

大众版 V1.3 正式排版版

老板让我负责 AI， 我该怎么办？

普通人也能上手的企业 AI 落地入门课

写给第一次负责 AI 项目的人

找问题 · 选场景 · 做 MVP · 算价值 · 小范围上线 · 复盘成长

目录

序言：这本书写给谁

第一章 老板让我负责 AI，我第一步该怎么办？

第二章 为什么负责 AI 的人，往往不是 AI 工程师？

第三章 别急着做 Agent，先找到真正值得做的事

第四章 什么事情适合交给 AI，什么事情不适合？

第五章 一周做出一个能演示的 AI MVP

第六章 老板问“这个值多少钱”，你怎么回答？

第七章 Demo 很好，为什么上线就没人用了？

第八章 从第一个 AI 项目开始，成为 AI 落地工程师

附录：你的第一个 AI 项目工具包

版本说明

序言：这本书写给谁

也许你并不是 AI 工程师。

你没有训练过大模型，也没有研究过深度学习。你可能只是一个产品经理、运营负责人、行政、人事、销售主管、IT 同事，或者一家中小企业的老板。

但某一天，老板在会上说：

“AI 最近这么火，我们公司也要研究一下。你先负责看看。”

于是，这件事落到了你身上。

你点点头，打开电脑，开始搜索“企业 AI 怎么落地”。很快，你看到一堆陌生词：大模型、智能体、Agent、RAG、知识库、工作流、插件、API、私有化部署、低代码平台……

信息越看越多，人却越来越慌。

你不是不知道 AI 很重要。你真正不知道的是：

我到底应该从哪一步开始？

这本书就是写给这个时刻的你。

这不是一本 AI 技术教程

市面上已经有很多内容在教你怎么写 Prompt、怎么调用模型 API、怎么搭建知识库、如何使用某个 Agent 平台、怎么比较不同大模型。

这些内容当然有用，但它们通常默认你已经知道自己要做什么。

可对大多数第一次负责 AI 项目的人来说，最大的问题并不是“怎么调用模型”，而是：

公司里到底哪个场景值得先做？这个场景适不适合用 AI？第一个 Demo 要做到什么程度？怎么判断它有没有价值？老板问“能省多少钱”时，怎么回答？上线以后没人用怎么办？出错了谁负责？

这些问题，才是企业 AI 落地真正绕不开的地方。

所以，这本书不会一开始就带你研究复杂技术。我们会先从最朴素的问题开始：

公司里有一件重复、耗时、让人头疼的工作，AI 能不能先帮一点忙？

只要能回答这个问题，你就已经走上了正确的路。

图1 企业 AI 落地最小路径图

先跑通一个小闭环，再逐步复制



不要一上来做大战略，先完成第一个可验证的小项目。

这张图就是本书的路线。先跑通一个小闭环，再谈更大的系统和战略。

全书贯穿的四个案例

为了让你更容易把方法放回真实工作里，V1.2 会反复使用四个小案例。它们不复杂，但足够典型。

案例线	业务部门	第一个小场景	主要帮助你理解什么
客服邮件分类	客服 / 销售支持	把客户邮件分成投诉、咨询、售后、合作	如何做分类类 MVP

案例线	业务部门	第一个小场景	主要帮助你理解什么
合同字段提取	法务 / 采购	从合同里提取金额、税率、付款方式、交付日期	AI 能做什么、不能做什么
报销单异常检查	财务	检查发票抬头、金额、日期、预算科目	规则、人工复核和 ROI
巡店报告生成	零售 / 门店管理	根据巡店照片和检查记录生成报告草稿	草稿生成、效率提升和上线

后面每章都会尽量回到这四类场景，而不是不断换新例子。这样你读完后，脑子里会有几条完整路径，而不是一堆散点。

这本书也不是写给世界 500 强 CTO 的

大型企业有自己的数字化团队、数据团队、安全团队、法务合规团队，也有充足预算去搭建 AI 平台和企业级系统。

但更多公司不是这样。

更多公司的现实是：没有专门 AI 团队，没有清晰 AI 战略，没有现成数据平台，没有充足预算，甚至连“谁负责 AI”都还没想清楚。

很多时候，就是某个部门里相对懂业务、愿意研究新东西的人，被推到了前面。

这本书写给的人，正是这样的人。

你不需要一开始就懂所有技术。你需要先学会做几件事：找到真实业务问题，判断它是否适合 AI，用最低成本做一个小验证，用数据说明它是否值得继续，上线时控制风险，并从第一个项目里积累经验。

这就是企业 AI 落地的起点。

你会在这本书里完成一次小小的成长

第一章，我们先解决最现实的问题：老板让我负责 AI，我第一步该怎么办？

第二章，我们讨论为什么负责 AI 的人，往往不是 AI 工程师。

第三章，你会开始学会找场景。

第四章，你会判断什么事情适合交给 AI，什么事情不适合。

第五章，你会用一周时间做出一个能演示的 AI MVP。

第六章，你会学会回答老板最关心的问题：这个东西到底值多少钱？

第七章，我们会讨论为什么很多 Demo 很好，上线后却没人用。

第八章，我们会把前面的内容串起来，告诉你如何从第一个小项目开始，成长为 AI 落地工程师。

读完这本书，你不会马上变成 AI 专家。但你会知道如何启动第一个 AI 项目。

这已经足够重要。

因为大多数人不是输在不会高深技术，而是输在第一步就走错了。

先做成一个小项目，比空谈 AI 战略更重要

企业 AI 落地，最怕一上来就讲得太大。

“我们要建设企业智能中台。”

“我们要打造公司级 AI 助手。”

“我们要让 AI 改造所有流程。”

这些目标听起来很振奋，但对第一次负责 AI 的人来说，往往太远。

真正可行的起点，通常很小：让 AI 帮客服先分一下邮件，让 AI 从合同里先提取几个关键字段，让 AI 帮财务先检查报销单有没有明显异常，让 AI 帮 HR 先从简历里提取学历、年限和技能，让 AI 帮区域经理先生成一版巡店报告草稿。

这些事情看起来不宏大，但它们真实、具体、容易验证。

而 AI 落地，正是从这些小事开始的。

这本书的核心观点

如果只用一句话概括这本书，我希望你记住：

AI 项目的第一步，不是研究工具，而是找到值得解决的业务问题。

工具会变，模型会变，平台会变。

但企业真正关心的事情不会变：能不能节省时间，能不能减少错误，能不能提高速度，能不能降低成本，能不能让员工少做重复劳动，能不能安全、可控地进入真实流程。

如果一个 AI 项目回答不了这些问题，再先进也只是演示。

如果一个小项目能回答这些问题，即使很简单，也已经开始创造价值。

现在，我们从最常见的场景开始。

老板让你负责 AI，你第一步到底该干什么？

第一章 老板让我负责 AI，我第一步该怎么办？

周一早会快结束时，老板突然看向你。

他说：

“最近 AI 很火，我们公司也得研究一下。你先负责看看，月底给我一个方案。”

你点点头，表面很镇定。

但回到工位后，真正的问题来了。

你打开搜索框，输入“企业 AI 怎么落地”。页面一下子跳出很多内容：大模型、智能体、Agent、知识库、RAG、工作流、自动化办公、私有化部署、低代码平台、企业 AI 助手……

你又打开几个 AI 工具试了试。它们确实很厉害。能写文案，能总结会议纪要，能分析表格，还能回答很多问题。

可问题是，你所在的公司不是发布会现场，也不是演示视频里的理想环境。

公司里有格式混乱的 Excel 表，有没人愿意维护的知识库，有各种版本的合同模板，有流程复杂的审批，有说不清楚但每天都在发生的重复工作，还有等着看结果的老板。

于是你很快发现，真正困难的不是“AI 能不能用”。真正困难的是：

公司到底应该从哪里开始用 AI？

这也是大多数人第一次负责 AI 项目时遇到的困境。

不是不努力，而是不知道第一步该往哪里走。

不要一上来就问“用哪个 AI 工具”

很多人接到 AI 任务后的第一反应，是先去研究工具。

今天看模型排行榜，明天试一个 Agent 平台，后天收藏一堆 Prompt 教程，再过几天，整理出一份工具清单。

这很正常。

因为工具最容易被看到。你可以立刻注册账号，立刻输入问题，立刻得到结果。这个过程反馈很快，也很容易让人产生一种感觉：

我已经开始做 AI 了。

但企业 AI 落地最容易出问题的地方，也正在这里。

工具不是起点。

业务问题才是起点。

如果你还不知道公司里哪个流程最浪费时间，哪个部门最痛苦，哪个环节最容易出错，那么过早研究工具，很容易变成“拿着锤子找钉子”。

最后的结果往往是：你做了一个看起来不错的 Demo，老板演示时觉得很新鲜，业务部门试用两天后，却发现它解决的不是自己的核心问题。

项目就这样停在了演示阶段。

很多 AI 项目失败，并不是因为模型不够聪明，而是因为一开始就没有选对问题。

也不要一开始就做“大而全”的 AI 系统

还有一种常见误区，是一上来就想做一个完整系统。

比如：

“我们要做一个公司级 AI 助手。”

“我们要做一个企业知识大脑。”

“我们要让 AI 自动处理所有客服问题。”

“我们要把采购、销售、财务、人事全部智能化。”

这些目标听起来很有想象力。但对第一次负责 AI 项目的人来说，往往太大了。

目标越大，牵涉的人越多。牵涉的人越多，问题就越复杂。

你会很快遇到一连串现实问题：数据在哪里？谁有权限？哪个系统能接？出错谁负责？业务部门愿不愿意配合？IT 部门是否允许接入？法务和合规有没有意见？预算谁批？验收标准是什么？

如果这些问题没有提前想清楚，一个“大而全”的 AI 项目很容易变成一个长期拖延的工程。

开会很多，演示很多，真正上线很少。

所以，第一次做 AI，不要从“大系统”开始。要从小而具体的工作开始。

不是：

“让 AI 改造整个财务部。”

而是：

“先让 AI 帮财务检查报销单里的发票抬头、金额和预算科目是否一致。”

不是：

“让 AI 接管所有客服。”

而是：

“先让 AI 把客户留言分成投诉、咨询、售后三类。”

不是：

“做一个企业知识大脑。”

而是：

“先让 AI 回答销售团队最常问的 20 个产品问题。”

小，不代表价值低。

小，意味着问题更清楚，结果更容易验证，风险也更可控。

更不要一开始就追求“全自动”

很多人对 AI 有一个误解：既然是 AI，就应该自动完成全部工作。

这很危险。

企业里的很多工作，不只是“做完”那么简单，还涉及责任、风险和判断。

比如合同终审、财务付款、客户退款、员工录用、法律意见、医疗建议。这些事情一旦出错，就不是“回答错了”那么简单，而是真正可能造成损失。

第一次做 AI 项目时，不要急着让 AI 自动执行。

更稳妥的做法是：

先让 AI 辅助人，而不是替代人。

它可以先做信息提取，可以先做初步分类，可以先生成草稿，可以先标记风险，可以先给出建议。

但最后的关键动作，仍然由人确认。

这不是保守。

这是专业。

一个能安全上线的 AI 项目，通常不是一开始就全自动，而是一步步推进。

第一步，只读。AI 只负责看资料、提取信息、生成摘要。

第二步，草稿。AI 可以生成邮件、报告、审核意见，但必须人工确认。

第三步，半自动。AI 可以发起待办、填写表单、创建流程，但关键节点保留人工审批。

这样做的好处是，业务部门更容易接受，风险也更容易控制。

你不是在对业务人员说：

“以后你不用干了。”

而是在说：

“重复、费时间、容易出错的部分，先让 AI 帮你处理。你来做最后判断。”

这句话，业务部门更愿意听。

正确的第一步：先研究工作，而不是先研究 AI

那么，第一步到底该做什么？

答案很简单：

不要先研究 AI。先研究工作。

你需要走进业务现场，看看大家每天到底在做什么。

哪些工作每天都在重复？哪些工作最耗时间？哪些工作经常出错？哪些工作让人觉得烦，但又不得不做？

哪些工作不需要特别深的专业判断，却需要处理大量信息？

这些地方，往往就是 AI 项目的入口。

贯穿案例：报销单异常检查

财务同事每月要处理大量报销单，其中很多时间花在检查发票抬头、金额、日期和预算科目上。这个场景不需要 AI 自动批准报销，只需要先帮财务标记明显异常。它真实、高频、规则相对清楚，也有人可以复核，所以很适合作为第一个候选场景。

比如：销售团队每天要整理客户邮件，财务每个月要审核大量报销单，HR 要从几十份简历里筛出候选人，采购要对比供应商报价，法务要检查合同里有没有漏填关键条款，区域经理要根据巡店照片写报告。

