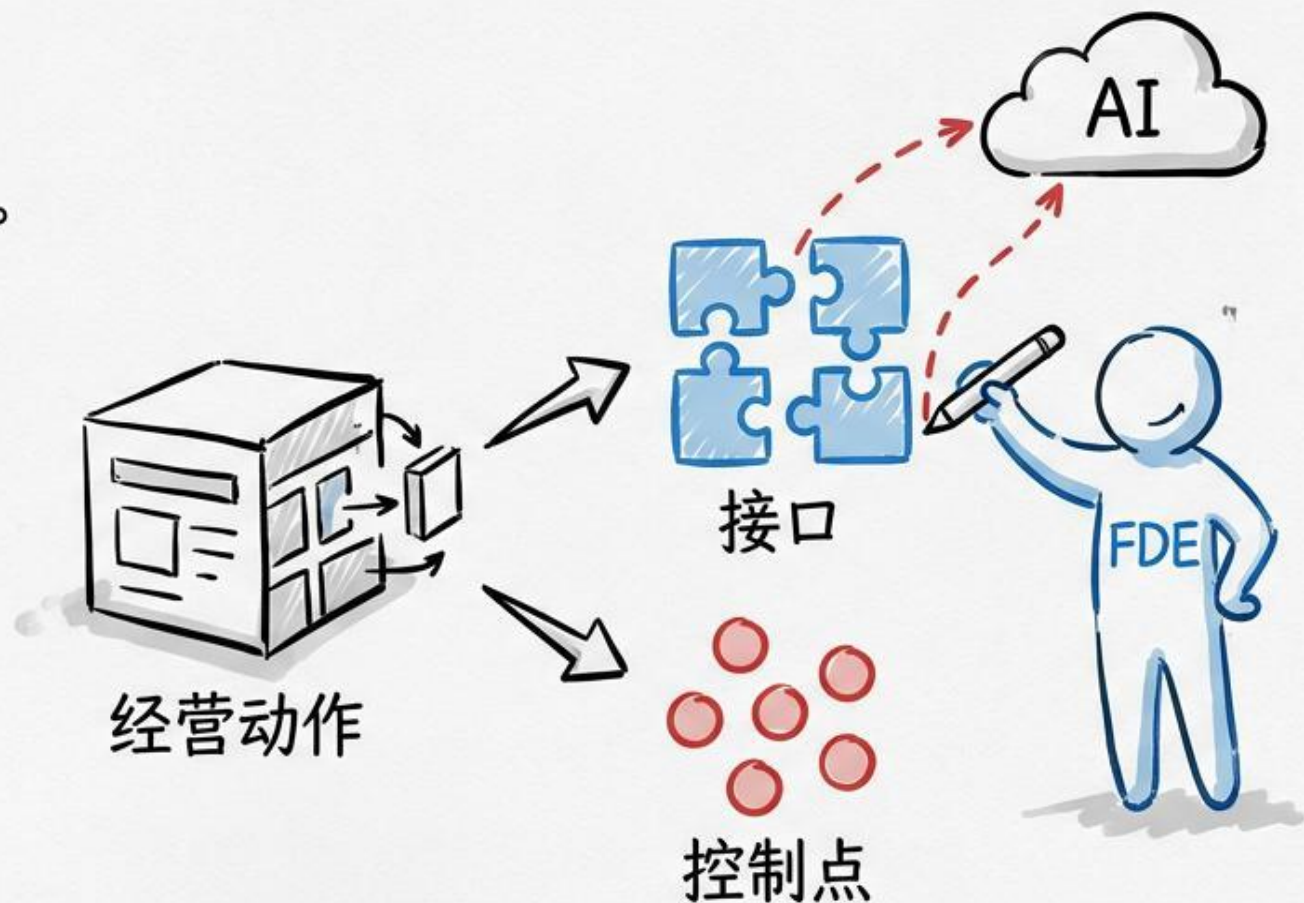
从输入到业务结果的证据链



原创概念五：经营动作接口

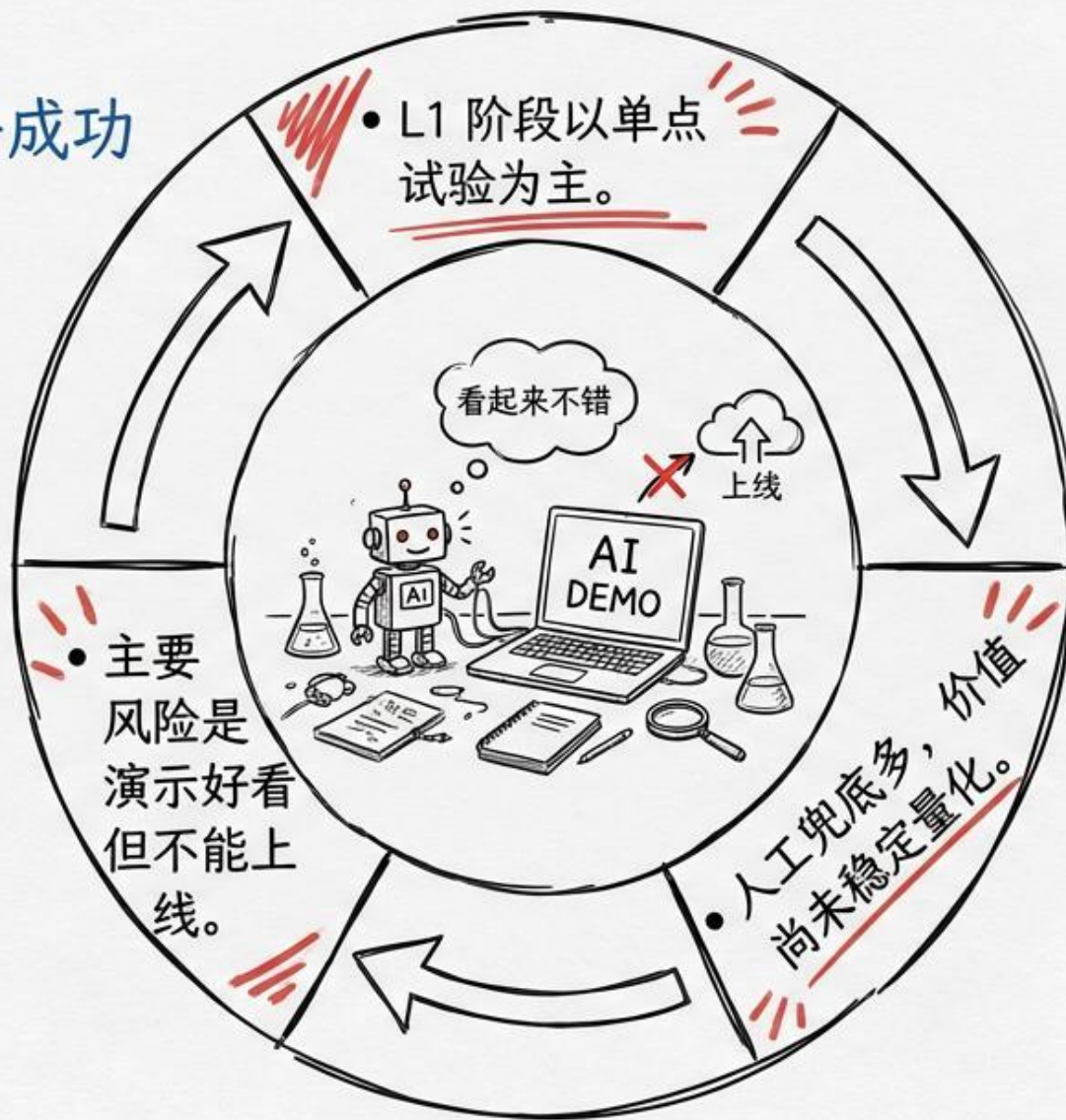
FDE 的对象是企业经营动作

- 企业经营动作需要被重新接口化。
- 模型、工具、人员和控制点要在同一流程内协同。
- FDE 是让 AI 进入经营动作的接口设计者。



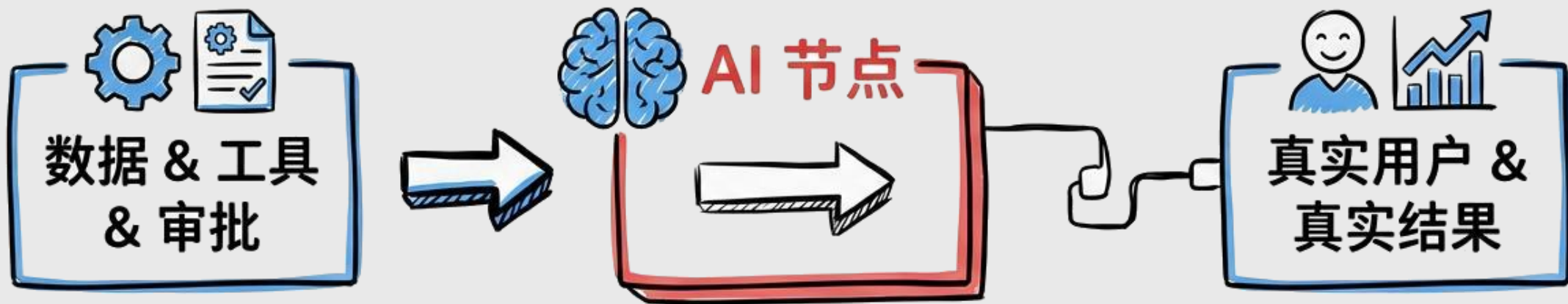
成熟度 L1: 实验型

