

为什么需要 Skill Engineering



让一次成功的任务，沉淀为可复用、可验证、可迭代的团队能力

当前现象

已经开始使用 Qoder

不少同事已把 AI 用到日常工作中

数据处理是主要场景

整理明细、汇总报表、解释结果

主要仍通过 Chat 完成工作

每次重新说明需求，再反复调整

成功经验尚未稳定沉淀

换一个 Chat，往往又要从头解释



下一步目标

感知 Skill 的作用

看见流程和规则复用带来的价值

一次性完成成为可复用能力

把成功方法留在完整的能力包中

个人经验成为团队共同方法

让另一位同事也能按同一口径完成

形成共同的工程语言

知道如何设计、执行、验证和改进

把 Chat 中成功的方法沉淀为 Skill，再通过验证与复用完善它。

从一次数据处理任务说起

收到客户的一份工单明细，需要生成每日运营报表。

同样的日报，每次都在重复这些说明

要说明什么	本次规则
统计范围	10 月 10 日创建的工单
去重方式	按工单 ID；内容冲突时停止
统计口径	完成、待处理、完成单平均耗时

教学示例：5 单，完成 3 单、待处理 2 单；完成单平均耗时 65 分钟。

下一次提供新数据与统计日期，沿用同一流程和口径，再核对结果。



一次 Chat

- 解释需求
确认日期、字段和规则
- 反复调整
检查去重与计算是否正确
- 交付一次
方法仍留在这次对话里

留下 Skill

- 流程说明
规定输入检查与交付顺序
- 业务规则
固定去重、日期和统计口径
- 处理与校验
保存可调用的计算脚本

Chat、Skill、Skill Engineering



把一次成功的任务，转化成能跨对话复用、执行关键操作、具备验证机制的工作能力。

Skill Engineering = Skill Design + Deterministic Execution + Validation + Evaluation + Reuse

技能设计

确定性执行

结果验证

效果评估

复用传播

Chat 使用

单次完成任务

提出问题，获得答案
解决当下的任务



Skill 使用

复用已有能力

调用已沉淀的流程和资源
按约定方法完成任务



Skill Engineering

工程化沉淀与持续优化

分清规则、实现与验证
换对话、换数据、换人复用

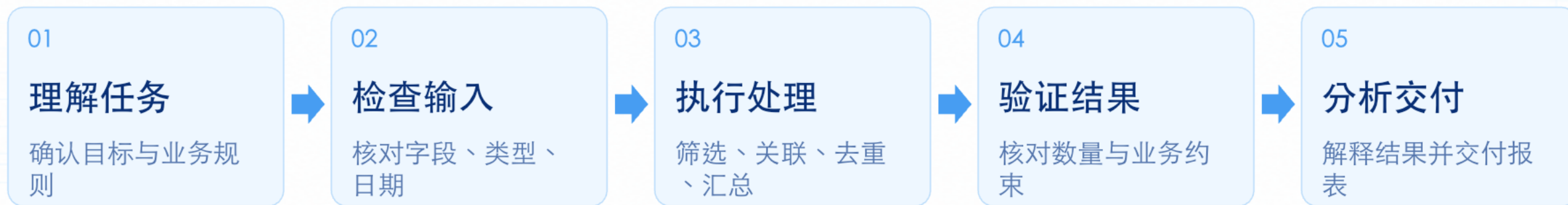
更稳定：新 Chat 也能复现。更可靠：关键计算可验证。更高效：团队可复用、可传播。