这些工作有一个共同特点：它们不是最复杂的工作，但足够高频；它们不是完全没有规则，但输入形式经常变化；它们不是完全不需要人，但有一部分步骤可以先交给 AI 做。

这类场景，最适合作为第一个 AI 项目。

因为它们容易验证，也容易让业务方看到价值。

一个好问题，胜过一个好模型

第一次负责 AI 项目的人，最容易高估模型，低估问题。

但企业真正买单的，不是模型有多厉害，而是它解决了什么问题。

同样是一个大模型，你让它“做一个公司智能助手”，它可能很难落地。但你让它“从客户邮件中识别投诉、咨询和售后问题，并生成处理建议”，它就变得具体很多。

你让它“帮助财务提升效率”，这个目标太空。但你让它“检查报销单中的发票抬头、金额、日期和预算科目是否匹配”，它就能开始发挥作用。

你让它“辅助销售”，范围太大。但你让它“根据客户背景和历史记录，生成一版跟进邮件草稿”，就更容易验证。

所以，AI 项目的第一步，不是找到最强模型。

而是找到一个足够具体的问题。

越具体，越容易做。越具体，越容易评估。越具体，越容易让老板看到结果。

第一个 AI 项目，应该满足五个条件

图2 第一个 AI 项目五条件

先挑容易验证的小场景



第一个项目不求大，先求真、先求小、先求能验证。

第一个项目不需要宏大，但最好同时满足这五个条件。缺少其中任何一项，都要谨慎推进。

如果你现在就要选一个场景，可以用下面五个条件判断。

第一，它要足够真实。

不要从想象中的需求开始。要从已经发生的工作开始，比如真实客户邮件、真实合同、真实报销单、真实客服记录、真实巡店报告、真实简历。

AI 项目不能建立在“我们觉得业务可能需要”上，而要建立在“业务每天真的在做”上。

第二，它要足够高频。

一个月只发生一两次的事情，不适合作为第一个项目。更适合的，是每天、每周、每月大量重复出现的任务。

高频，意味着节省一点点时间，也能积累出明显价值。

第三，它要足够耗时。

如果一个任务原本只需要 10 秒，哪怕 AI 做得再好，也很难产生明显价值。

你要找的是那些每次看起来不长，但累计起来很浪费时间的事情。

第四，它要有规则。

AI 不是魔法。如果连业务人员都说不清楚判断标准，AI 也很难稳定完成。

适合 AI 的场景，通常有相对明确的规则。比如合同必须包含金额、税率、交付日期、付款方式；发票抬头必须与报销主体一致；客户邮件可以分成投诉、咨询、售后、合作几类。

规则越清楚，AI 越容易做出稳定结果。

第五，它要有人愿意复核。

第一个 AI 项目不要追求无人值守。你需要一个业务方负责人，愿意提供样本，愿意试用结果，也愿意告诉你哪里错了。

没有业务人员参与，AI 项目很难成功。

这个人，就是你的第一个业务伙伴。

你第一周真正该做的事

如果老板今天让你负责 AI，你第一周不要急着提交一份宏大的方案。

更好的做法，是完成四件小事。

第一，找三个业务部门聊一聊。

不用开正式大会。你可以找销售、财务、人事、客服、采购、运营中的几个人，问他们：最近最烦的重复工作是什么？每个月大概要做多少次？每次大概要花多久？有没有固定规则？如果让 AI 先帮你做一部分，你最希望它帮哪一步？

这一步的目标不是立刻定方案，而是找到真实痛点。

第二，收集 10 到 20 个真实样本。

不要只听描述。一定要看真实材料。

AI 项目不能建立在想象上。真实样本会告诉你，业务到底有多复杂。

如果样本里涉及客户隐私、员工信息或商业敏感内容，要先做脱敏处理，不要随意把敏感数据上传到不确定的工具里。

第三，选一个最小场景。

不要选“整个流程”，只选其中一步。

不是“自动审核合同”，而是“提取合同中的关键字段”。不是“自动处理客服”，而是“先判断客户问题类型”。不是“自动写巡店报告”，而是“先根据照片生成检查项草稿”。

第一个场景越小，越容易做成。

第四，做一页纸方案。

这页纸不需要复杂，只要讲清楚：要解决什么问题，现在谁在做，每月发生多少次，每次耗时多久，AI 先帮哪一步，成功标准是什么，需要谁配合。

这比一份几十页的 AI 战略 PPT 更有价值。

因为它能真正推动项目开始。

这一章读完，你可以马上做什么

1. 找一个业务同事聊 20 分钟，问他最近最烦的重复工作是什么。
2. 收集 10 到 20 个真实样本，比如邮件、合同、报销单或巡店报告。
3. 用“一页纸启动模板”写下一个候选场景，不要超过一页。

本章小结

当老板让你负责 AI 时，第一反应不要是慌，也不要立刻去研究工具。

你真正要做的，是先把问题找清楚。

AI 项目的起点，不是模型，不是平台，也不是 Prompt，而是一个真实、重复、高频、有人愿意配合的业务问题。

第一次做 AI，不要追求大而全。先找一个小场景，先做一个小验证，先让业务看到一次具体改善。

这才是最稳的开始方式。

下一章，我们要解决另一个问题：

可我不是 AI 工程师，我真的能负责这件事吗？

第二章 为什么负责 AI 的人，往往不是 AI 工程师？

上一章，我们说到一个关键判断：第一次做 AI，不要先研究工具，要先找到业务问题。

但很多人看到这里，心里还是会有一个疑问：

可我不是 AI 工程师，我真的能负责这件事吗？

这个问题非常真实。

你可能是产品经理、运营负责人、行政、人事、销售主管、财务经理、IT 同事，甚至是一个中小企业老板。

你会用 AI 写东西，也试过几个工具。但你没有训练过大模型，也不懂深度学习，更没写过复杂的 Agent 系统。

所以当老板说“AI 这件事你来负责一下”时，你第一反应可能不是兴奋，而是心虚。

你会想：这不是技术部门的事吗？我又不会算法，怎么负责 AI？万一别人问我模型原理，我答不上来怎么办？是不是得先学 Python、RAG、LangChain、MCP，才能开始？

其实不用。

至少在企业 AI 落地的早期阶段，不用。

因为企业真正需要的第一个人，往往不是最懂模型的人，而是最能把业务问题讲清楚的人。

这也是为什么，最后负责 AI 的人，常常不是 AI 工程师。

很多 AI 项目失败，不是因为技术没人懂

企业里做 AI 项目时，最常见的分工是这样的：业务部门提出需求，技术部门负责实现，老板负责看结果。

听起来很合理。

但实际推进时，很快就会出问题。

业务部门说：“我们想用 AI 提升效率。”

技术部门问：“具体提升哪个流程？”

业务部门说：“就是很多重复工作。”

技术部门继续问：“重复工作有哪些？输入是什么？输出是什么？判断规则是什么？”

业务部门往往答不上来。

这不是业务部门不专业，而是他们平时并不会用“系统需求”的方式描述自己的工作。

他们知道自己痛苦，但不一定能把痛苦拆成可开发、可评估、可上线的任务。

另一边，技术人员懂系统、懂接口、懂模型调用，但未必知道业务现场真正发生了什么。

他们可能会把需求理解成一个功能：做一个智能客服，做一个合同审核助手，做一个知识库问答系统。

但业务真正需要的，可能只是其中一个很小的环节。

客服不是需要 AI 回答所有问题，而是先帮他们判断客户问题属于投诉、售后还是咨询。

法务不是需要 AI 代替律师审核合同，而是先帮他们把合同里的金额、付款时间、违约条款提取出来。

财务不是需要 AI 自动批准报销，而是先帮他们检查发票抬头、金额、日期是否异常。

很多项目不是缺技术，而是缺一个人站在中间，把业务语言翻译成 AI 能处理的问题。

这个人，就是 AI 落地项目里最关键的角色。

企业 AI 落地，需要的是“连接者”

第一次负责 AI 项目的人，不一定要成为算法专家。

更准确地说，他首先要成为一个连接者。

连接谁？

连接业务、技术和管理层。

业务部门关心的是：这东西能不能帮我少干点重复活？出错了谁负责？会不会反而增加工作量？我原来的流程要不要改？

技术部门关心的是：数据在哪里？接口有没有？权限怎么开？系统怎么部署？日志怎么记录？

老板关心的是：要花多少钱？多久能看到结果？能省多少人力？风险大不大？以后能不能复制到别的部门？

如果没有人把这三方的问题放在同一张桌子上，AI 项目很容易变成各说各话。

业务觉得技术不懂现场。技术觉得业务需求不清楚。老板觉得项目推进太慢。

而 AI 项目负责人要做的，正是把这些问题串起来。

他不一定亲自写所有代码，但要知道业务到底痛在哪里。

他不一定亲自训练模型，但要知道 AI 适合解决哪一段流程。

他不一定亲自部署系统，但要知道上线前必须有哪些风险控制。

他不一定是财务专家，但要能算出项目大概值不值得做。

这就是企业 AI 落地真正需要的能力。

负责 AI，不等于一个人把所有事做完

很多人一听“你来负责 AI”，就会误以为所有事情都要自己会。

于是开始焦虑：要学模型原理，要学代码，要学提示词，要学低代码平台，要学私有化部署，要学数据安全，要学项目管理。

最后越学越慌，反而迟迟不敢开始。

其实，负责 AI 项目，并不意味着你要成为所有领域的专家。

更像是你要成为这个项目的“主驾驶”。

主驾驶不一定会修发动机，但他要知道车往哪里开，什么时候该刹车，什么时候该加油，什么时候必须找维修人员。

AI 项目也是一样。

你不需要一个人完成全部工作，但你需要知道：哪些问题要问业务，哪些问题要找 IT，哪些问题要请法务确认，哪些问题要向老板汇报，哪些地方不能让 AI 自动决定，哪些结果必须人工复核。

AI 项目负责人最重要的能力，不是“什么都会”，而是“知道每件事该找谁、怎么判断、怎么推进”。

这比单纯掌握一个工具更重要。

这个角色，后来有了一个名字：AI 落地工程师

如果你正在做这些事情：和业务部门一起梳理流程，找到适合 AI 的小场景，用工具搭一个原型，拿真实样本测试效果，算出大概节省多少时间和成本，推动小范围试用，设计人工兜底，把结果汇报给老板——那你其实已经站在一个新角色的入口。

这个角色可以叫：

AI 落地工程师。

这里的“工程师”，不只是写代码的人。

它更像是一种工程化思维：把一个模糊想法，变成具体问题；把一个具体问题，变成可验证方案；把一个可验证方案，变成可上线流程；把一个上线流程，变成持续创造价值的能力。

这才是“落地”的意思。

会用 AI，不等于能落地 AI。

能让 AI 稳定进入业务流程，才叫真正落地。

图3 AI 负责人三能力模型

企业需要的是连接业务、工具和价值的人



第一次负责 AI，不必先成为专家，但要先成为**连接者**。

你不需要一开始成为算法专家。早期最重要的是把业务、工具和价值连接起来。

这个角色的能力可以先简化成三件事。

第一，听懂业务痛点。

AI 项目的起点不是模型，而是业务。

很多业务人员说需求时，表达会很模糊：我们客服太忙了，合同审核效率太低，财务报销太麻烦，销售跟进客户很费时间，HR 筛简历筛不过来。

这些话还不是 AI 项目，只是痛点。

你需要继续追问：客服忙在哪里？合同审核慢在哪里？报销麻烦在哪里？销售跟进费时间是写邮件，还是查客户背景？HR 筛简历，是看学历、年限、技能，还是判断候选人与岗位匹配度？

只有把模糊痛点拆成具体步骤，AI 才有机会介入。

第二，搭出一个能看的原型。

听懂业务之后，下一步不是马上做大系统，而是搭一个小原型。

这个原型不需要完美，只需要让业务人员看到：原来 AI 真的能帮我处理一部分工作。

你可以用现成平台搭一个简单分类器，可以用表格加 AI 插件做信息抽取，可以用知识库工具做简单问答，也可以用工作流工具串起“上传文件—提取字段—生成结果”的流程。

第三，讲清楚业务价值。

很多 AI 项目演示时很热闹，但最后没有预算。

原因很简单：负责人讲的是技术亮点，老板听的是投入产出。

不要只讲模型能力很强，要讲每月能处理多少单；不要只讲 AI 可以自动生成，要讲每次节省多少分钟；不要只讲准确率很高，要讲哪些错误可以接受，哪些错误必须人工复核。

企业愿意持续投入 AI，不是因为 AI 很新，而是因为它能创造可衡量的价值。

三类人，最适合转向 AI 落地

很多人以为，AI 落地工程师一定来自技术背景。

其实不一定。

在企业里，最容易转向这个角色的，往往是下面三类人。

第一类，是业务运营和管理人员。

这类人的优势是懂业务。他们知道每天哪些事情最烦，哪些流程最慢，哪些地方最容易出错。他们也知道一线员工真正愿不愿意用一个新工具。

这类人最需要补的是基本 AI 工具使用能力、低代码平台操作能力、简单数据整理能力和项目推进能力。

第二类，是产品经理和项目经理。

这类人的优势是会拆需求。他们习惯把一个模糊目标拆成用户、场景、流程、功能和验收标准，也擅长跨部门沟通。

这类人最需要补的是 AI 能力边界、评估方法、错误处理、人工在环和基础模型选型。

第三类，是 IT 或开发人员。

这类人的优势是能落地。他们知道系统怎么接，数据怎么流，权限怎么开，日志怎么记，也更容易理解 API、数据库、部署和安全问题。

这类人最需要补的是业务访谈、ROI 计算、场景优先级判断、非技术沟通和项目汇报。

不管你从哪里开始，都有机会进入这个角色。

业务背景的人，不要因为不会代码就放弃。技术背景的人，也不要以为会开发就足够。产品背景的人，更要理解 AI 项目和普通软件项目不一样。

真正成熟的 AI 落地工程师，最后都会长成复合型人才。

你不需要一开始就懂所有技术，但必须知道 AI 的边界

不懂底层模型，不代表你不能开始做 AI 项目。

当然，技术能力越强，未来能走得越远。但在第一个 AI 项目里，你更需要先掌握的是：看懂一个业务流程，找到一个小场景，准备一批真实样本，用工具做一个简单演示，让业务人员试用，记录反馈，算出大概价值。

这些事情，很多非技术背景的人也能做。

你可以不会写复杂代码，但你要会问问题。你可以不了解模型训练，但你要知道 AI 可能会出错。你可以不懂私有化部署，但你要知道涉及敏感数据时不能随便上传。

AI 很强，但它不是万能的。

它可能会理解错材料，可能会生成看似合理但实际错误的内容，可能会在格式复杂的文件里漏掉信息，也可能对模糊问题给出不稳定答案。

所以，负责 AI 的人必须有一个基本判断：

AI 的结果，不能在没有任何验证的情况下直接进入高风险决策。

AI 可以帮你初筛合同，但不应该独立做合同终审。AI 可以帮你检查报销单异常，但不应该自动批准大额付款。AI 可以帮客服生成回复草稿，但敏感投诉最好仍由人工确认。