Demo 成功不等于部署成功



成熟度 L2: 嵌入型

AI 进入一个真实 workflow

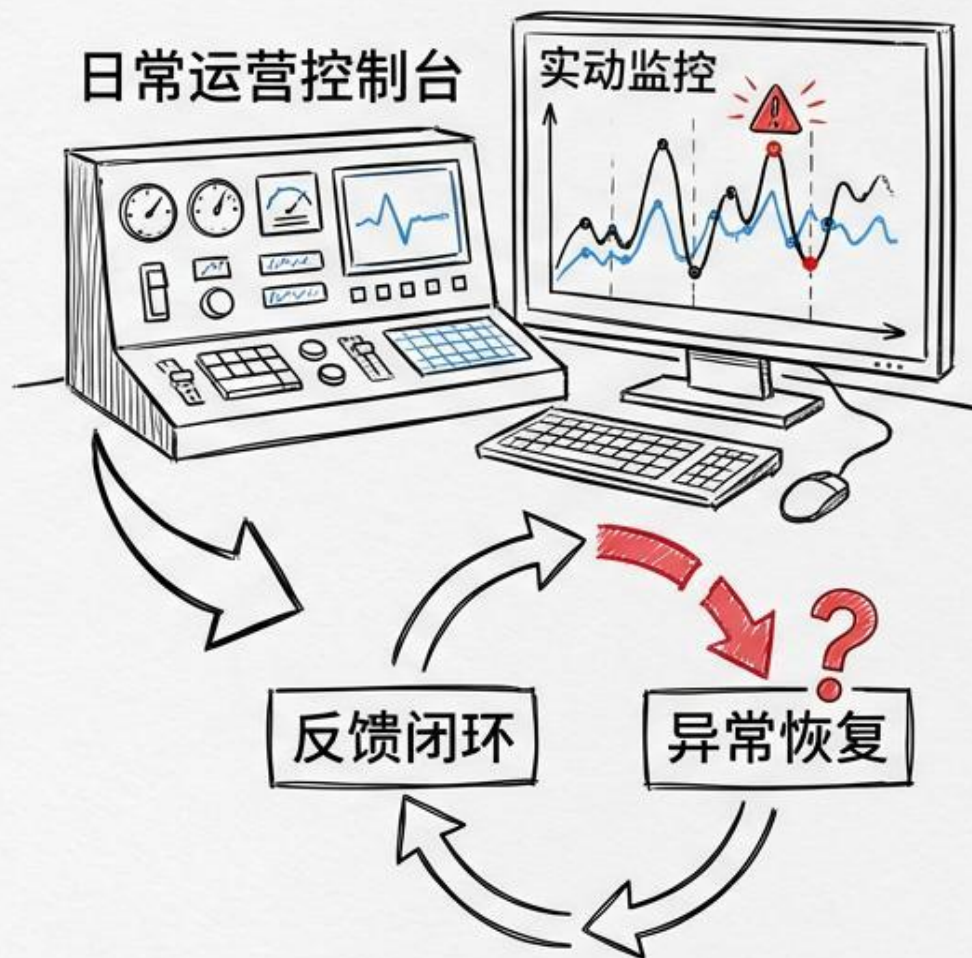


! 主要风险: 权限和责任边界不稳定

成熟度 L3: 运营型

AI 成为日常运营系统的一部分

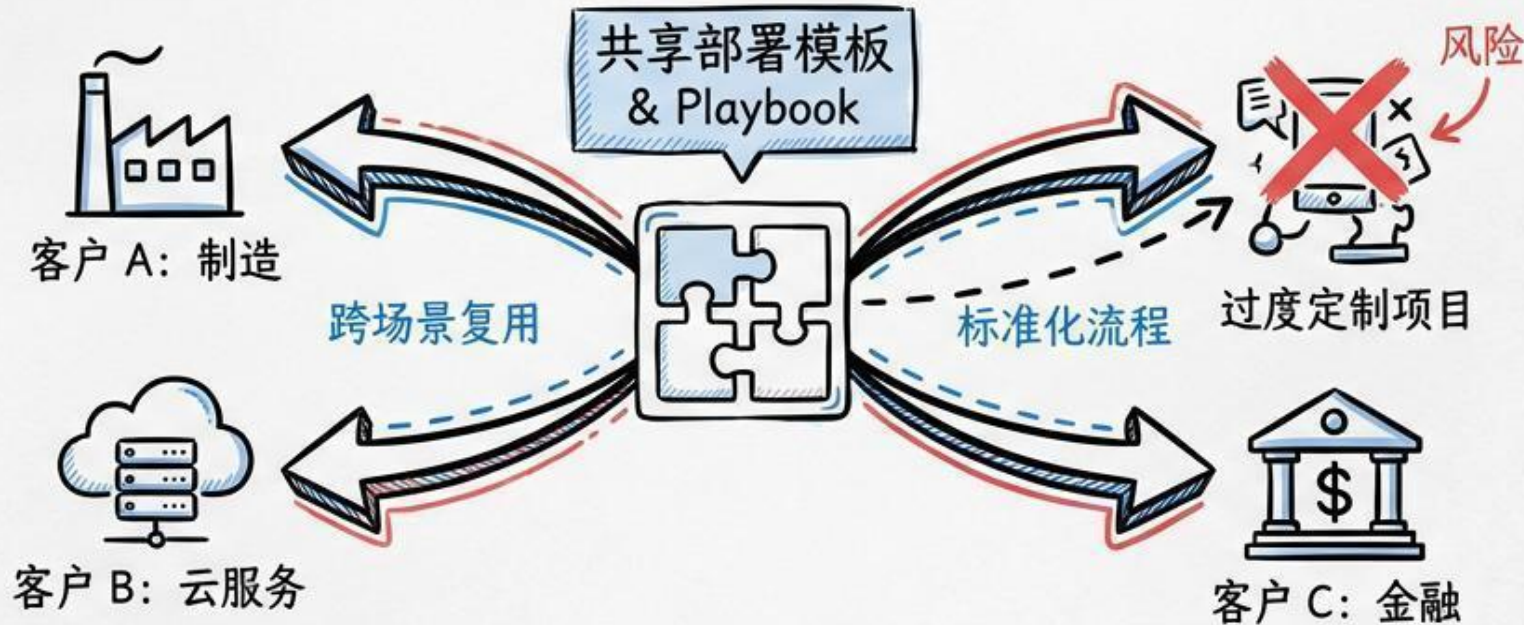
- L3 阶段已有日志、评测、回滚和监控。
- 系统可以在日常流程中运行。
- 主要风险是反馈闭环断裂或异常恢复不足。