一个数据处理 Skill 的两个部分



Workflow 安排任务怎样推进，Data Processing 完成数据的解析、转换与计算。

Workflow：任务怎样推进



Workflow（工作流）

决定步骤顺序、如何处理歧义、何时继续、何时停止。

Data Processing（数据处理）

按确认的规则完成运算，并返回结果、异常与校验信息。

Workflow 和数据处理都可能需要推理，也都可能需要确定性实现。

AI 推理 vs 确定性执行



在工作流与数据处理两个层次上，分别确定 AI 和固定代码的职责。

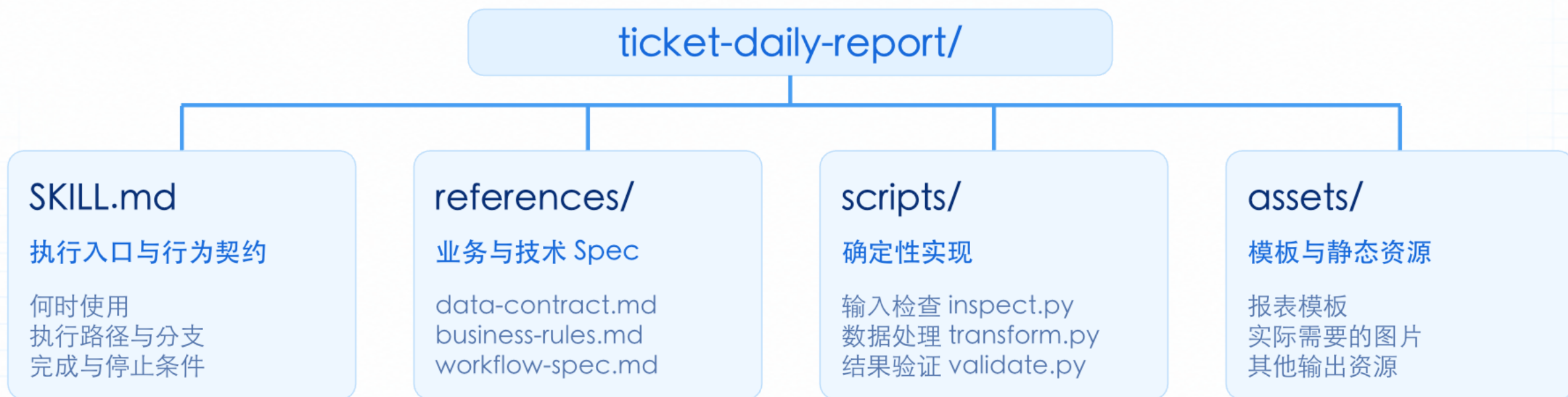
工作层次	AI 推理负责	确定性实现负责
Workflow 工作流	识别任务意图 判断异常原因 选择处理策略	固定步骤顺序 状态检查 必经校验、停止条件
Data Processing 数据处理	理解字段含义 识别非标准格式 解释异常	筛选、关联、去重 汇总、金额计算 格式验证

让 Agent 处理不确定性，让代码处理确定性，让验证机制决定结果是否可信。

一个好的 Skill 长什么样



SKILL.md 保留入口与行为契约，详细 Spec 与 Implementation 按职责分开。



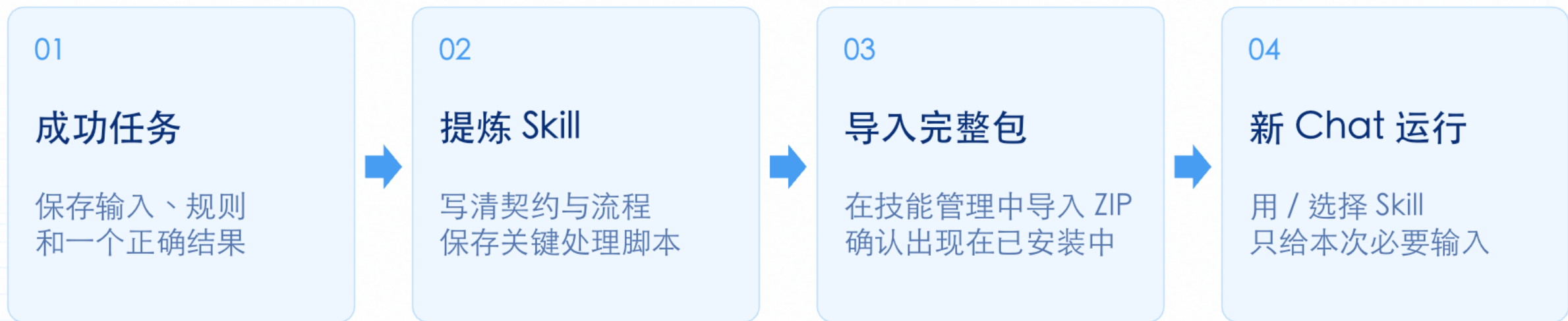
按需扩展：tests/ 验证程序，evals/ 验证 Agent 的完整行为。

主文件简洁，细节按需加载（Progressive Disclosure）。简单任务可以只有 SKILL.md。

Qoder 实操：沉淀一个 Skill



以已经做对的工单日报为起点，把任务方法整理成能独立执行的能力包。



给 Qoder 的提示词：把刚才成功的数据处理沉淀为

Skill：写清输入、业务规则、输出和停止条件；关键计算保存成脚本；不要依赖这次聊天历史。

先手动选择验证执行，再用自然语言验证触发。Qoder 操作入口按当前产品版本核对。

如何证明 Skill 可靠

跨新 Chat 的可重复性，是 Skill Engineering 的第一道门槛。



Discovery

能否找到正确的 Skill

新 Chat 中手动选择
再测试自然语言触发



Execution

能否独立完成任务

不依赖历史聊天
按契约调用已有脚本



Verification

能否证明结果正确

核对关键结果
失败时明确停止

正常输入

结果与独立预期一致

可处理的变化

换日期、字段顺序或数据

应停止的异常

缺字段、ID 冲突不补造

程序 Tests 检查计算；Agent Evals 检查实际选择、执行与停止行为。最后请另一位同事使用同版本再试。

开始实践与问题反馈

把日常工作中的一个成功方法，变成下一次可以继续使用的 Skill。



今天就可以开始的一次实践

- 01 选一件重复工作
规则清楚，能确认正确结果
- 02 完成三类验证
正常、变化、应停止的异常
- 03 换人复用
另一位同事使用同版本再试

问题反馈

遇到问题，联系张坚、李琦

脱敏输入样例
预期结果与实际结果
出错截图与校验记录

业务同事沉淀方法与验收标准，技术同事协助固定关键处理，团队共同验证和改进。