你不一定要知道模型内部怎么工作，但你必须知道：哪里可以试，哪里要谨慎，哪里必须保留人工。

这是 AI 项目负责人最基本的责任。

这一章读完，你可以马上做什么

1. 判断自己更接近业务型、产品型还是技术型背景。
2. 写下自己最需要补的一项能力：业务理解、原型搭建，还是价值表达。
3. 找到一个愿意配合你的业务伙伴，因为 AI 项目不能只靠一个人闭门研究。

本章小结

负责 AI，不代表你必须成为算法专家。

在企业 AI 落地的早期，最重要的人往往不是最懂模型的人，而是最能连接业务、技术和价值的人。

这个人要听懂业务痛点，能搭出简单原型，也能向老板讲清楚项目价值。

他要知道 AI 能做什么，也要知道 AI 不能做什么。

他要推动项目开始，而不是沉迷工具试用。

如果你是那个被老板叫去“研究 AI”的人，不要急着否定自己。

你可能正站在一个新角色的起点：AI 落地工程师。

下一章，我们开始真正动手：公司里这么多事情，到底哪个适合先做？

第三章 别急着做 Agent，先找到真正值得做的事

很多人第一次负责 AI 项目时，脑子里会冒出一个很大的目标：

“我们能不能做一个公司级 AI 助手？”

这个想法很自然。

因为你看到的很多演示，都是一个无所不能的 AI 助手。它能回答问题，能查资料，能写报告，能做分析，还能调用各种工具。

看起来，只要公司也有这样一个助手，效率就能立刻提升。

但真正进入企业后，你会发现事情没那么简单。

销售说，希望 AI 帮他们整理客户资料。

财务说，希望 AI 帮他们审核报销。

人事说，希望 AI 帮他们筛简历。

采购说，希望 AI 帮他们比价和跟进订单。

客服说，希望 AI 帮他们回复客户。

老板说，最好能把这些都做了。

于是，一个原本想从小处开始的 AI 项目，很快变成了一个巨大的愿望清单。

每个部门都想要。每个需求都有道理。每个场景看起来都能用 AI。

但你很快就会陷入新的问题：

到底先做哪个？

这是 AI 落地中非常关键的一步。

因为第一个场景选错了，后面做得再努力，也很容易失败。

第一个项目，不要选“看起来最厉害”的

很多人选第一个 AI 场景时，会犯一个错误：专挑最炫、最复杂、最能体现 AI 能力的场景。

比如，让 AI 自动审核所有合同，让 AI 独立处理全部客服问题，让 AI 自动生成经营分析报告，让 AI 替老板做决策，让 AI 接管一个完整部门的流程。

这些场景听起来很有想象力，也很适合放在 PPT 里。

但它们通常不适合作为第一个项目。

原因很简单：越复杂，越难验证；越复杂，越难上线；越复杂，越容易涉及责任风险。

第一次做 AI，你真正需要的不是证明 AI 有多强，而是证明：

AI 能在一个具体环节里，稳定帮人节省时间或减少错误。

这才是第一个项目的目标。

所以，不要问“哪个场景最震撼”，要问“哪个场景最容易验证价值”。

这两个问题的答案，往往不是同一个。

好场景，通常藏在那些“烦但不难”的工作里

真正适合作为第一个 AI 项目的场景，通常不神秘。

它们往往就藏在日常工作里。

这些工作有一个共同特点：

烦，但不一定难。

客服每天要给客户留言分类。这件事不难，但量一大就很烦。

财务每个月要检查大量发票和报销单。每张单据都不复杂，但重复处理几千次，人就会崩溃。

HR 要从简历里找出学历、工作年限、技能关键词。一份简历不难看，但几十份、几百份叠在一起，就是负担。

销售要根据客户信息写跟进邮件。一封邮件不难写，但每天都要写，而且还要看起来不那么敷衍。

区域经理要根据巡店照片写报告。判断标准并不复杂，但照片多、门店多、格式多，很耗时间。

这些工作不像“AI 替代高管决策”那么吸引眼球，但它们更真实，也更容易落地。

因为业务人员已经在做这些事情，规则大致存在，样本也容易拿到。

如果 AI 能帮一部分，业务方很快就能感受到变化。

第一个 AI 项目，应该优先从这类工作里找。

贯穿案例：客服邮件分类

客服主管每天要先把客户邮件分成投诉、咨询、售后、合作等类型，再转给不同同事处理。每封邮件不难判断，但数量一多就很耗时间。这个场景不需要 AI 一开始自动回复客户，只要先做“分类建议”，就能验证价值。

不要找“一个部门”，要找“一个动作”

很多初学者在找 AI 场景时，会说：我们想做财务 AI，我们想做销售 AI，我们想做 HR AI，我们想做客服 AI。

这些说法太大。

一个部门里面有很多流程，一个流程里面又有很多动作。

如果一开始就说“做财务 AI”，很快就会不知道从哪下手。

正确的方式是继续往下拆。

不要找一个部门，要找一个流程。

不要只找一个流程，还要继续找其中一个动作。

不是“做财务 AI”，而是“报销审核里的发票抬头检查”。

不是“做合同 AI”，而是“从合同中提取金额、税率、付款时间和交付日期”。

不是“做客服 AI”，而是“把客户留言分成投诉、咨询、售后和合作四类”。

不是“做 HR AI”，而是“从简历里提取岗位、年限、技能和所在地”。

不是“做销售 AI”，而是“根据客户背景生成一版跟进邮件草稿”。

场景一旦拆到“动作”层面，就变得清楚了。

清楚，就能验证。

能验证，才有机会成功。

一个适合 AI 的场景，通常满足四个条件

图4 AI 场景筛选漏斗



筛选场景不是为了证明 AI 无所不能，而是为了找到最容易做成的第一步。

如果你手里有一堆候选场景，可以先用四个条件筛一遍：

高频、重复、规则清晰、输入多样。

第一，高频。

一个任务每月发生 3000 次，每次节省 1 分钟，一个月就是 3000 分钟。如果一个任务一年只发生两次，即使 AI 做得很好，也很难证明价值。

所以，当你听到一个需求时，第一句可以问：这个事情每个月大概发生多少次？

第二，重复。

AI 很适合处理重复性强的工作，比如分类、提取、检查、整理、生成草稿。

这些工作虽然每次内容不同，但任务模式类似。客户邮件内容不一样，但分类逻辑类似；合同文本不一样，但关键字段类似；简历经历不一样，但筛选维度类似。

这种“内容变化，但任务模式稳定”的工作，很适合 AI。

第三，规则清晰。

AI 不是魔法。

如果业务人员自己都说不清楚什么叫“好”、什么叫“不合格”、什么叫“异常”，AI 也很难稳定完成任务。

比如财务能说清楚：发票抬头不一致要退回，金额和报销单不一致要退回，没有审批记录要退回，超预算但没有说明要退回。这就是规则清晰。

规则越清楚，AI 越容易做。

第四，输入多样。

如果规则固定、输入也固定，其实不一定需要 AI。

订单金额超过 5000 元就提醒，库存低于 100 件就补货，日期超过 30 天就标红，这些用普通代码、Excel 公式、RPA 都能解决。

AI 真正擅长的，是处理输入不那么规整的内容，比如邮件、聊天记录、PDF、Word、图片、照片、语音转文字、合同文本、简历、报告。

这些内容格式不统一，表达方式多变，普通代码处理起来麻烦，但大模型更有优势。

所以，一个好场景往往是：

规则能说清楚，但输入很复杂。

可以先从四类场景里找

完全不知道从哪里开始时，可以先看四类场景。

第一类，分类。

分类是最适合入门的 AI 场景之一，因为它简单、清楚、容易评估。

比如客户留言分类、工单分类、销售线索分类、投诉等级分类、简历匹配分类、合同风险等级分类。

分类任务的好处是，结果通常很明确。对就是对，错就是错。你很容易拿历史数据来测试。

第二类，信息抽取。

很多企业都有大量非结构化文件：合同、发票、简历、报价单、报告、会议纪要、采购单。

这些文件里有很多关键信息，但人工提取很耗时间。

比如从合同中提取合同金额、税率、付款节奏、交付日期、质保期、违约责任；从简历中提取姓名、学历、工作年限、技能和最近一家公司。

输入格式多样，但目标字段明确，这类场景非常适合 AI。

第三类，检查和比对。

比如报销单和发票是否一致，合同和招标文件是否一致，报价单和客户需求是否一致，订单和出库记录是否一致，简历和岗位要求是否匹配。

这类任务的核心，不是让 AI 做最终决定，而是让它先找出可能有问题的地方。

第四类，草稿生成。

很多工作不是不会写，而是写起来耗时间。

比如客服回复草稿、销售跟进邮件、会议纪要、巡店报告、日报周报、整改建议。

这类任务可以先让 AI 生成初稿，再由人修改确认。

但草稿生成不适合一上来就完全自动发送。尤其是客户回复、法律意见、对外公文、员工沟通等场景，都应该保留人工确认。

不要只听需求，要看真实样本

找场景时，有一个非常重要的动作：一定要看真实样本。

不要只听业务人员描述。

因为业务人员描述出来的流程，往往比真实情况干净很多。

他说“我们就是审核一下报销单”，你拿到真实材料后，可能会发现：有的发票是图片，有的是 PDF，有的截图模糊，有的金额没问题但预算科目错了，有的审批记录在另一个系统里。

他说“我们就是看一下合同”，你拿到真实合同后，可能会发现：有的是扫描件，有的条款顺序不一样，有的字段写在附件里，有的付款条件藏在补充协议里。

真实样本会告诉你三件事：输入到底有多乱，规则到底有多清楚，AI 的第一步应该做到哪里。

没有样本，就不要急着做方案。

找场景时，至少要访谈三类人

一个 AI 场景能不能做，不能只问一个人。

最好至少访谈三类人。

第一类，真正做这件事的人。

他们是一线用户，比如客服专员、财务审核员、销售助理、人事招聘专员、采购跟单员。

他们最清楚每天的工作细节。你要问他们：这件事每天做几次？最耗时间的是哪一步？最容易出错的是哪一步？如果 AI 只能帮一步，你最希望它帮哪一步？

第二类，这个流程的负责人。

比如客服主管、财务经理、HR 负责人、采购经理、销售运营负责人。

他们更关心流程效率和管理结果。你要问他们：这个流程现在有没有统计数据？错误会造成什么影响？有没有历史样本可以提供？哪些结果必须人工确认？如果做试点，谁可以配合验收？

第三类，IT 或系统负责人。

很多 AI 项目不是想法不好，而是数据拿不到。

所以，最好早点问 IT：数据在哪个系统里？能不能导出？有没有接口？是否涉及敏感数据？是否允许使用云端工具？如果先做 MVP，能不能先用脱敏样本？

不要等 Demo 做完才发现数据不能用。

那会很尴尬。

一个真实例子：从“做财务 AI” 拆到“发票异常检查”

假设老板说：“财务这块能不能用 AI 提效？”

这句话太大。

如果你直接接下来，很容易陷入“做一个财务智能体”的大坑。

更好的做法，是去找财务部门聊。

你可能会发现：财务每月处理 3000 张报销单，其中很多时间花在检查发票和报销单是否一致。

常见问题包括发票抬头不对、金额不一致、日期超期、预算科目填错、缺少审批记录、附件不完整。

这时，场景就逐渐清楚了。

你可以把“做财务 AI”拆成几个小任务：发票字段识别，报销单与发票金额比对，发票抬头检查，预算科目匹配，异常报销提醒，退回原因草稿生成。

其中，最适合作为第一个 MVP 的，可能不是“自动审核报销”，而是：

报销单与发票信息一致性检查。

因为它足够具体，有真实样本，规则也相对清楚，出错后可以人工复核。

这就是一个更好的起点。

场景优先级，不靠感觉，靠打分

当你收集到多个候选场景后，不要靠感觉决定先做哪个。

可以用一个简单打分表。

看五项：频率、耗时、规则、样本、风险。

前四项分数越高越好。最后一项要看是否可控。

客户邮件分类错了，可以人工改回来，风险较低。合同金额识别错了，如果只是初筛，也可人工复核，风险可控。

但如果让 AI 自动批准付款，风险就很高，不适合作为第一个项目。

所以，第一个场景不一定要总分最高，而是要选择：

价值明显、规则清楚、风险可控、样本容易拿到的场景。

有些场景，第一阶段应该主动放弃

找场景时，不只是要知道什么适合做，还要知道什么暂时不要做。

第一类，责任太重的最终决策。

比如合同终审、法律意见定稿、大额付款审批、员工录用决定、医疗诊断建议、信贷最终审批。

这些场景不是 AI 完全不能参与，而是不能让 AI 在第一阶段做最终决策。更适合的做法是 AI 做初筛、提取、提醒、比对，人做最终判断。

第二类，规则说不清的场景。

如果业务人员自己都说不清判断标准，只说“凭经验”，那就要谨慎。

你可以先把它拆小。不要直接判断“客户有没有潜力”，先提取客户行业、规模、预算、历史沟通记录。不要直接判断“候选人好不好”，先提取岗位匹配信息。

第三类，数据拿不到的场景。

有些场景看起来很好，但数据在系统里拿不到，或者涉及敏感信息，无法用于试点。

作为第一个 MVP，可以先用脱敏样本或模拟数据。如果连脱敏样本都拿不到，就不要急着做。

第四类，老板很想做，但一线没人愿意用的场景。

老板觉得某个 AI 工具很酷，但一线员工觉得没必要。如果没有真实用户愿意配合，项目很难成功。

AI 项目不是做给 PPT 看的，是做给真实工作流用的。

最小场景的标准：一句话能说清楚

判断一个场景是否足够小，有一个简单方法：

能不能用一句话说清楚？

“用 AI 做财务提效”，不够小。

“用 AI 检查报销单和发票金额是否一致”，就清楚很多。

“用 AI 做销售助手”，不够小。

“用 AI 把客户邮件分成投诉、咨询、售后三类”，就可以开始。

第一个 AI 项目的场景，最好能写成这样的格式：

用 AI 帮谁，在什么材料里，完成哪一个具体动作。

例如：用 AI 帮客服，从客户邮件里，判断问题类型。

用 AI 帮财务，从发票和报销单里，检查金额是否一致。

用 AI 帮 HR，从简历里，提取岗位匹配信息。

用 AI 帮区域经理，从巡店照片里，生成报告初稿。

只要能写出这句话，你的 AI 项目就已经开始变清楚了。

这一章读完，你可以马上做什么

1. 列出 5 个公司里高频、重复、让人烦的工作。
2. 用“用 AI 帮谁，从什么材料里，完成哪一个动作”这个句式描述每个场景。
3. 用场景优先级打分表筛出最适合作为 MVP 的一个。