成熟度 L4：模式型

部署经验可以跨场景复用

- L4 阶段多个部门或客户出现重复模式。
- 团队能形成 playbook、模板和评测套件。
- 主要风险是 过度定制项目。



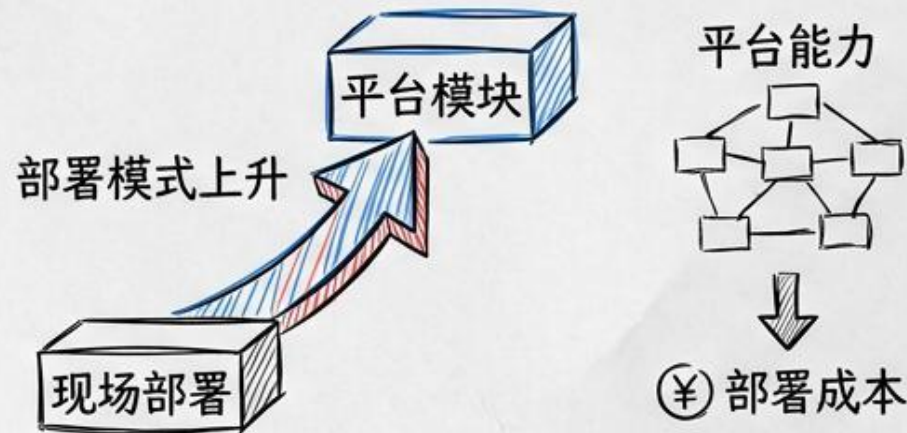
成熟度 L5：产品型

现场模式进入平台产品

✓ L5 阶段部署经验被产品化。

! 主要**风险**是产品抽象失去现场适配能力。

• 平台能力降低下一次部署成本。

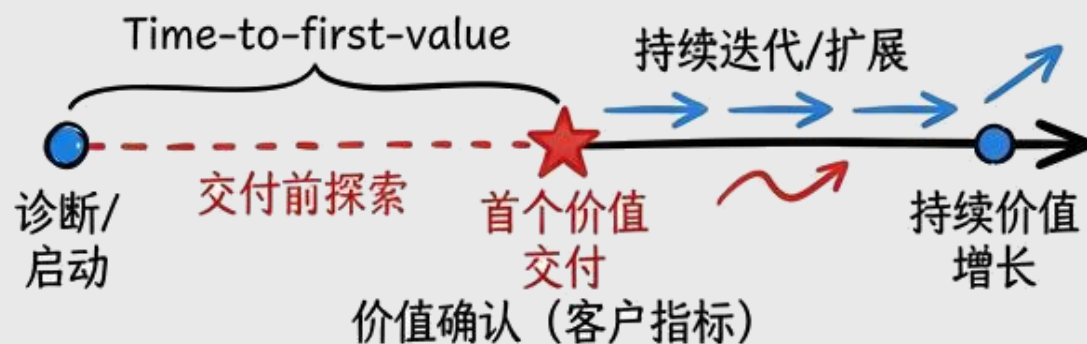


指标一：Time-to-first-value

衡量从诊断到首个价值的速度

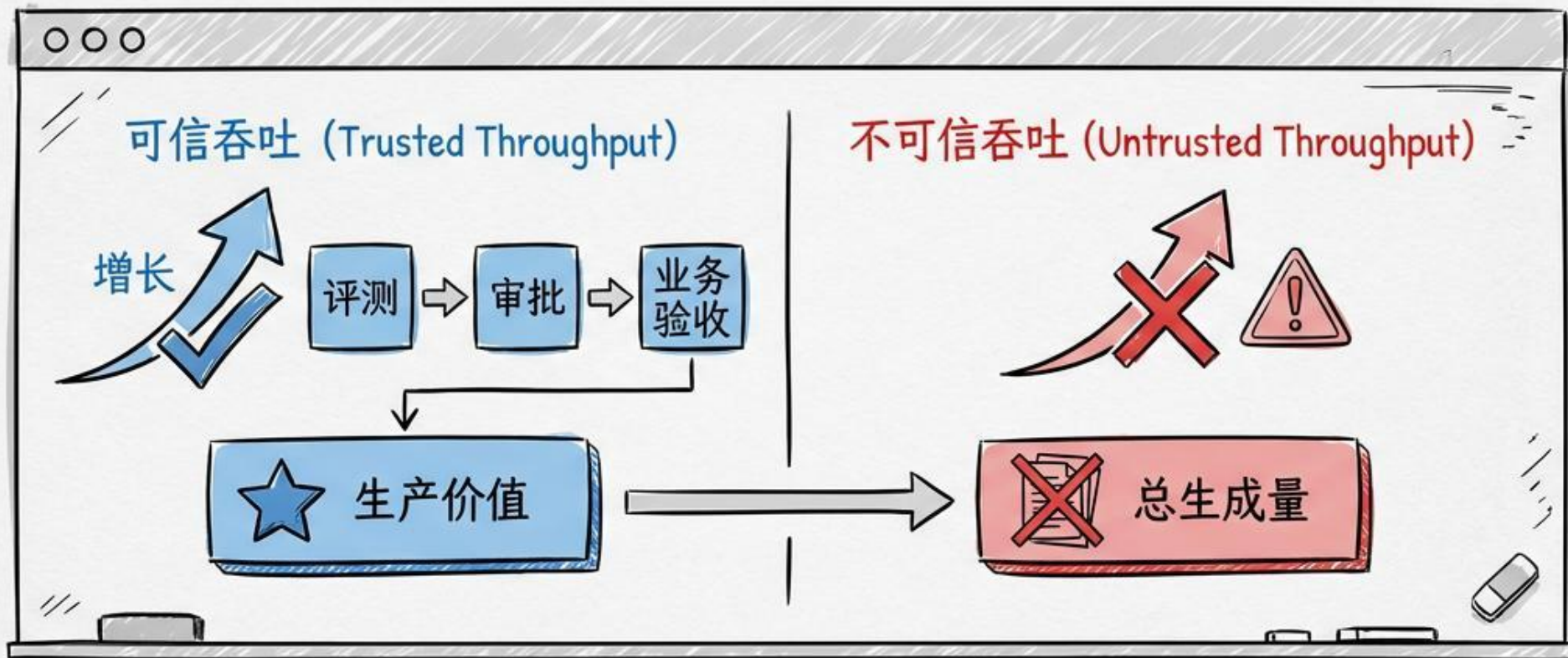
- FDE 追求**早交付**和**快速迭代**。
- 首个价值应由**客户业务指标**确认。
- 速度必须与**风险边界**共同衡量。

