

用 AI 放大每个人的专业能力

更好的方法 · 更强的团队 · 更大的可能

Skill Engineering 培训系列

从 Chat 使用者, 到 Skill 工程实践者

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

PEOPLE
PROCESS
SKILLS
BETTER TOGETHER

从尝试
到沉淀
让 AI 真正创造持续价值

IDEA
EXECUTION
VALIDATION
REUSE



核心公式

TURN SUCCESSFUL TASKS INTO LASTING TEAM CAPABILITIES

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

带来的价值

MORE STABLE · MORE RELIABLE · MORE EFFICIENT



适用于 Qoder / QoderWork 场景下的数据处理、 workflow 复用与团队培训

A MORE CAPABLE TEAM
WITH AI

沉淀方法 · 复用能力 · 共建未来

让 AI 能力, 成为团队的长期竞争力

出品: AI 经营交付效能专项组

为什么需要 Skill Engineering

Skill Engineering 自学课件

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。

出品：AI 经营交付效能专项组02 / 10

本页讲解

比如，今天花了几轮对话把日报做好，明天换个 Chat，同样的日期、去重和平均耗时又得重新解释。我们可以把今天用过的数据、确认的规则和核对过的结果留下来，再请 AI 把处理方法整理成 Skill。以后拿到同类数据，就有一套现成的方法可以继续用。交给同事时，他也能顺着这些资料看懂这份日报是怎么算出来、怎么验过的。

继续阅读与实践：点击打开在线深入阅读

出品：AI 经营交付效能专项组

02 / 10

从一次数据处理任务说起

Skill Engineering 自学课件

从一次数据处理任务说起

拿到下面这份工单明细，怎样让 AI 把它整理成 10 月 10 日的日报？

输入：7 行虚构工单明细

工单 ID	创建日期	状态	耗时 / 分钟
T001	2026-10-10	已完成	60
T002	2026-10-10	待处理	0
T003	2026-10-10	已完成	45
T004	2026-10-10	已完成	90
T004	2026-10-10	已完成	90
T005	2026-10-10	待处理	0
T006	2026-10-11	已完成	30

T004 重复一行；T006 属于 10 月 11 日。

这份明细要变成一份日报：去重、按创建日期筛选，再统计数量与完成单平均耗时。



交给 AI 的任务

请统计 10 月 10 日创建的工单。相同 ID 完全一致时保留一条，日期、状态或耗时不一致时停下来等我确认。告诉我总数、完成数、待处理数和完成单平均耗时。

应该交回的日报

5 单：完成 3 单，待处理 2 单
完成单平均： $(60 + 45 + 90) \div 3 = 65$ 分钟

出品：AI 经营交付效能专项组

03 / 10

本页讲解

这 7 行就是我们交给 AI 的原始明细。T004 出现了两次，同一张工单只能算一次；T006 是 10 月 11 日创建的，也不该进入 10 日的日报。处理后剩下 5 张工单，其中 3 张已完成、2 张待处理。完成的 3 张分别用了 60、45、90 分钟，加起来 195 分钟，除以 3，平均就是 65 分钟。我们希望 AI 把明细整理成这样的日报，并让人能核对它怎么算的。把这些规则保存成 Skill，下次拿到新数据，就有了可以沿用的方法。

继续阅读与实践：点击打开在线深入阅读

Chat、Skill、Skill Engineering

Skill Engineering 自学课件

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。

出品：AI 经营交付效能专项组

04 / 10

本页讲解

刚才那份日报，在一个 Chat 里做对，就完成了当次任务。把规则和处理脚本保存成 Skill，下一次便有了可调用的方法。不过，另一位同事拿到新数据，也能做对吗？还要实际试一试。我们讲的 Skill Engineering，就是在沉淀方法之后，继续检查换对话、换数据、遇到异常时会发生什么。这样才知道，保存的方法离开原来的聊天还能不能工作。