本章小结

第一次做 AI，不要急着搭 Agent，也不要一上来就做公司级助手、大而全系统或全自动流程。

你真正要做的，是先找到一个值得做的小场景。

一个好场景，通常满足四个条件：高频、重复、规则清晰、输入多样。

它最好能从一个具体动作开始，而不是从一个完整部门开始。

不要说“做财务 AI”，要说“先检查报销单和发票金额是否一致”。

不要说“做销售 AI”，要说“先把客户邮件分成投诉、咨询和售后”。

AI 落地的第一步，不是让 AI 变得无所不能，而是让它在一个小地方，真正帮上忙。

下一章，我们要判断：这个场景到底该不该用 AI？

第四章 什么事情适合交给 AI，什么事情不适合？

找到一个看起来不错的 AI 场景以后，新的问题马上出现了：

这件事真的适合用 AI 吗？

这是很多人容易忽略的一步。

前面几章，我们一直在强调：不要一上来就选模型，也不要一开始就做一个大而全的系统。你要先找到一个具体场景。

但场景找到了，并不等于马上就该上 AI。

因为企业里有很多事情，虽然可以自动化，却不一定需要 AI。

有些事情，用 Excel 公式就能解决。有些事情，用普通代码更稳定。有些事情，用 RPA 就够了。还有些事情，AI 可以辅助，但绝不能独立决定。

如果不做判断，很容易出现一种尴尬情况：明明几行规则代码就能完成的事，非要用大模型；明明需要专业人员负责的事，却让 AI 自动拍板；明明只是一个简单流程，却被包装成复杂 Agent。

最后，成本变高了，风险变大了，业务反而更不信任 AI。

所以，真正成熟的 AI 项目负责人，不是到处使用 AI，而是知道：

什么时候该用 AI，什么时候不该用 AI。

不是所有自动化，都需要 AI

很多人第一次接触 AI 时，会产生一种冲动：既然 AI 这么强，那公司里所有重复工作是不是都可以交给 AI？

答案是：不一定。

企业自动化有很多种方式。AI 只是其中一种。

比如，你要判断一个订单是否超时。规则是：如果发货日期超过承诺日期 3 天，就标记为异常。

这件事不需要 AI。

普通代码、Excel 公式、系统规则都能稳定完成。

再比如，你要判断一个客户是否欠费。规则是：如果应收款超过账期 30 天，就提醒销售跟进。

这也不需要 AI。

因为输入固定，规则固定，结果也固定。

在这类场景里，用 AI 反而是“杀鸡用牛刀”。成本更高，结果还可能不稳定。

普通代码每次都会给出一样的结果，AI 则可能因为输入表达方式不同，产生不同回答。

所以，判断一个场景是否适合 AI，第一步不是问“AI 能不能做”，而是问：

有没有更简单、更稳定、更便宜的方法？

如果普通规则就能解决，那就优先用普通规则。

这不是保守。

这是专业。

AI 最适合的，是“规则能说清楚，但输入很复杂”的事情

什么事情适合用 AI？

可以记住一句话：

规则能说清楚，但输入很复杂。

什么叫规则能说清楚？

就是业务人员能够告诉你：什么是正确，什么是错误，什么是异常，什么需要提醒，什么需要人工确认。

比如财务审核报销单，规则可能是：发票抬头必须和报销主体一致，发票金额必须和报销金额一致，发票日期不能超过规定时间，超预算必须有审批记录。

这些规则是能说清楚的。

什么叫输入很复杂？

就是每次来的材料不一样。有的是 PDF，有的是图片，有的是扫描件，有的是邮件附件，有的是 Word 文档，有的是聊天记录截图。

这时候，普通代码处理起来很麻烦。但 AI 可以阅读、理解、提取、归纳这些非结构化内容。

这就是 AI 的优势。

贯穿案例：合同字段提取

每份合同的写法都不一样，但采购和法务最关心的字段往往相对固定：金额、税率、付款方式、交付日期、质保期。AI 很适合先帮人把这些信息找出来，但不适合直接判断“这份合同能不能签”。

再看合同场景。

合同审核的规则可能是：必须有合同金额，必须有付款节奏，必须有交付日期，必须有违约责任，必须有保密条款。

这些规则能说清楚。但每份合同的表达方式都不一样，有的写在正文，有的写在附件，有的条款名字不同，有的顺序不同，有的是扫描件。

这就适合让 AI 先做信息提取和差异标记。

注意，是“先做”，不是让 AI 做最终法律判断。

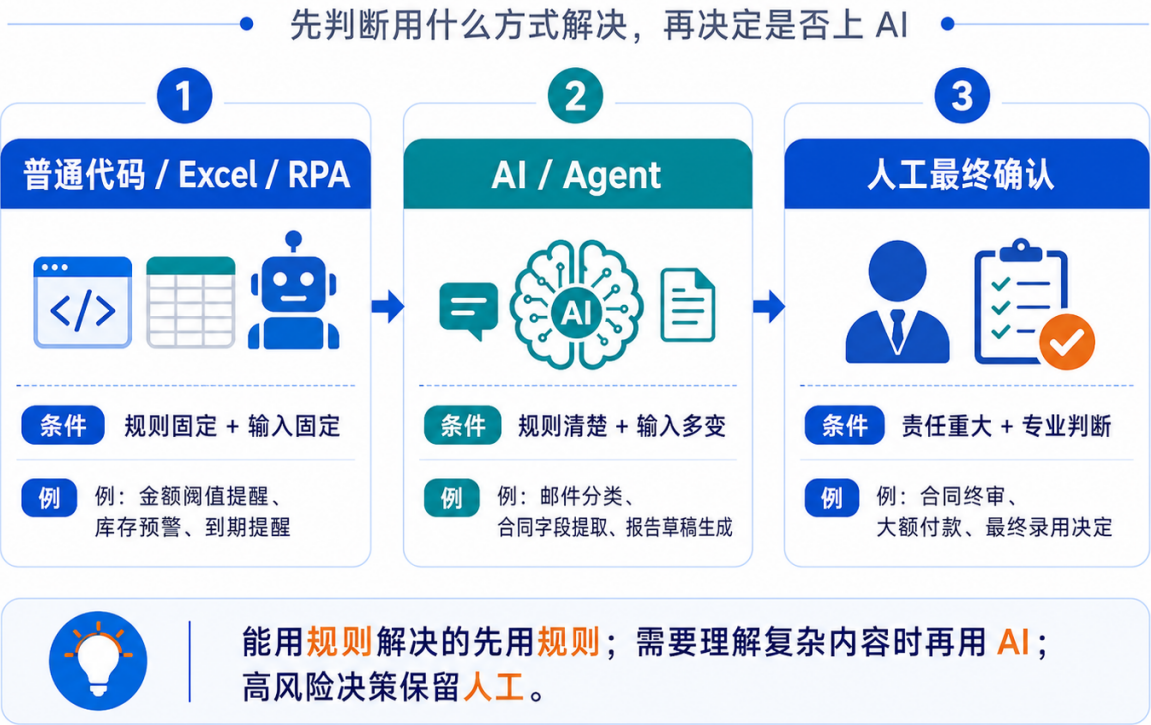
一个简单矩阵：规则和输入

判断是否适合 AI，可以看两个维度：

规则是否清楚？输入是否多变？

这两个问题，基本能帮你筛掉大部分不合适的场景。

图5 AI / 代码 / 人工三分法



真正成熟的 AI 项目，不是什么都交给 AI，而是把每一段工作放到最合适的位置。

场景类型	规则	输入	推荐方式
规则固定，输入固定	清楚	固定	普通代码 / Excel / 系统规则
规则清楚，输入多变	清楚	多变	适合 AI / Agent
规则不清，输入固定	不清	固定	先梳理规则，不急着做 AI
规则不清，输入多变	不清	多变	暂不适合作为第一个项目

规则固定、输入固定的任务，比如订单金额超过 5000 元提醒、库存低于 100 件补货、客户超过 30 天未付款提醒，都不需要 AI。

规则清楚、输入多变的任务，比如客户邮件分类、合同关键字段提取、发票和报销单比对、简历信息抽取、巡店照片生成报告草稿，才是 AI 最适合的区域。

规则不清时，不要急着开发。先把“凭经验”拆成几个可以观察的信号。

规则不清、输入又多变的场景，尤其不适合作为第一个项目。

判断一个任务，先问三个问题

当你拿到一个候选场景时，可以先问三个问题。

第一，这个任务的判断规则能不能说清楚？

如果规则说不清，AI 就很难稳定。

你可以问业务人员：你现在是怎么判断的？什么情况算通过？什么情况算异常？什么情况必须退回？什么情况必须上级审批？有没有历史案例？有没有标准模板？

如果对方只说“这个主要靠经验”“这个很难说”“你看多了就知道”，那就要小心。

第二，输入是不是经常变化？

如果输入固定，普通规则可能更合适。

系统表单里的金额、日期、状态、库存数量，这些字段很规则。如果只是判断数值大小、日期是否超期、字段是否为空，不需要 AI。

但如果输入是邮件、合同、图片、聊天记录、报告、简历、发票扫描件，AI 的价值就会更明显。

第三，出错以后，损失是否可控？

一个任务适合 AI，不代表它适合自动执行。

客户邮件分类错了，可以人工改回来，风险较低。

合同字段提取漏了一项，如果只是初筛，法务还能复核，风险可控。

报销单异常没有识别出来，如果只是提醒工具，财务还能最终审核，风险可控。

但如果 AI 自动批准付款、自动签合同、自动拒绝候选人、自动给客户退款，就完全不一样。

所以，判断 AI 能不能做，还要判断：

AI 可以做到哪一步？

它可以看，可以提取，可以分类，可以提醒，可以生成草稿，但不一定能批准、发送、支付、签署、拒绝。

三种处理方式：代码、AI、人工

很多 AI 场景，最后会分成三种处理方式。

不是所有流程都给 AI，而是一部分给普通代码，一部分给 AI，一部分保留人工。

第一种，确定性代码。

金额大于 5000 元就标红，日期晚于交付期就提醒，字段为空就退回，库存小于安全值就补货。

这类规则非常适合用普通代码、系统配置或 Excel 完成。它们便宜、稳定、可解释。

第二种，AI / Agent。

AI 适合处理输入复杂、表达多样、需要理解语义的部分。

比如识别客户邮件意图，从合同里找付款条款，从简历里提取项目经验，从会议纪要里整理行动项，根据巡店照片生成问题描述。

如果任务还需要分步骤完成，比如先读取文件，再提取字段，再调用系统查询，再生成结果，就可以进一步发展成 Agent。

在大众版里，你可以先把 Agent 理解成：

一个能围绕目标，分步骤调用工具完成任务的 AI 应用。

第三种，人工判断。

有些事情必须保留人工，特别是涉及责任、伦理、法律、财务、安全的关键动作。

比如合同最终签署、大额付款审批、员工录用或淘汰、法律意见定稿、医疗诊断、重要客户投诉处理、信贷最终决策。

AI 可以辅助这些事情，但不应该独立承担最终责任。

更好的方式是：AI 提取信息，AI 标记风险，AI 生成建议，人做最后确认。

AI 最适合“辅助判断”，不适合“替人负责”

这是企业 AI 落地里非常重要的一条原则。

AI 可以帮人看得更快，但不能替人承担责任。

AI 可以帮人发现问题，但不能替人决定所有问题。

AI 可以帮人生成草稿，但不应该在高风险场景里自动发出。

AI 可以帮人排序候选项，但不应该不经确认就淘汰某个人。

第一次做 AI 项目时，可以把 AI 的角色设定为：助手，而不是负责人；初筛员，而不是审批人；草稿生成器，而不是最终发布者；风险提醒器，而不是责任承担者。

这个定位越清楚，项目越容易上线。

因为业务人员不会觉得自己被取代，管理层也更容易接受风险，IT 和合规也更容易放行。

用一个报销场景来看怎么拆

假设我们要做一个“报销审核 AI”。

这个说法太大。

一个报销审核流程里，可能有很多步骤：读取报销单，识别发票信息，检查发票抬头，比对金额，判断费用科目，检查预算是否充足，判断是否需要上级审批，批准或退回报销。

这些步骤不是都要交给 AI。

步骤	推荐方式	原因
读取报销单	AI / OCR	材料格式可能多样
识别发票信息	AI / OCR	图片、PDF、扫描件不统一
检查发票抬头	普通规则 + AI	规则明确，但输入可能复杂
比对金额	普通代码	数值比对更稳定
判断费用科目	AI 辅助 + 规则	描述可能不标准
检查预算	系统规则	数据在预算系统中
判断是否需审批	普通规则	金额、部门、规则明确
批准或退回	人工确认	涉及责任，需人工