首个价值时间轴



指标二：可信吞吐

衡量能被接受的任务完成量



— 总生成量不能代表生产价值。

— 可信吞吐关注通过评测、审批和业务验收的任务量。

— 高风险流程更需要可信而非全自动。

指标三：评测覆盖率

衡量关键场景是否进入测试

- 评测应覆盖常规样本、边界样本和失败样本。
- John Deere 案例显示定制评测对准确性改进重要。
- 没有覆盖率，部署风险无法度量。



评测矩阵		
常规	边界	失败
✓	✓	✗

覆盖不同场景



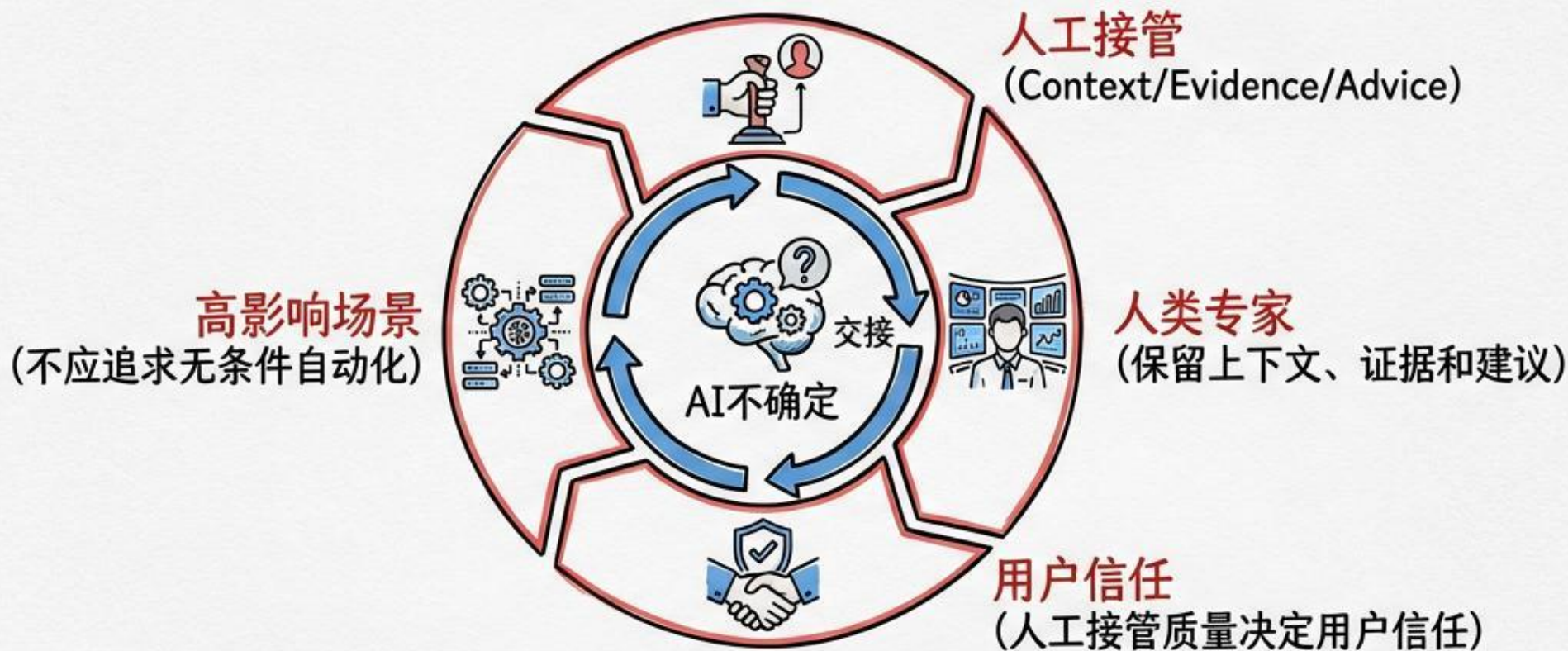
定制评测 → 准确性提升



无覆盖率 = 无法度量风险

指标四：人工接管质量

衡量 AI 不确定时如何交给人