这才是现实中的 AI 落地。

不是让 AI 从头到尾接管全部流程，而是把每一步放到最合适的位置。

能用规则的，用规则。需要理解复杂材料的，用 AI。涉及责任的，留给人。

这一章读完，你可以马上做什么

1. 拿刚选出的场景，回答“规则是否清楚、输入是否多变、出错是否可控”三问。
2. 把流程拆成三类：普通代码能做的、AI 适合做的、必须人工确认的。
3. 写一张“AI 不能做什么”清单，先把边界说清楚。

本章小结

找到一个场景以后，不要立刻上 AI。

先判断它到底适不适合。

不是所有自动化都需要 AI。规则固定、输入固定的事情，普通代码更便宜、更稳定。规则说不清、责任又重的事情，不适合作为第一个项目。

AI 最适合的，是那些规则能说清楚、输入形式多样、任务高频重复、结果可以复核、错误风险可控的场景。

第一次做 AI，不要追求让它独立负责。

更好的方式是让 AI 先看一遍、提取一遍、分类一遍、生成一版草稿、标记一批风险。

人做最终判断。

下一章，我们开始做第一个东西：一周做出一个能演示的 AI MVP。

第五章 一周做出一个能演示的 AI MVP

前面几章，我们一直在做准备。

你已经知道，第一次做 AI 不要先研究工具，也不要一上来做大系统。

你要先找到一个具体场景，再判断它到底适不适合 AI。

现在，真正的动作开始了。

你选中了一个场景。比如客服邮件分类、合同字段提取、报销单异常检查、简历信息抽取、巡店报告草稿生成。

老板也说：

“这个方向可以，你先做一个看看。”

这时候，很多人又会掉进一个新坑：开始设计一个完整系统。

要有登录页面，要有权限管理，要接公司数据库，要打通 OA，要接企业微信，要有漂亮的后台，要能自动执行，要能生成报表。

听起来很完整。

但如果这是你的第一个 AI 项目，这么做通常太早了。

因为你还没有证明一件最重要的事：

AI 在这个场景里，到底能不能帮上忙？

所以，第一个阶段不应该做完整系统。

你真正要做的是一个 MVP。

MVP 不是简陋版系统，而是最小验证

MVP 是 Minimum Viable Product，最小可行产品。

你可以把它理解得更简单：

用最小成本，验证一个关键问题。

在 AI 项目里，MVP 要验证的不是“系统能不能做得很完整”，而是：

AI 能不能在一个具体环节上，稳定地帮人节省时间、减少错误，或者提高处理速度。

这件事验证清楚之前，不要急着接系统、不要急着上线、不要急着做完整流程。

因为你可能会花很多时间，把一个本来就不值得做的场景做得很漂亮。

一个失败的 AI MVP，只花一周就知道方向不对，是好事。

一个失败的 AI 系统，花三个月才发现没人用，是灾难。

第一个 MVP，只需要证明一件小事

很多人做 MVP 时，还是忍不住想多做一点。

做客户邮件分类，就想顺便自动回复、自动派单、自动生成工单、自动统计报表。

做合同字段提取，就想顺便审核风险、生成法律意见、发起审批流程。

做巡店报告，就想顺便识别照片、写整改建议、创建待办、追踪闭环。

这些未来都可以做。

但第一周不要做这么多。

第一个 MVP 只需要证明一个最小能力。

客户邮件分类 MVP，只证明 AI 能不能把客户邮件分成投诉、咨询、售后、合作几类。

合同字段提取 MVP，只证明 AI 能不能从合同里提取金额、税率、付款时间、交付日期。

报销单检查 MVP，只证明 AI 能不能识别发票抬头、金额、日期，并发现明显异常。

巡店报告 MVP，只证明 AI 能不能根据照片和检查记录生成一版可修改的报告初稿。

这些 MVP 都很小。

但正因为小，才容易完成，也容易判断效果。

第一个 MVP 不一定要接入公司系统

很多人一开始就卡在系统接入上。

客户邮件在邮箱里，合同在 OA 里，报销单在财务系统里，简历在招聘系统里，巡店照片在企业微信或钉钉里。

一听要接系统，IT 就要评估接口、权限、安全、合规、预算。

项目还没开始，就已经进入漫长流程。

第一个 MVP 不一定要这样。

你可以先用导出的样本。比如导出 50 封历史客户邮件，拿 20 份脱敏合同，准备 30 张发票截图，整理 20 份简历，拿 10 份巡店照片和历史报告。

先离线验证。

不用一开始就接系统，不用一开始就自动执行，不用一开始就涉及敏感数据。

MVP 的目标是证明可行性，不是完成生产部署。

等结果证明值得做，再进入系统集成和上线阶段。

这会让项目推进快很多。

一个合格的 AI MVP，应该包含五样东西

第一个 AI MVP 不需要复杂，但不能只有一个“能聊的机器人”。

它至少应该包含五样东西。

第一，真实样本。

AI 项目不能只用你自己编的测试数据，必须使用真实业务样本。涉及隐私和敏感信息时，需要脱敏。姓名、手机号、身份证号、银行卡号、客户名称、合同金额等敏感信息，可以先替换或打码。

第二，明确输入。

你要说清楚，AI 接收什么材料。是一封邮件，一份合同，一张发票，一段聊天记录，一组照片，还是一个 Excel 表。

不要写“输入业务资料”，这太模糊。要写“输入客户原始邮件正文和标题”或“输入 PDF 合同正文”。

第三，明确输出。

AI 最后应该给出什么结果？是分类、字段表格、异常清单、报告草稿，还是回复建议？

输出最好结构化。

客户邮件分类，不要只让 AI 写一段话，可以让它输出邮件类型、紧急程度、关键词、建议处理人、回复建议。

合同字段提取，也最好输出成表格：合同金额、税率、付款方式、交付日期、违约金。

结构化输出，才容易检查、评估和汇报。

第四，人工标准答案。

你不能只让 AI 跑一遍，然后凭感觉说“还不错”。

你需要有人给出标准答案。20 封客户邮件，先让客服主管标注正确类别；20 份合同，先让法务或采购人员标注关键字段；20 张发票，先让财务人员标注正确金额、抬头和日期。

有了标准答案，你才能比较 AI 对了多少、错了多少、错在哪里、哪些错误可以接受、哪些错误不能接受。

第五，简单评估结果。

第一个 MVP 不需要复杂评测系统，但至少记录几个指标：准确率大概多少，每条处理耗时多少，人工原来需要多久，AI 处理一次大概成本多少，业务人员觉得结果能不能用。

比如：20 封客户邮件，AI 分类正确 17 封，准确率 85%；人工每封邮件分类约 1 分钟，AI 每封约 5 秒；3 个错误中，2 个是“投诉”和“售后”边界不清；客服主管认为，可以先用于初筛，但不能自动回复。

这就是一个非常有价值的 MVP 结论。

它不完美，但它真实。

第一周不要追求“惊艳”，要追求“可信”

很多 AI 演示喜欢追求惊艳。页面很漂亮，回答很流畅，效果像魔法。

但企业真正看重的不是惊艳，而是可信。

老板会问：这个结果稳定吗？

业务会问：它错的时候怎么办？

IT 会问：数据放在哪里？

财务会问：到底能省多少？

所以，第一个 MVP 的重点不是炫技，而是让大家相信：

这个场景值得继续做。

怎么建立可信？靠真实样本、人工标准答案、错误记录、业务人员反馈和明确的下一步计划。

一个能诚实展示错误的 MVP，比一个只展示成功样本的 Demo 更专业。

因为真正做过项目的人都知道，AI 一定会出错。

关键不是假装它不会错，而是知道它会在哪里错，以及怎么控制这些错误。

三种最适合第一周做的 MVP

第一种，分类器。

分类器是最适合入门的 MVP，因为它简单、清楚、容易评估。

输入可以是客户邮件、工单内容、销售线索、客户留言、投诉记录、候选人简历。

输出是类别，比如投诉 / 咨询 / 售后 / 合作 / 垃圾信息，高意向 / 中意向 / 低意向，紧急 / 普通 / 可延后。

第二种，信息抽取。

很多企业都有大量文件，但关键信息散落在里面。

输入可以是合同、简历、发票、报价单、采购单、会议纪要、巡店记录。

输出是结构化字段，比如从合同里提取金额、税率、付款节奏、交付日期，从简历里提取学历、年限、技能、城市。

第三种，草稿生成。

很多人每天都在写类似内容。

输入可以是客户问题、会议记录、巡店照片、检查表、销售沟通记录、产品资料。

输出可以是回复草稿、报告初稿、整改建议、会议纪要、周报、跟进邮件。

草稿生成见效快，但不适合一开始就自动发送。尤其是对客户、员工、合作伙伴的正式内容，必须人工确认。

一个例子：客户邮件分类 MVP

假设第一个场景是客户邮件分类。

业务情况是：客服团队每天收到大量客户邮件，有人咨询价格，有人反馈产品问题，有人投诉服务，有人申请售后，有人寻求合作。

现在客服主管每天要先手工分拣，再转给不同人员处理。

这件事不难，但很耗时间。

你可以这样做 MVP。

第一步，收集样本。

找客服主管要 50 封历史邮件，去掉客户姓名、电话、订单号等敏感信息，保留邮件标题和正文。

第二步，定义类别。

不要一开始分太细。先定义 4 到 5 类，比如投诉、售后、产品咨询、商务合作、无关信息。每一类写一句简单说明。

第三步，人工标注。

让客服主管先给这 50 封邮件标注正确类别。这一步非常重要，不要省。

第四步，让 AI 分类。

把邮件输入 AI，让它输出邮件类别、判断理由、紧急程度、建议处理人。第一版可以先只做分类，分类稳定后，再加后面的字段。

第五步，对比结果。

把 AI 分类结果和人工标准答案对比，记录正确多少、错误多少、哪些类别最容易混淆。

你可能发现：50 封邮件里，AI 判断正确 43 封，7 封错误中有 5 封是“投诉”和“售后”混淆。客服主管认为，如果 AI 先做初筛，人工再确认，可以接受。

这就是一个可用结论。

第六步，准备演示。

演示时不要只展示成功案例。可以展示一个 AI 判断正确的样本，一个 AI 判断错误但可纠正的样本，一个边界情况需要人工确认的样本。

这样反而更可信。

你可以对老板说：

目前 50 条样本中，AI 分类准确率约 86%。主要错误集中在投诉和售后的边界判断。建议第一阶段只用于初筛，不自动回复客户。如果接入邮件系统，预计可以减少客服主管 40% 到 60% 的分拣时间。下一步建议扩大到 200 条样本继续验证。

这就是一个很专业的 MVP 汇报。

图6 一周做出 AI MVP 时间线

● 第一个原型，不求完美，先求可验证 ●



一周 MVP 的目标不是做完整系统，而是证明一个小动作是否值得继续。

第 1 天，确定一个最小场景。

当天要产出一句话：用 AI 帮谁，从什么材料里，完成哪一个具体动作。

第 2 天，准备真实样本。

准备 10 到 50 条真实样本。数量不用太多，但必须真实，同时完成脱敏。

第 3 天，定义输入、输出和标准答案。

明确输入是什么，AI 输出什么，并让业务人员标注一批标准答案。

第 4 天，用工具搭出原型。

可以用通用 AI 工具、低代码平台、表格加 AI 功能、简单脚本，或者任何你们公司允许使用的工具。

工具不是重点，重点是让流程跑通：输入样本、AI 处理、输出结果、结果可查看。

第 5 天，跑测试，记录错误。

不要只看成功案例，要认真记录分类错、字段漏、格式不对、内容编造、输出不稳定等问题。

第 6 天，优化规则和展示方式。

根据错误做一轮优化，比如调整分类标准、补充例子、固定输出格式、增加“无法判断”选项。

第 7 天，演示和复盘。

演示时展示业务问题、现在怎么做、AI 先帮哪一步、测试了多少样本、效果如何、主要错误是什么、风险怎么控制、下一步建议是什么。

一个好的 MVP 演示，不是说“你看 AI 多厉害”，而是说：

我们用真实样本验证了一个小场景，它目前能解决这些问题，也有这些限制。建议下一步这样推进。

MVP 的结果，不是只有成功或失败

AI MVP 更常见的是四种结果。

第一种，效果很好，可以继续扩大样本。

比如 50 条样本，准确率达到 90% 以上，错误也比较轻微。

第二种，部分可用，需要缩小范围。

比如合同字段提取中，金额、日期、税率很准，但违约条款不稳定。可以先只做基础字段提取，把复杂条款留给人工。

第三种，方向有价值，但规则不清。

比如客户意向判断结果不稳定，你发现不是 AI 不行，而是销售团队自己对“高意向客户”的定义也不统一。这时下一步不是继续调 Prompt，而是先统一业务规则。

第四种，暂时不适合做。

比如数据拿不到，样本太少，错误风险太高，业务方也不愿意复核。

这种结果也很有价值，因为它帮你避免了一个更大的失败。

一个诚实的“不适合现在做”，比硬撑一个假成功更好。

这一章读完，你可以马上做什么

1. 填写 AI MVP 定义表，明确输入、输出、样本数量和成功标准。
2. 约业务人员标注一批标准答案，不要只靠自己判断 AI 对不对。
3. 按一周时间线排出第一个 MVP 的 7 天计划。

本章小结

第一个 AI 项目，不要急着做完整系统。

你真正需要的是一个 MVP。

MVP 的目标不是完美，而是验证 AI 能不能在一个具体环节上帮人节省时间、减少错误或提升速度。

一个合格的 AI MVP，至少要有真实样本、明确输入、明确输出、人工标准答案和简单评估结果。

第一周最适合做的 MVP，通常是分类器、信息抽取和草稿生成。

不要一开始就接系统，不要一开始就自动执行，不要一开始就追求全流程。

先用真实样本，验证一个小动作。

只要这个小动作证明有效，你就有机会继续往下走。

下一章，我们面对老板最现实的问题：这个东西到底值多少钱？

第六章 老板问“这个值多少钱”，你怎么回答？

你做出了第一个 AI MVP。

演示当天，效果还不错。

AI 能把客户邮件分成投诉、咨询、售后几类，也能从合同里提取金额、税率、付款时间，还能根据巡店照片和检查记录生成一版报告草稿。

业务同事觉得有点惊喜，老板也点了点头。

然后，他问了一个很现实的问题：

“这个东西如果继续做，到底值多少钱？”

很多 AI 项目，就是卡在这句话上。

负责人可以讲模型很强、工具很好、未来空间很大、同行都在做。

但老板真正关心的是：花多少钱？省多少钱？多久能看到效果？风险大不大？值不值得继续投？

如果你答不上来，项目很容易停在 Demo 阶段。

不是因为 AI 没用。

而是因为你没有把“有用”翻译成老板听得懂的语言。

这一章要解决的，就是这件事。

AI 项目不能只讲先进，要会算账。

老板不是不相信 AI，而是不知道该怎么投

很多人会误解老板，觉得老板不批预算，是因为不懂 AI。

有时候确实如此。

但更多时候，老板不是不相信 AI，而是不知道这个项目应该怎么判断。

传统软件项目至少有一些熟悉的评估方式。买一套财务系统，能规范账务流程；买一套 CRM，能管理客户线索；上线一个审批系统，能减少纸质流转。

但 AI 项目不一样。

它看起来很聪明，却也有不确定性。它能生成结果，但不一定每次都对。它能节省时间，但节省多少要算。它能提升效率，但是否值得持续投入，也要验证。

所以老板会本能地追问：

“这个项目继续做下去，会不会只是一个玩具？”

你要做的，不是说服老板“AI 很重要”。这句话他大概率已经知道。

你要做的，是说清楚：

这个具体场景，为什么值得继续做。

注意，是具体场景。

不是“AI 很有前景”，不是“未来企业都要用 AI”，而是客服邮件分类这个场景，每月处理多少邮件，人工现在花多少时间，AI 能节省多少，继续做要花多少钱，什么时候能回本。

这才是企业决策语言。

第一次算 ROI，不要搞得太复杂

ROI 是 Return on Investment，投资回报率。

在第一个 AI 项目里，你不用一开始就做得很复杂。

你只需要回答一个基本问题：

继续投入这件事，值不值？

最简单的算法是：

$$\text{月净收益} = \text{当前人工成本节省} - \text{AI 成本} - \text{开发和维护成本}$$

如果月净收益为正，而且风险可控，就值得继续试点。

如果月净收益很低，或者风险很高，就要谨慎。

如果根本算不出收益，说明这个场景还需要重新定义。

第一次算 ROI，不追求精确到小数点，但必须有大致数字。

没有数字，讨论就会变成感觉。

有了数字，讨论就会具体很多。

算账之前，先别急着说“节省多少人”

很多人一谈 AI 价值，就会直接说：

“可以节省几个人。”

这句话要谨慎。

不是不能算人力节省，而是第一次做 AI 项目时，不一定要把价值表达成“减少多少岗位”。

这样说很容易引发业务部门抵触。

更好的表达是：节省多少重复工作时间，缩短多少处理周期，减少多少低级错误，让员工把时间放回更重要的工作上。

比如，客服主管原来每天花 2 小时分拣邮件，现在减少到 30 分钟。

财务原来每月花 80 小时检查发票，现在减少到 25 小时。

区域经理原来写一份巡店报告要 2 小时，现在只需要 25 分钟修改 AI 草稿。

这些表达更容易被接受。

因为你不是在说“AI 要替代谁”，而是在说：

AI 先替人拿走低价值、重复、耗时的部分。

这才是早期 AI 落地更现实的价值。

AI 项目的价值，不只有省人力

人力时间是最容易计算的一类价值，但 AI 项目的价值不止这一种。

第一类，节省时间。

比如邮件分类从人工 1 分钟一封，变成 AI 5 秒初筛；合同字段提取从人工 25 分钟一份，变成 AI 2 分钟生成初稿；巡店报告从人工 2 小时一份，变成 AI 生成初稿后人工修改 25 分钟。

第二类，提升速度。

有些项目不一定直接减少很多人工，但能让流程更快。

比如客户投诉更快被识别出来，销售线索更快分配给销售，合同风险更早被发现，采购报价更快完成初步对比。

速度提升有时比节省时间还重要，因为它可能影响客户体验、成交速度和内部协作效率。

第三类，减少错误。

很多重复工作，人不是不会做，而是容易疲劳。

比如报销单金额看错，合同字段漏看，客户邮件分错部门，简历关键词漏掉，巡店报告漏写问题。

AI 如果能做第一轮检查，就可能减少漏检。

第四类，增加收入。

有些 AI 项目不只是降本，还可能增收。

比如销售跟进邮件生成得更快，让销售能跟进更多线索；客户咨询更快回复，提高转化率；老客户需求被更及时识别，带来复购机会。

增收类价值通常比降本更难算，因为影响收入的因素很多。

第一次汇报时要保守，不要轻易说“用了 AI，收入能提高 30%”。更稳的说法是：

“这个项目主要价值在于提高线索响应速度。后续可以跟踪响应时间、跟进数量和成交转化变化。”

先跟踪，再下结论。

最适合新手的 ROI 算法：四步法

图7 简版 ROI 公式图

老板最关心的，不是先进，而是值不值



第一次算 ROI，不求特别复杂，但一定要把账算清楚。

第一次算 ROI 不求精确到小数点，但必须让讨论从“感觉有用”变成“数字上是否值得”。

第一步，算现在花了多少时间。

公式是：

$$\text{月人工耗时} = \text{月处理量} \times \text{单次人工耗时}$$

例如，每月处理 400 份合同，每份合同人工初筛 25 分钟。

$$400 \times 25 \text{ 分钟} = 10000 \text{ 分钟，约等于 } 166.7 \text{ 小时。}$$

这就是当前每月人工耗时。

第二步，算这些时间大概值多少钱。

公式是：

$$\text{当前人工成本} = \text{月人工耗时} \times \text{人员小时成本}$$

人员小时成本不需要一开始算得特别精确，可以用月薪粗略折算。

比如一个员工综合月成本 12000 元，每月工作约 176 小时，那么每小时成本约 68 元。为了简单，可以按 60 元或 70 元估算。

继续刚才的合同例子：166.7 小时 $\times$ 60 元/小时，大约是 10000 元。

第三步，估算 AI 后能节省多少。

注意，不要假设 AI 节省 100%。AI 上线后，人通常仍然需要复核。

比如原来每份合同人工初筛 25 分钟，AI 提取关键字段后，人工复核变成 8 分钟。每份节省 17 分钟，每月 400 份，就是 6800 分钟，约等于 113.3 小时。

按 60 元/小时算，每月节省人工时间价值约 6800 元。

第四步，扣掉 AI 成本和开发成本。

AI 不是免费用的。至少会有工具订阅费、模型调用费、平台费用、开发或配置成本、后续维护成本。

比如工具和模型调用每月 800 元，初期开发和配置花了 24000 元，计划按 12 个月分摊，每月 2000 元。

$$\text{月净收益} = 6800 - 800 - 2000 = 4000 \text{ 元。}$$

这就是一个简单的 ROI 结论。

你可以汇报：

在当前假设下，该场景每月可节省约 113 小时人工复核时间，折算价值约 6800 元。扣除工具和开发分摊后，预计月净收益约 4000 元。建议继续扩大样本，并在一个部门做小范围试点。

这比说“AI 很智能”有说服力得多。

ROI 不要只算一次，要分阶段算

AI 项目的 ROI，不应该只在立项时算一次。

最好分三个阶段算。

第一阶段，MVP 前，算预估账。

这时数据不一定准确，但要做一个粗略判断：这个场景有没有可能值得做？如果连粗略账都不划算，就不要花太多时间。

第二阶段，MVP 后，算验证账。

MVP 做完后，你有了真实测试数据，比如 AI 准确率、处理速度、人工复核时间、错误类型和业务反馈。这时要更新 ROI，不要继续用之前的想象数字。

第三阶段，试点后，算真实账。

小范围上线后，你要用真实使用数据重新计算：实际用了多少次，人工时间是否真的减少，业务人员是否愿意用，错误率是否可接受，模型调用成本是否超预期，有没有新增维护工作。

很多项目在 MVP 阶段看起来很美，试点后才发现业务不用、成本偏高、错误太多，或者流程反而变复杂。

所以 ROI 要持续更新。

不要只做乐观估算，要做三种版本

向老板汇报时，不要只给一个最乐观数字。

最好准备三种估算：保守、基准、乐观。

比如合同字段提取场景，每月 400 份合同，原来每份 25 分钟。

版本	AI 后每份人工复核时间	每月节省时间	说明
保守	15 分钟	约 67 小时	AI 只提取基础字段
基准	8 分钟	约 113 小时	AI 提取字段较稳定
乐观	5 分钟	约 133 小时	流程和模板进一步优化

这样汇报更可信。

老板会觉得你不是只挑好看的数字，而是已经考虑了不确定性。

AI 项目本来就有不确定性。成熟的做法不是回避它，而是把它说清楚。

一个好的汇报，不是要证明所有 AI 都值得做

你不是 AI 销售。

你不是为了证明“所有业务都该用 AI”。

你是项目负责人。

你要帮助公司判断：哪些值得做，哪些不值得做，哪些现在不做但未来可以做。

所以，汇报时不要把所有场景都说成高价值。

更好的方式是分类汇报：

我们调研了 6 个候选场景。其中 2 个适合优先做 MVP，2 个适合轻量使用现成工具，1 个需要先梳理规则，1 个风险较高，暂不建议推进。

这样的汇报，老板会更信任你。

因为你不是在追热点，而是在做判断。

这一章读完，你可以马上做什么

1. 补齐一个场景的月处理量、单次人工耗时和人员小时成本。
2. 算出保守、基准、乐观三档收益，不要只给最漂亮的数字。
3. 给老板准备一句清楚结论：继续试点、轻量使用、先补规则，还是暂缓。

本章小结

AI 项目要想继续推进，不能只靠演示。

老板最终会问：这个值多少钱？

你要能回答。

第一次算 ROI，不需要复杂财务模型。只要算清楚：现在每月处理多少，人工每次花多久，AI 能节省多少时间，工具和开发要花多少钱，风险是否可控。

AI 的价值也不只有省人力。它还可能提升速度、减少错误、改善体验、增加收入。

但越是难量化的价值，越要谨慎表达。

一个成熟的 AI 项目负责人，不是把所有项目都说成值得做，而是能判断：哪些值得投，哪些轻量试，哪些先补规则，哪些暂时别做。

下一章，我们讲最容易翻车的一步：Demo 很好，为什么上线就没人用了？

第七章 Demo 很好，为什么上线就没人用了？

很多 AI 项目最让人沮丧的地方，不是做不出 Demo。

恰恰相反，Demo 往往很好做。

你上传一份合同，AI 能提取关键字段；你输入一封客户邮件，AI 能判断是投诉还是咨询；你放几张巡店照片，AI 能写出一段报告草稿；你把会议录音转成文字，AI 能整理出待办事项。

演示当天，老板觉得不错，业务同事也觉得新鲜。

大家甚至会说：

“这个可以啊，继续推进。”

于是你以为项目快成功了。

但真正上线后，问题开始出现。

业务人员用了一两次，就不再打开。AI 偶尔出错，大家开始不信任。原本说好要配合试点的人，后来忙自己的事去了。IT 说正式接系统还要走流程。老板问用了多少次，你拿不出数据。

最后，这个项目没有明确失败，却也没有真正成功。

它只是慢慢没人用了。

这就是很多 AI 项目从 Demo 到落地时最常见的结局。

所以这一章要讲一个关键问题：

为什么 Demo 看起来很好，上线后却没人用？

答案通常不是 AI 不够强。

而是它没有真正进入工作流程。

Demo 成功，不等于项目成功

Demo 的目标，是证明“AI 能不能做”。

上线的目标，是证明“人会不会用”。

这两件事不一样。

Demo 可以在一个理想环境里完成。样本是你挑过的，流程是你设计好的，演示时间只有几分钟，错误可以提前避开，老板只看到最顺的一面。

但真实工作不是这样。

真实业务里，材料格式很乱，用户没有耐心，流程会被打断，有人会上传奇怪文件，有人会忘记点确认，有人会质疑 AI 的结果，也有人会担心出错后责任归谁。

还有人会觉得：

“我自己做也挺快，为什么还要多用一个工具？”

所以，AI 项目从 Demo 到上线，真正要解决的不是“它能不能跑”，而是：它能不能融入原来的工作方式，能不能让人更省事而不是更麻烦，出错时有没有办法处理，结果能不能被信任，有没有人持续负责。