- 高影响场景不应追求无条件自动化。
- 接管必须保留上下文、证据和建议。
- 人工接管质量决定用户信任。

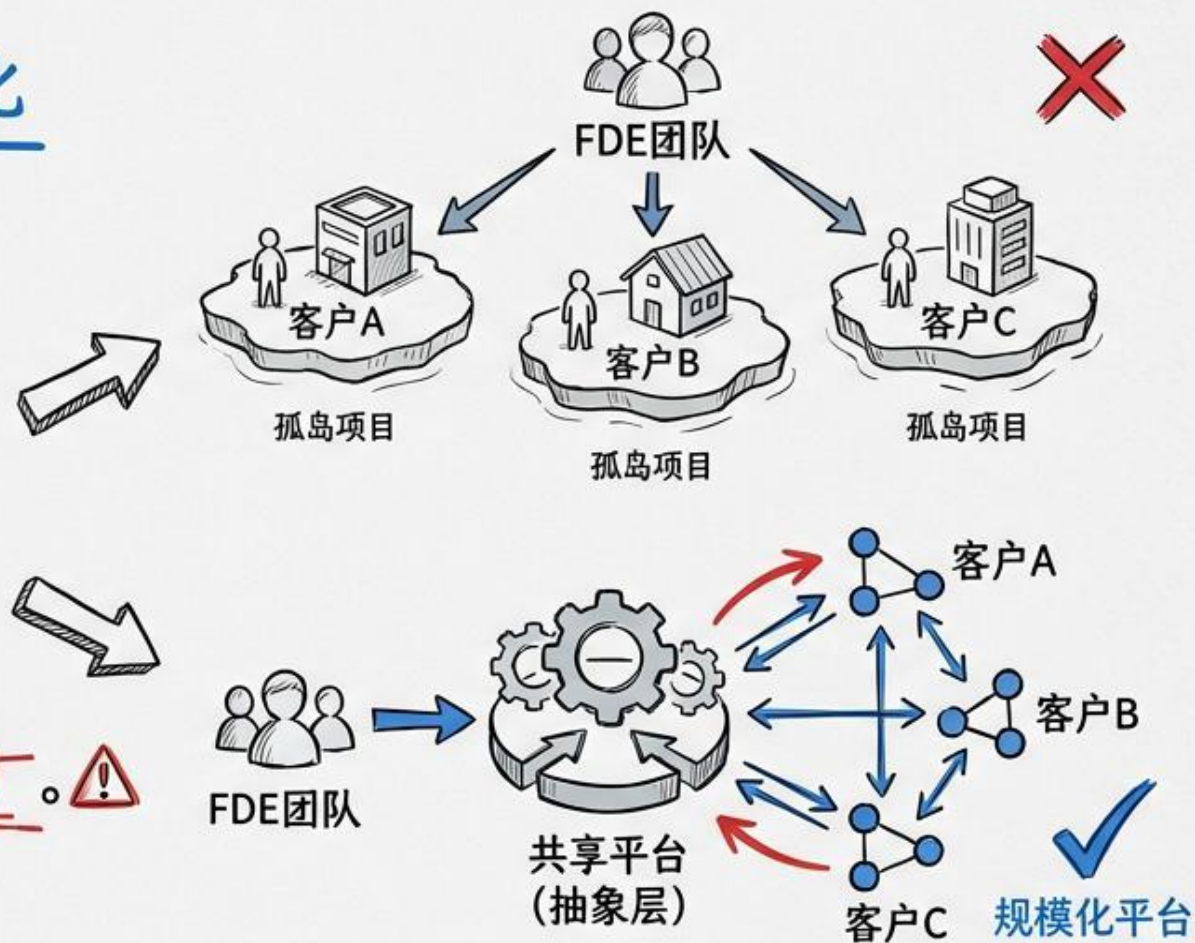
风险一：定制过拟合

每个客户都变成孤岛会伤害规模化

→ FDE 需要定制，但不能停在定制。

→ 重复模式必须及时抽象。

→ 否则部署公司会变成高成本项目工厂。⚠



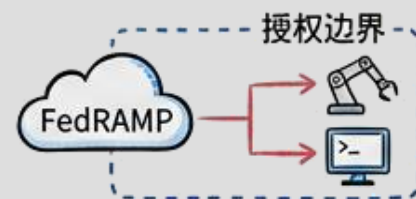
风险二：合规瓶颈

模型能力不能替代授权和审计

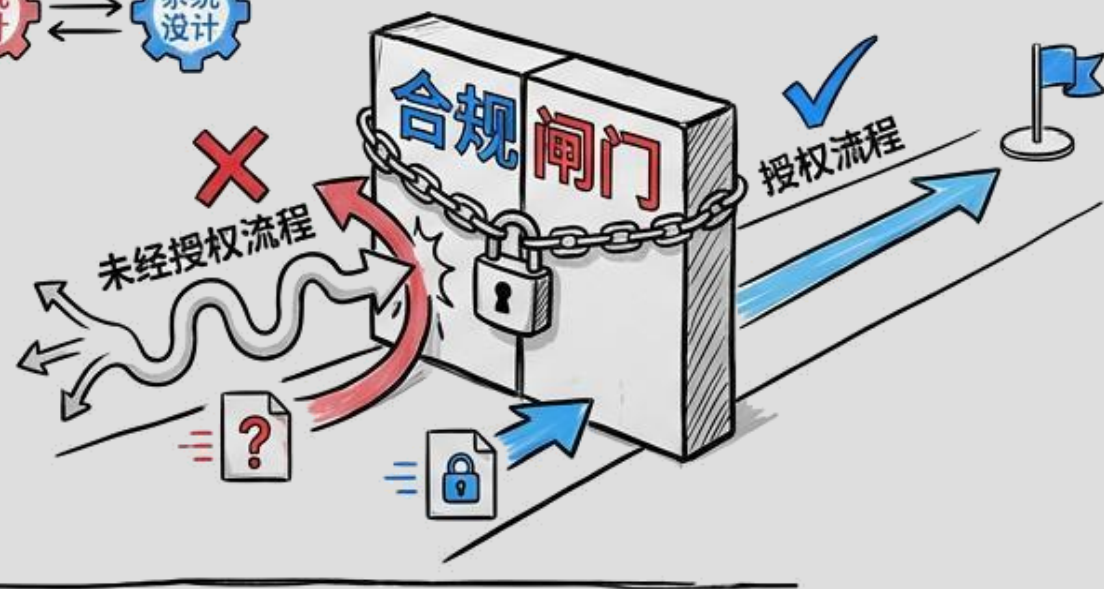
✓ — 政府和金融场景需要**合规证据**。



✓ — FedRAMP 场景说明功能和端点会受**授权边界**影响。