如果这些问题没解决，Demo 再漂亮，也很难长期有人用。

很多 AI 上线失败，是因为增加了工作量

这是第一个常见原因。

你以为 AI 是来减负的。

但业务人员实际感受到的是：我又多了一个系统要打开，我又多了一个结果要检查，我还要复制粘贴资料进去，AI 错了我还要改，最后责任还是我的。

如果 AI 没有减少原有步骤，反而增加了新步骤，业务人员很快就会放弃。

比如客服邮件分类。

如果客服原来在邮箱里处理邮件，你做了一个 AI 工具，但要求客服每天先把邮件复制到另一个页面，再等 AI 分类，再把结果复制回表格。

这就很麻烦。

哪怕 AI 分类还不错，业务人员也可能不用。

因为它打断了原来的工作流。

再比如合同字段提取。

如果法务原来在 OA 里看合同，你让他先下载合同，再上传到 AI 工具，再复制结果到审批系统。

他可能会觉得：还不如我自己看。

所以，上线前要问一个很朴素的问题：

AI 到底减少了步骤，还是增加了步骤？

如果只是把工作换了个地方做，价值就会很弱。

真正好的 AI 上线方式，是尽量贴近原来的工作流。

用户不需要改变太多习惯。最好是在他已经使用的地方，看到 AI 的结果。

比如在邮件旁边显示分类建议，在合同审批页面显示字段提取结果，在巡店表单里自动生成报告草稿，在企业微信或钉钉里推送待确认事项。

AI 越贴近原来的工作位置，越容易被使用。

第二个原因：业务人员不信任它

AI 很强，但它会出错。

而且它有时候错得很自信。

这会让业务人员不安。

第一次错，大家会说“还可以理解”。第二次错，大家开始怀疑。第三次错，很多人就会回到原来的人工流程。

尤其是在合同、财务、客户沟通这类场景里，信任非常重要。

如果 AI 给出一个结果，但没有说明来源、理由或置信程度，业务人员就很难放心使用。

比如合同字段提取。

AI 说付款方式是“验收后 30 天内支付”。

但法务会想：它从哪里看到的？是正文里写的，还是附件里写的？有没有漏掉补充协议？如果它提错了，最后是谁负责？

如果这些问题没有答案，业务人员就会重新自己看一遍。

这样 AI 就没有节省多少时间。

所以，让业务人员信任 AI，不能只靠一句“模型很强”。

你需要让结果更可检查。

提取字段时，附上原文位置。分类邮件时，给出判断理由。生成报告时，标明依据来自哪些照片或检查项。不确定时，不要硬给答案，而是标记“需人工确认”。

这会让 AI 从一个“黑盒答案机器”，变成一个“可复核的助手”。

可信，才会被使用。

第三个原因：一开始就让 AI 做得太多

很多项目上线失败，是因为第一版就想做全自动。

比如 AI 自动回复客户，AI 自动审批报销，AI 自动判断合同风险并给出结论，AI 自动筛掉候选人，AI 自动创建任务并下发整改。

听起来很高效。

但对第一次上线来说，风险太大。

因为业务人员还没有建立信任，管理者也不知道 AI 的错误边界，IT 和合规可能也没准备好。

更稳的方式，是分阶段上线。

第一阶段，只读。AI 只负责看材料、提取信息、生成建议。

第二阶段，草稿。AI 可以生成邮件、报告、审核意见，但必须人工确认。

第三阶段，半自动。AI 可以创建待办、填写表单、发起流程，但关键节点仍然保留人工确认和抽检。

不要一开始就让 AI 替人执行。

先让 AI 帮人看，再让 AI 帮人写，最后才让 AI 帮人做一部分动作。

这条路径看起来慢，但更容易成功。

因为它让信任逐步建立。

第四个原因：没有明确谁负责

AI 上线后，一定会出现问题。

某个字段提取错了，某类邮件总是误判，某些文件上传失败，AI 生成的报告格式不符合要求，业务人员不知道怎么反馈，成本突然变高，使用量突然下降。

这时候，如果没有明确负责人，项目很快就会失控。

业务说：“这是技术的问题。”

技术说：“这是业务规则不清楚。”

老板说：“那到底谁来改？”

于是问题堆在那里，没人持续处理。

所以，AI 项目上线前，必须明确三个角色。

第一，业务负责人。

他负责判断 AI 结果是否可用，提供反馈，确认流程是否适合业务。

第二，技术或工具负责人。

他负责处理系统、工具、接口、账号、权限、日志等问题。

第三，项目负责人。

他负责统筹进度、记录问题、汇报效果、推动下一步。

小项目也需要负责人。

不是为了开更多会，而是为了让问题有人接。

没有负责人，就没有持续改进。

没有持续改进，AI 很快就会变成一个没人维护的工具。

第五个原因：没有使用数据

老板问项目效果时，很多人只能说：“大家反馈还可以” “应该有用” “业务觉得挺方便” “最近比较忙，还没怎么推广”。

这些话都不够。

上线以后，至少要记录几类基础数据：用了多少次，谁在用，每次处理什么任务，AI 输出是否被采纳，人工修改了多少，出错主要集中在哪里，有没有节省时间。

如果没有这些数据，项目就无法判断。

它到底是真的有用，还是只是大家客气一下？到底是没人知道，还是没人愿意用？到底是功能不行，还是流程不顺？

即使是大众版阶段的小项目，也要有最基本的记录。

不用复杂系统，Excel 表也可以，飞书多维表格也可以，工具后台日志也可以。

关键是要留下痕迹。

AI 项目上线以后，不能只看感觉，要看使用。

上线的核心原则：不动核心系统，从边缘切入

第一次上线 AI 项目，不要急着改企业核心系统。

不要一开始就改 ERP，不要一开始就改财务系统，不要一开始就改 OA 审批主流程，不要一开始就让 AI 操作关键数据。

更稳妥的方式是：

从边缘切入，先做辅助层。

什么叫边缘？

就是不直接改变核心业务结果的地方。

比如在合同审批前，先生成字段提取表；在客服正式回复前，先生成分类建议；在财务审核前，先生成异常提醒；在巡店报告提交前，先生成草稿；在会议结束后，先整理待办清单。

这些都属于辅助层。

它们不会直接替代原系统，也不会一上来改变最终决策。

这样上线阻力小很多。业务敢试，IT 更容易接受，风险也可控。

等这个辅助层真正证明有价值，再考虑进一步接入系统。

第一次上线，范围一定要小

很多人做出 Demo 后，马上想全公司推广。

这很危险。

第一个上线范围越大，问题暴露得越快，也越难收拾。

更好的方式是小范围试点。

不是所有客服都用，先选一个客服小组。不是所有合同都用，先选低风险合同。不是所有门店都用，先选 5 到 10 家门店。不是所有报销单都用，先选一个部门的差旅报销。

小范围试点有几个好处。

第一，反馈更快。你可以直接找到使用者，问哪里不好用。

第二，风险更小。即使出错，也不会影响全公司。

第三，改进更容易。流程还没定死，可以快速调整。

第四，更容易形成样板。当一个小团队真的用起来以后，再推广到其他团队，会更有说服力。

第一次上线，不要追求覆盖面，要追求可控性。

图8 三阶段上线图



不要跳过 v1，也不要一上来就追求全自动。上线的第一目标是可控。

我们可以把 AI 上线分成三个阶段。

第一阶段，只读。

AI 不做任何实际操作，只负责看资料，生成结果。比如从合同中提取字段，从邮件中判断类别，从简历中提取基础信息，从会议纪要中整理待办，从报销单中识别异常。

这一阶段的重点是建立信任。

第二阶段，草稿。

AI 可以生成内容，但不自动发出。比如生成客户回复草稿、合同审核意见草稿、巡店报告草稿、整改建议草稿、销售跟进邮件草稿。

人负责修改和确认。

第三阶段，半自动。

AI 可以开始触发一些低风险动作，比如创建待办、填写表单草稿、把邮件分配给对应小组、把异常记录加入表格、向负责人推送提醒。

但关键动作仍然要人确认。

不要跳过只读阶段，不要一上来就半自动。

如果只读阶段都没人用，半自动只会更危险。

上线前必须定义“AI 不能做什么”

很多项目只定义 AI 要做什么，却没有定义它不能做什么。

这会带来风险。

客服场景要明确：AI 不自动发送客户回复，不承诺退款，不处理法律威胁类邮件，遇到辱骂、维权、媒体曝光等关键词，必须转人工。

合同场景要明确：AI 不给最终法律意见，不判断合同是否可以签署，不替代法务审核，只做字段提取和差异提醒。

财务场景要明确：AI 不自动批准报销，不自动触发付款，不修改财务系统数据，只做异常提醒。

这些边界越清楚，项目越安全，也越容易被业务和管理层接受。

上线后，第一周最重要的是观察

很多人上线后，会马上等结果。

但第一周最重要的不是结果，而是观察。

你要看：谁在用，谁不用，为什么不用，哪个步骤最卡，用户有没有看 AI 的结果，他们有没有修改，AI 的错误集中在哪里，有没有人绕开系统回到原来的流程。

这些观察比后台数据更重要。

因为早期项目失败，往往不是模型问题，而是使用方式问题。

你可能发现，AI 分类效果不错，但入口太深，用户懒得点。报告草稿质量还行，但格式不符合公司模板。合同字段提取准确，但输出不方便复制到审批系统。客服回复草稿可用，但语气不符合品牌规范。

这些都是可以改的。

但前提是你要看见。

这一章读完，你可以马上做什么

1. 把你的 MVP 拆成 v1 只读、v2 草稿、v3 半自动三个阶段。
2. 确定第一个试点小组，不要一上来全公司推广。
3. 约定上线后一到两周的复盘时间，并提前准备试点复盘表。

本章小结

Demo 成功，不等于上线成功。

Demo 证明的是 AI 能跑，上线证明的是 AI 能被真实业务使用。

很多 AI 项目上线后没人用，原因通常不是模型不够强，而是增加了业务工作量、结果不够可信、一开始做得太自动、没有明确负责人、没有使用数据、没有失败处理方式。

第一次上线 AI，不要动核心系统，先从边缘切入。

不要全公司推广，先小范围试点。

不要一上来全自动，先只读，再草稿，最后半自动。

最重要的是：让 AI 成为业务人员愿意使用的助手，而不是一个增加负担的新系统。

只有这样，AI 才能真正进入工作流程。

下一章，我们完成大众版的最后一步：从第一个 AI 项目开始，成为 AI 落地工程师。

第八章 从第一个 AI 项目开始，成为 AI 落地工程师

如果你是从第一章一路读到这里，回头看，会发现自己已经走过了一条完整路径。

一开始，你可能只是被老板叫去“研究一下 AI”。

你不知道该先看模型，还是先试工具。

你担心自己不是工程师，不懂算法，也不会搭复杂系统。

你看到很多 AI 案例，但回到自己公司，还是不知道从哪里开始。

这很正常。

几乎所有第一次负责 AI 项目的人，都会经历这个阶段。

但现在，你已经不一样了。

你知道，AI 项目的第一步不是选工具，而是找到真实业务问题。

你知道，一个好场景要高频、重复、规则清晰、输入多样。

你知道，不是所有自动化都需要 AI，有些事情普通代码更合适，有些事情必须保留人工判断。

你知道，第一个项目不要一上来做完整系统，而是先做一个小 MVP。

你知道，Demo 之后要算 ROI，因为老板最终关心的是投入产出。

你也知道，Demo 成功不等于上线成功，AI 真正要进入工作流程，才算落地。

这意味着，你已经不只是“在研究 AI”。

你开始具备一种新的能力：

把 AI 变成业务结果的能力。

这就是 AI 落地工程师的起点。

你不需要一开始就成为专家

很多人一听“AI 落地工程师”，会觉得这个词离自己很远。

好像必须懂算法、懂模型、懂架构、懂代码、懂私有化部署，才配得上这个名字。

其实不是。

至少在第一阶段，不是。

AI 落地工程师不是从论文开始的，也不是从模型参数开始的。

而是从一个很普通的场景开始：客服每天都在重复分类邮件，财务每个月都在核对报销单，HR 每周都在筛简历，销售每天都在写跟进邮件，区域经理每次巡店都在整理报告。

你能发现这些工作里的重复部分，能判断哪一步适合 AI，能做出一个小验证，能让业务人员试用，能向老板讲清价值。

这就已经是 AI 落地工程师最核心的能力雏形。

专家不是一天长成的。

但项目是从第一步开始的。

AI 落地工程师不是“会用 AI 工具的人”

现在，会用 AI 工具的人越来越多。

会让 AI 写文案，会让 AI 生成图片，会让 AI 总结会议，会让 AI 改简历，会让 AI 写代码。

这些能力有用，但还不够。

企业真正需要的，不只是一个会用工具的人。

而是一个能回答下面这些问题的人：这个场景值不值得做？业务人员真的痛吗？每月处理量有多少？规则是否清楚？数据能不能拿到？AI 先帮哪一步？哪些动作必须人工确认？做完以后能省多少时间？上线后谁负责维护？出错了怎么办？

这些问题，比“哪个工具更好用”更重要。

因为工具会变，模型会变，平台会变。

但企业落地 AI 的基本逻辑不会变。

真正稀缺的能力，是把一个模糊想法变成一个可验证、可上线、可持续改进的项目。

从工具试用员，到项目负责人

第一次负责 AI 时，很多人会停留在“工具试用员”状态。

这个工具试一下，那个平台看一下，今天研究一个模型，明天收藏一个教程，后天整理一个工具清单。

这不算错。

了解工具是必要的。

但如果一直停在这里，就很难产生业务价值。

AI 落地工程师要完成一次转变：

从“我试过哪些工具”，转向“我推动了哪些业务改变”。

这两种汇报方式完全不同。

工具试用员会说：本周测试了 8 个 AI 工具，其中 3 个支持知识库，2 个支持工作流，1 个可以生成图片。

项目负责人会说：本周访谈了客服、财务和 HR 三个部门，找到 6 个候选场景；初步判断客户邮件分类和报销单异常检查最适合做第一个 MVP；已收集 50 封客户邮件，完成脱敏，并由客服主管标注标准答案；下周将验证 AI 分类准确率，并估算节省时间。