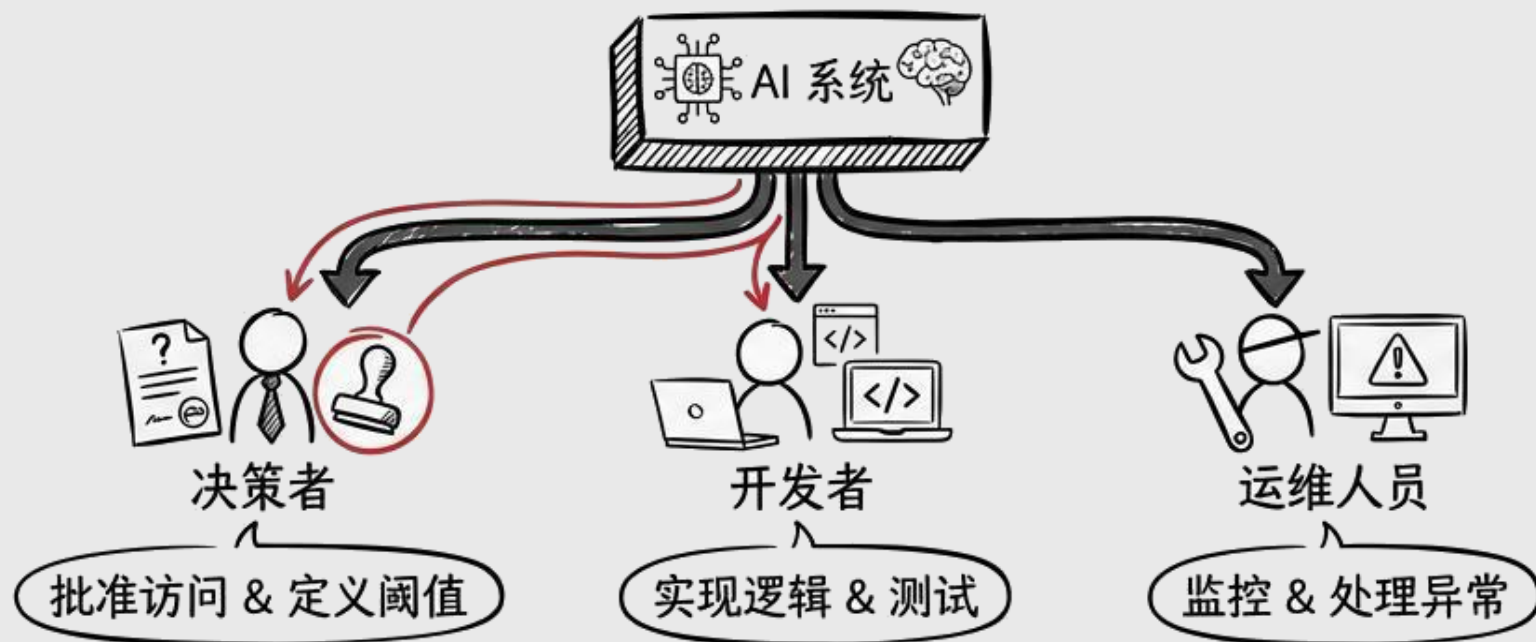


✓ — **合规设计**必须与系统设计同步。



风险三：责任集中

- AI 深入核心流程后，问责会集中到部署决策

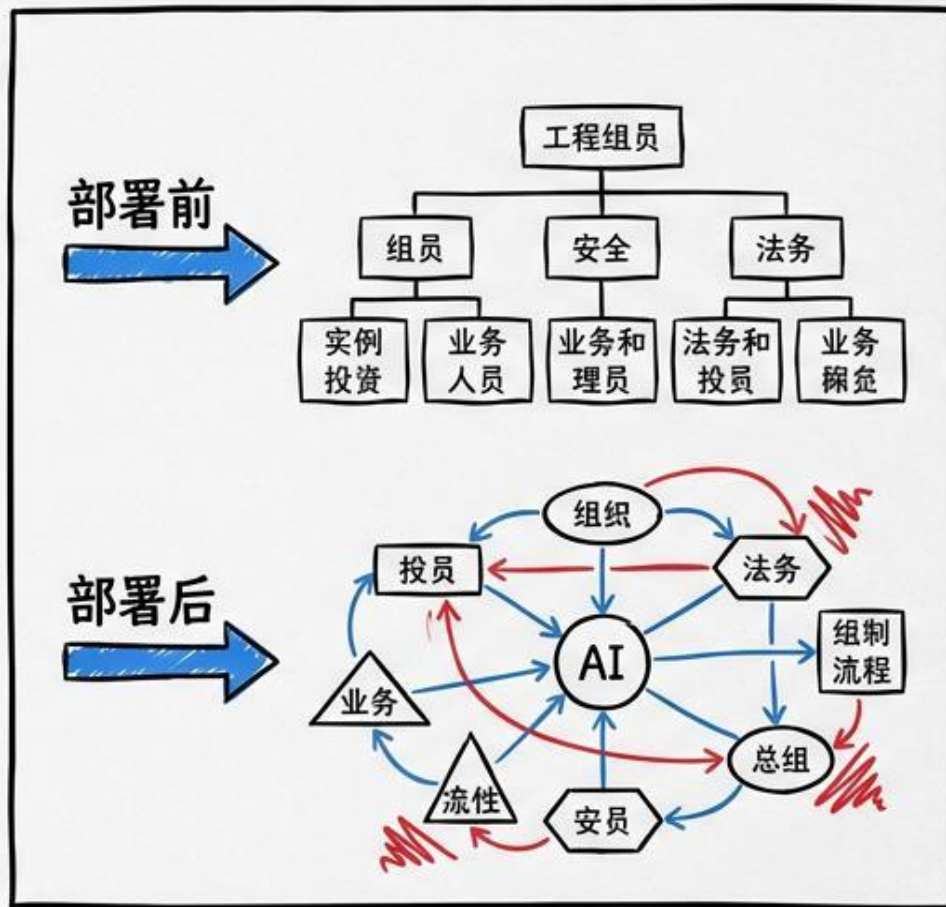


- 谁批准工具访问，谁定义阈值，谁处理异常都需要明确。
- FDE 应提前设计责任链。
- 否则系统 越强，组织风险越高。

风险四：变革阻力

部署 AI 会改变岗位、流程和权力边界

- FDE 不只是工程问题，也是 组织变革 问题。
- 业务、IT、安全、法务和一线人员必须 共同参与。
- 忽视人和流程 会让技术系统无法持续使用。





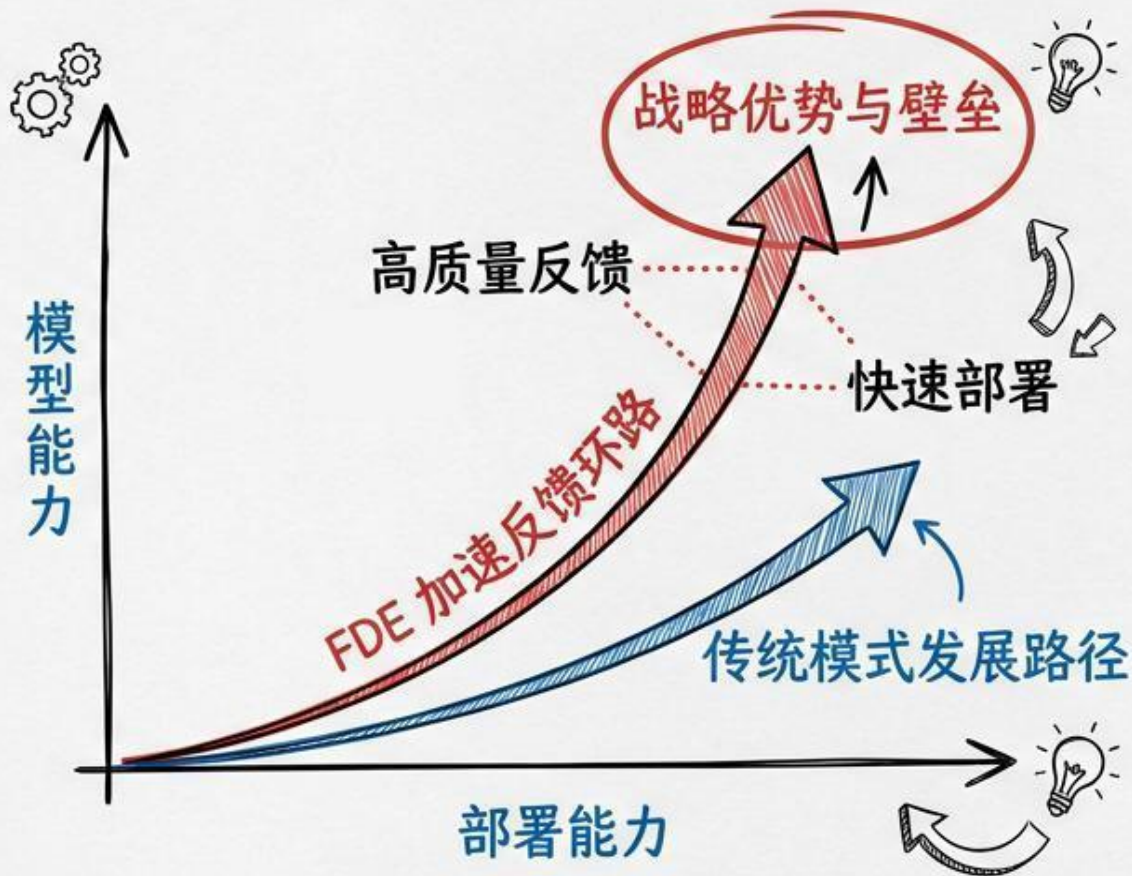
FDE 改变企业 AI 竞争



优势来自部署速度与产品化回流

- 模型领先只是起点。
- 谁能把模型更快带入真实流程，谁能获得更高质量反馈。
- FDE 让部署能力成为 AI 公司的战略壁垒。

竞争曲线：模型能力与部署能力双轴



FDE可能重塑软件交付方式

软件从许可交付转向现场共同构建



许可证交付

👉 企业 AI 不会只靠**自助控制台**完成。

👉 复杂流程需要**现场诊断、构建和迭代**。

👉 FDE让**产品与服务边界**重新融合。

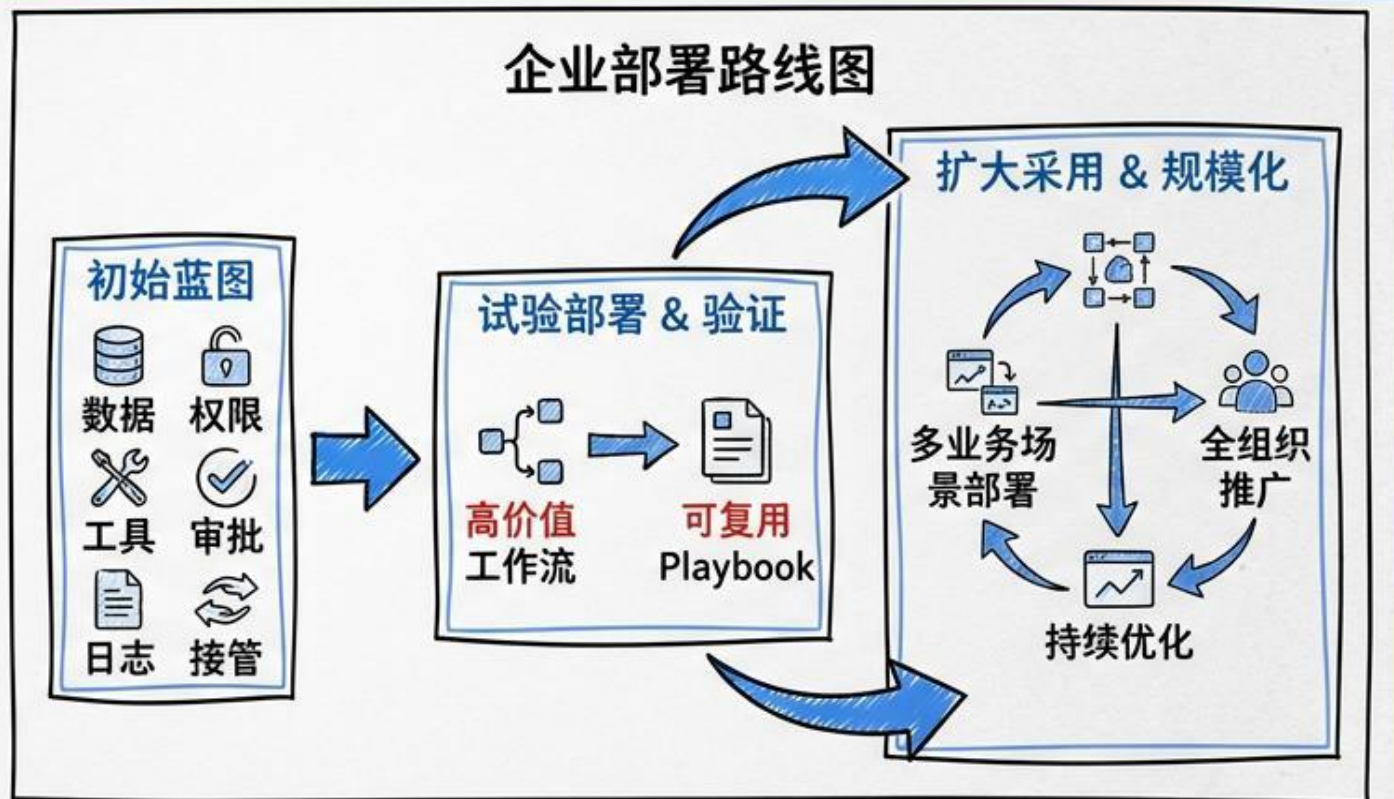


现场共同构建
(FDE)