后者才像真正能推进事情的人。

老板也更愿意继续给这样的人资源。

第一个项目，不是终点，而是起点

大众版这本书，只帮你完成一件事：跑通第一个小型 AI 项目。

但一个项目成功，并不代表你已经完成了 AI 落地。

它只是说明，你已经知道怎么开始。

真正的成长，发生在第二个、第三个、第四个项目里。

因为第一个项目往往有很多偶然性。

可能业务方特别配合，场景刚好很简单，样本很干净，老板期待不高。

但当你开始做第二个项目时，问题会变复杂。

另一个部门的数据可能更难拿，另一个场景的规则可能更模糊，另一个流程可能涉及更多审批，另一个负责人可能没有那么配合，另一个项目的 ROI 可能没那么明显。

这时，你就不能只靠热情和工具。

你需要方法、标准、复盘。

你需要把第一个项目里的经验沉淀下来。

这就是从“做成一个项目”到“形成一种能力”的区别。

每做完一个 AI 项目，都要留下三样东西

如果你想真正成长为 AI 落地工程师，每做完一个项目，都不要只留下一个 Demo。

至少要留下三样东西。

第一，场景说明。

记录这个项目解决的是什么业务问题，为什么选它，它的频率、耗时、规则、风险是什么，AI 先做哪一步，人工保留哪一步。

当你以后做第二个项目时，这份说明会变成参考。

第二，评估结果。

记录用了多少真实样本，人工标准答案是什么，AI 正确多少，错误多少，错误集中在哪里，哪些错误可以接受，哪些错误必须处理。

AI 项目不是“感觉好用”就够了。

它必须可验证。

第三，复盘结论。

记录业务人员为什么愿意用，为什么不愿意用，MVP 哪部分最有价值，哪部分浪费了时间，上线前最容易忽略什么，下一次应该提前做什么。

不复盘，一个项目只是一个项目。

复盘后，一个项目才会变成能力。

AI 落地工程师的三条成长路线

不同背景的人，成长路径不一样。

业务型的人，优势是懂业务。

如果你来自运营、管理、咨询、财务、人事、销售、客服等岗位，你知道哪里最痛，也知道一线员工真正怎么工作。

你的成长重点是熟悉常见 AI 工具，学会低代码平台，理解输入、输出、样本、评估这些基本概念，掌握简单 ROI 计算，学会把业务流程拆成 AI 可以处理的小动作。

产品型的人，优势是会拆需求。

如果你是产品经理、项目经理、业务分析师，你习惯把模糊问题拆成用户、场景、流程、功能、指标和验收标准，也比较擅长跨部门沟通。

你的成长重点是理解 AI 能力边界，学会做小样本评估，学会错误分类，学会设计人工在环，学会控制上线范围。

技术型的人，优势是能实现。

如果你来自 IT、开发、数据、自动化、系统运维等岗位，你懂接口、权限、系统和数据流，也更容易理解模型 API、工作流、部署、日志等技术问题。

你的成长重点是学会业务访谈、场景优先级判断、ROI 计算、向非技术人员解释风险和价值，以及克制自己不要什么都用复杂技术。

不管你是哪一类，都可以进入 AI 落地这条路。

只是补课重点不同。

图9 30 天成长路线图

• 从被老板点名，到完成第一个 AI 小项目 •



30 天不要求你成为专家，只要求你跑通第一个小闭环。

如果你现在要继续往前走，可以用一个 30 天计划。

这个计划不追求你成为专家，只追求一件事：

让你能独立推动一个小型 AI 项目从想法走到验证。

第 1 周：找到一个场景。

访谈 3 个业务人员，收集 5 到 10 个候选场景，用场景发现表做初步筛选，选出 1 个最适合 MVP 的场景。

产出是一页纸项目启动说明。

第 2 周：做出一个 MVP。

准备 10 到 50 条真实样本，完成脱敏，定义输入和输出，让业务人员标注标准答案，用现成 AI 工具或低代码平台跑通流程。

产出是一个可演示 MVP。

第 3 周：评估效果。

对比 AI 输出和人工标准答案，记录正确率、错误类型、处理耗时，找业务人员试用，收集反馈。

产出是一份简单评估报告。

第 4 周：算 ROI，做汇报。

计算当前人工耗时，估算 AI 后节省时间，估算工具和开发成本，列出风险和上线边界，准备下一步建议。

产出是一页纸项目汇报。

结论可以是继续扩大样本、进入小范围试点、先补充规则、暂时不建议做。

敢给出判断，比盲目推进更重要。

成为 AI 落地工程师，需要长期补四门课

大众版到这里，已经帮你完成了第一步。

接下来，如果你想继续做深，需要系统补四门课。

第一门课，业务审计。

你要学会更系统地进入业务现场，找到真正值得做的场景，包括怎么访谈、怎么画流程、怎么找瓶颈、怎么评估优先级、怎么判断价值和风险。

第二门课，效果评估。

你要学会证明 AI 到底做得好不好，包括怎么准备黄金数据集、怎么设计标准答案、怎么统计正确率、怎么分类错误、怎么判断哪些错误能接受、怎么与人工基线对比。

第三门课，部署上线。

你要学会让 AI 安全进入生产流程，包括权限、日志、人工确认、失败处理、数据脱敏、系统接入、小范围试点、上线复盘。

第四门课，运营和复制。

你要学会让 AI 项目持续产生价值，包括成本监控、使用数据、错误复盘、模型和工具调整、模板沉淀、跨部门复制、组织协作。

这就是商业版会继续展开的内容。

大众版让你敢开始。

商业版会帮助你做深、做稳、做成体系。

一个成熟的 AI 负责人，敢说三句话

随着经验增加，你会发现，一个真正成熟的 AI 负责人，不是永远说“可以做”。

他要敢说三句话。

第一句：这个适合先做。

当一个场景高频、重复、规则清晰、样本容易拿到、风险可控，你要敢于推动。

很多 AI 项目不是死于失败，而是死于迟迟不开始。

第二句：这个现在不适合做。

当一个场景规则不清、风险太高、数据拿不到、业务没人配合，你要敢于暂缓。

AI 项目不是越多越好。选错场景，比不做更糟糕。

第三句：这个可以做，但只能做到这一步。

合同可以做字段提取，但不做终审。报销可以做异常提醒，但不自动付款。客服可以做回复草稿，但不自动承诺赔偿。HR 可以做简历信息提取，但不自动淘汰候选人。

AI 落地工程师的专业性，很多时候就体现在边界感上。

你知道 AI 能帮什么，也知道它不能越过哪里。

这一章读完，你可以马上做什么

1. 建立自己的“AI 落地档案”，把第一个项目的场景、样本、测试结果和复盘记录下来。
2. 用 30 天行动表安排下一个小项目，重点不是更复杂，而是更可复制。
3. 选一个你最需要补的方向：业务审计、效果评估、部署上线，或运营复制。

最后的话：从一个小项目开始

AI 听起来很大。

大模型、智能体、自动化、数字员工、企业智能化，每个词都很宏大。

但真正的落地，往往从一个很小的地方开始。

一封客户邮件。

一张发票。

一份合同。

一份简历。

一张巡店照片。

一段会议记录。

一个重复到让人厌烦的动作。

你不需要一开始改变整个公司。

你只需要先找到一个真实问题，让 AI 在其中帮上一点忙。

然后验证，复盘，改进，再复制。

这就是 AI 落地最朴素，也最可靠的路径。

一开始，你只是那个被老板叫去“研究一下 AI”的人。

但当你能找到问题、做出验证、算清价值、控制风险、推动上线时，你已经开始进入一个新的角色。

企业需要的，从来不只是会写 Prompt 的人。

而是能把 AI 放进真实工作里，让它持续创造价值的人。

这，就是 AI 落地工程师。

附录：你的第一个 AI 项目工具包

下面这些模板，可以直接复制到 Word、飞书、Notion、Excel 或项目文档里使用。它们对应本书的完整路径：找问题、选场景、做 MVP、算价值、小范围上线、复盘成长。

附录一：AI 项目启动三问

- 1. 公司里哪个流程每天、每周或每月都在重复发生？
- 2. 这个流程现在谁最痛苦？
- 3. 如果 AI 只能先帮一步，最应该先帮哪一步？

附录二：一页纸启动模板

问题	填写内容
业务场景	例如：客户邮件分类
当前负责人	例如：客服主管 / 销售助理
当前痛点	例如：每天邮件太多，人工分拣慢
月处理量	例如：每月 3000 封
单次耗时	例如：每封 1 分钟
现有规则	例如：分为投诉、咨询、售后、合作
AI 先做哪一步	例如：自动分类并生成处理建议
人工保留哪一步	例如：客服确认后再回复
成功标准	例如：分类准确率达到 85%，节省 50% 分拣时间
需要谁配合	例如：客服主管提供 100 封历史邮件

附录三：AI 场景发现表

部门	候选场景	每月发生次数	每次耗时	是否重复	规则是否清楚	输入是否多样	是否有人复核	备注
销售	客户邮件分类	3000 次	1 分钟	是	是	是	是	可先做分类
财务	报销单检查	2000 次	3 分钟	是	是	是	是	可先检查金额和抬头
HR		500 份	5 分钟	是	是	是	是	

部门	候选场景	每月发生次数	每次耗时	是否重复	规则是否清楚	输入是否多样	是否有人复核	备注
	简历信息提取							可先提取字段
采购	报价单比对	300 份	10 分钟	是	部分清楚	是	是	需明确规则
法务	合同终审	100 份	30 分钟	是	部分清楚	是	必须	不做终审，只做初筛

附录四：AI 适用性三问

问题	如果答案是“是”	如果答案是“否”
规则能不能说清楚？	可以继续评估	先梳理规则
输入是不是经常变化？	更适合 AI	可能用普通代码即可
出错后是否可控？	可以做试点	只能辅助，不能自动执行

附录五：AI / 代码 / 人工判断表

任务	规则是否清楚	输入是否多变	错误风险	推荐方式
客户邮件分类	是	是	低	AI
发票金额比对	是	否	中	普通代码 + 人工复核
合同关键字段提取	是	是	中	AI + 人工复核
大额付款审批	是	否	高	人工确认
简历字段提取	是	是	中	AI
最终录用决定	否	是	高	人工
库存低于阈值提醒	是	否	低	系统规则
巡店报告草稿	是	是	低	AI + 人工修改

附录六：AI MVP 定义表

问题	填写内容
MVP 名称	例如：客户邮件自动分类 MVP
业务使用者	例如：客服主管、客服专员
当前痛点	每天人工分拣邮件耗时
AI 先做哪一步	判断邮件类型并给出建议处理人
输入材料	邮件标题、正文

问题	填写内容
输出结果	类别、紧急程度、判断理由、建议处理人
样本数量	50 封历史邮件
标准答案	客服主管人工标注
成功标准	分类准确率 $\geq 85\%$ ，人工认可可用于初筛
人工保留动作	客服确认分类后再处理
暂不做什么	不自动回复客户，不自动创建工单

附录七：MVP 测试记录表

样本编号	人工标准答案	AI 输出	是否正确	错误类型	备注
001	投诉	投诉	是	-	判断理由准确
002	售后	投诉	否	类别混淆	客户语气强烈
003	产品咨询	产品咨询	是	-	可用
004	商务合作	无关信息	否	识别错误	邮件内容较短
005	投诉	需人工确认	部分正确	不确定	可接受

附录八：简版 ROI 计算表

项目	计算方式	示例
月处理量	每月发生多少次	400 份
单次人工耗时	原来每次花多久	25 分钟
当前月人工耗时	月处理量 $\times$ 单次人工耗时	166.7 小时
人员小时成本	月综合成本 $\div$ 月工作小时	60 元/小时
当前人工成本	当前月人工耗时 $\times$ 人员小时成本	10000 元
AI 后单次人工耗时	AI 辅助后人工复核时间	8 分钟
AI 后人工耗时	月处理量 $\times$ AI 后单次人工耗时	53.3 小时
节省人工时间	当前耗时 - AI 后耗时	113.4 小时
节省时间价值	节省时间 $\times$ 小时成本	6800 元
AI 月成本	工具费 + 调用费	800 元
开发月分摊	开发成本 $\div$ 分摊月数	2000 元
月净收益	节省时间价值 - AI 成本 - 开发分摊	4000 元

附录九：上线前检查清单

检查项	是否完成
是否只选择了一个小范围试点？	☐ 是 ☐ 否
是否明确 AI 只做哪一步？	☐ 是 ☐ 否
是否明确 AI 不能做什么？	☐ 是 ☐ 否
是否保留人工确认？	☐ 是 ☐ 否
是否有业务负责人？	☐ 是 ☐ 否
是否有技术或工具负责人？	☐ 是 ☐ 否
是否有问题反馈渠道？	☐ 是 ☐ 否
是否记录输入、输出和人工修改？	☐ 是 ☐ 否
是否定义了高风险内容转人工规则？	☐ 是 ☐ 否
是否先用脱敏或低风险数据试点？	☐ 是 ☐ 否
是否明确试点周期？	☐ 是 ☐ 否
是否约定复盘时间？	☐ 是 ☐ 否

附录十：三阶段上线表

阶段	AI 做什么	人做什么	风险水平	适合目标
v1 只读	提取、分类、摘要、提醒	查看和判断	低	建立信任
v2 草稿	生成回复、报告、审核意见	修改和确认	中	提升效率
v3 半自动	创建待办、填写表单、分配任务	抽检和兜底	中高	接入流程

附录十一：30 天 AI 落地行动表

时间	目标	关键任务	产出
第 1 周	找场景	访谈业务、收集候选场景、选择最小任务	一页纸项目说明
第 2 周	做 MVP	准备样本、定义输入输出、搭建原型	可演示 MVP
第 3 周	做评估	对比标准答案、记录错误、收集反馈	简单评估报告
第 4 周	做汇报	计算 ROI、定义风险边界、提出下一步	项目汇报页

版本说明

本稿为《老板让我负责 AI，我该怎么办？》大众版 V1.3 正式排版版。

本版本完成了：

- 在 V1.2.2 图片版基础上完成 PDF 排版；
- 增加封面、目录、页眉页脚和统一版式；
- 正式嵌入 9 张核心图示；
- 保留每章行动页和“第一个 AI 项目工具包”附录；
- 适配 PDF 小册子、公众号连载、飞书文档和课程讲义。

后续适合继续推进：

- V1.4：公众号连载版，按章节拆分并增加每篇开头钩子；
- 模板包：将附录表格拆成 Excel、飞书表格或 Notion 模板；
- 商业技术书：在大众版基础上展开 Audit、Evaluation、Deployment、Operation、Scale。