给企业的三条建议

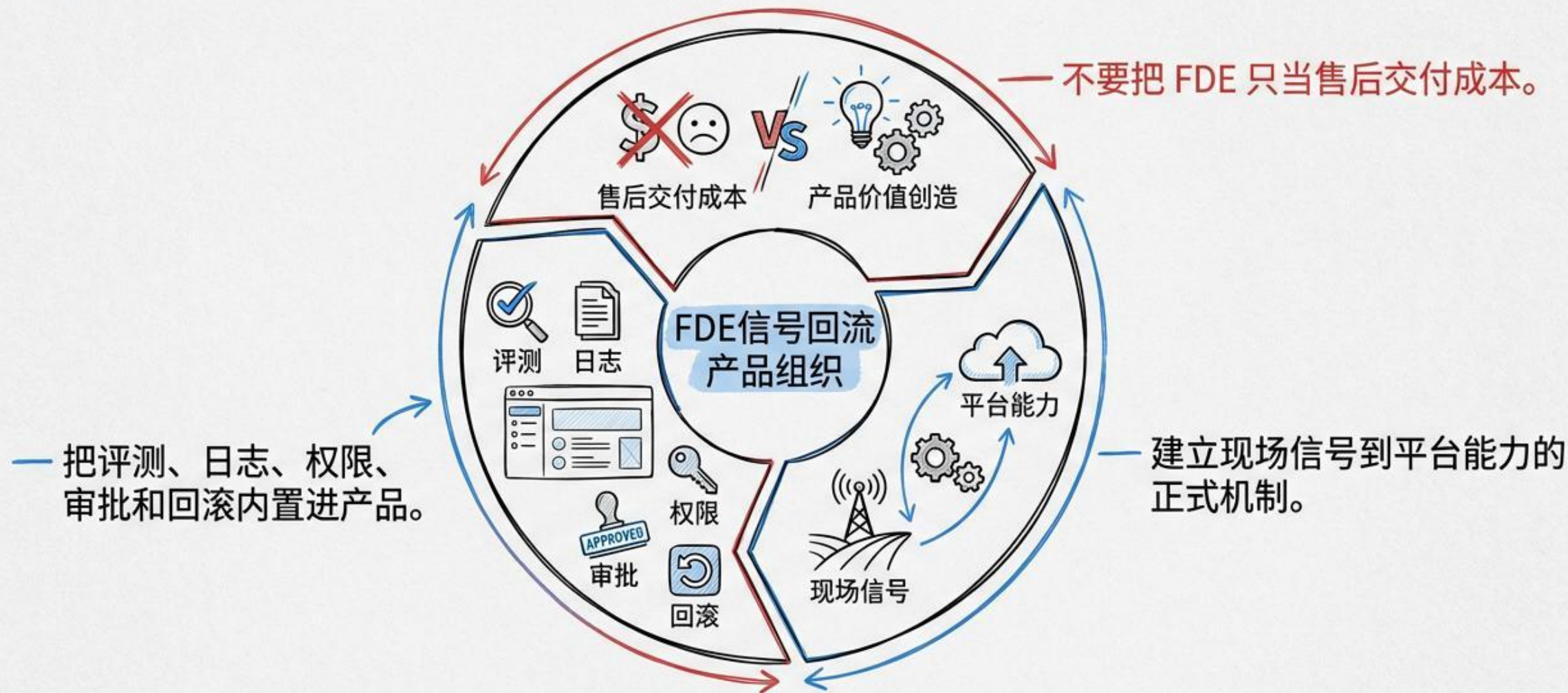
先建部署蓝图，再扩大采用

- 优先选择高价值、可测量、可控风险工作流。
- 把数据、权限、工具、审批、日志和接管写进蓝图。
- 要求每次部署产出可复用 playbook。



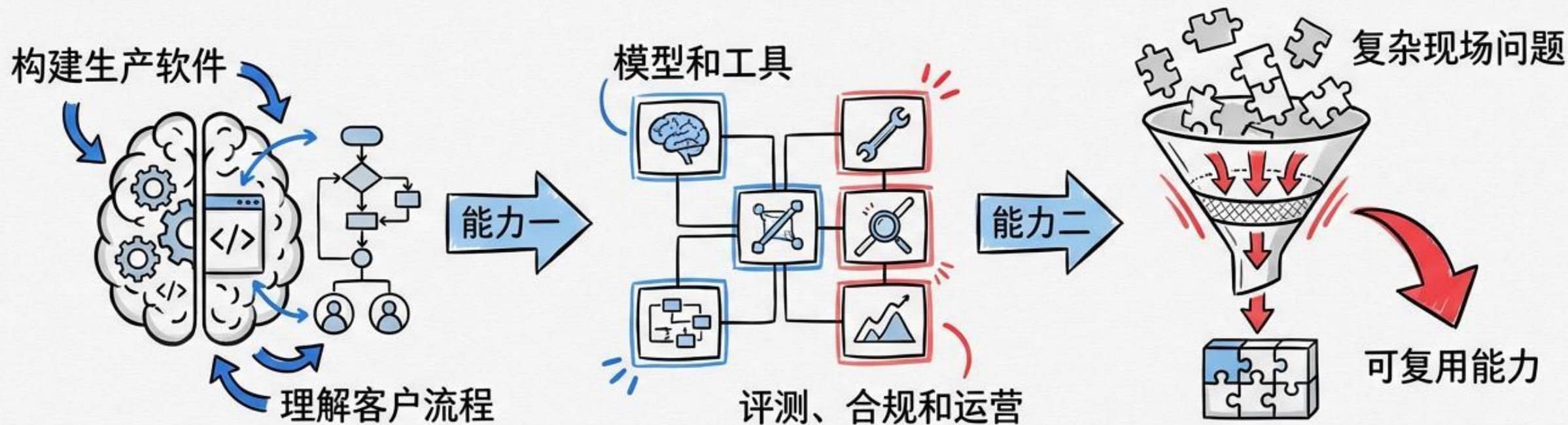
给 AI 厂商的三条建议

把 FDE 视为产品组织的一部分



给人才三条建议

FDE 人才需要多栈能力



FDE人才能力地图

最重要的是把复杂现场问题
抽象成可复用能力。

部署本身就是产品

OpenAI FDE 的关键启示

- 未来企业 AI 的竞争不只是模型榜单。
- 真正的竞争在于谁能把 AI 可靠部署到真实运营。
- FDE 代表从“会用模型”到“围绕智能重构组织”的转变。