继续阅读与实践：点击打开在线深入阅读

出品：AI 经营交付效能专项组

04 / 10

一个数据处理 Skill 的两个部分

Skill Engineering 自学课件

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

01
理解任务
确认目标与业务规则

02
检查输入
核对字段、类型、日期

03
执行处理
筛选、关联、去重、汇总

04
验证结果
每次核对，失败停止

05
分析交付
解释结果并交付报表

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 安排执行与停止；每次运行通过结果校验后，再分析交付。

出品：AI 经营交付效能专项组

05 / 10

本页讲解

算出 65 分钟之前，还有不少判断要做。统计哪一天、字段齐不齐、重复工单能不能合并，都影响最后的结果。假如同一个 ID 的两行耗时不同，就要停下来弄清原因，不能悄悄选一个。图里的 Workflow 管这些推进和停止的决定，Data Processing 按确认的规则去重、筛选和计算。两部分接起来，结果验过了，才把日报交出去。

继续阅读与实践：点击打开在线深入阅读

出品：AI 经营交付效能专项组

05 / 10

AI 推理 vs 确定性执行

Skill Engineering 自学课件

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。

出品：AI 经营交付效能专项组

06 / 10

本页讲解

拿刚才的平均耗时来说，195 除以 3，适合交给固定脚本，下一次也按同一规则计算。AI 则可以帮助理解需求：客户说的日期是创建日期，还是完成日期？字段名称换了，含义是否相同？把这些问题说清楚之后，再交给工具执行。脚本也可能写错分母，所以我们还要拿已知结果的样例核对，不能因为用了代码就省掉验证。

继续阅读与实践：点击打开在线深入阅读

一个好的 Skill 长什么样

Skill Engineering 自学课件

一个好的 Skill 长什么样

参考工单 Skill：入口、规则、实现、测试和评估场景各有用途。

ticket-daily-report/

SKILL.md

执行入口与行为契约

使用条件与输入
执行路径
完成与停止条件

references/

业务规则与数据契约

data-contract.md
business-rules.md

scripts/

确定性实现

pipeline.py
检查、计算、校验

tests/

程序测试与共用样例

fixtures/
expected/
check-regression.py

evals/

AI 任务与核对标准

evals.json
README.txt

按需增加 assets/ 存放模板；evals.json 保存场景，当前没有自动评估程序。

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

出品：AI 经营交付效能专项组

07 / 10

本页讲解

参考包的入口 SKILL.md 说明何时使用、怎样完成和停止；业务规则在 references，处理代码在 scripts/pipeline.py。tests 保存程序检查、输入样例和已确认答案；evals/evals.json 保存让 Agent 完成的任务与核对项，直接引用 tests 的样例和答案。它可以由人或负责评审的 LLM 读取。任务实际执行后，再把输出和工具执行记录交给评审方；仅有配置文件不会产生评估结果。当前参考包没有自动评估程序，tests/check-regression.py 只检查固定脚本。简单任务仍可只有 SKILL.md，模板等资源按需放 assets。

继续阅读与实践：点击打开在线深入阅读

出品：AI 经营交付效能专项组

07 / 10

Qoder 实操：沉淀一个 Skill

Skill Engineering 自学课件

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。



- 01 成功任务**
保存输入、口径和核对过的结果
- 02 提炼 Skill**
写清契约与流程
保存关键处理脚本
- 03 导入完整包**
在技能管理中导入 ZIP
确认出现在已安装中
- 04 新 Chat 运行**
用 / 选择 Skill
只给本次必要输入

给 Qoder 的提示词：把刚才做对的日报方法整理成完整 Skill ZIP，保存规则和处理脚本，让新 Chat 只给数据、日期和输出位置就能复用。

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。

出品：AI 经营交付效能专项组

08 / 10

本页讲解

先在普通 Chat 中做对日报，再把规则 and 实际使用的脚本整理成 Skill ZIP，安装后开新 Chat。第一次还没保存规则，需要把业务口径说明白；安装后只需说“用 normal.json 生成 2026-10-10 的工单日报，保存到新的结果文件夹”。每次都重讲算法，就无法看出 Skill 是否真的保存了方法。先用 / 选择技能检查执行，再不选择技能，用自然语言试它能否自动找到。

继续阅读与实践：点击打开在线深入阅读

如何验证并改进 Skill



本页讲解

工单跟练先验证结果与安装复用。更完整的 Task Evals 使用固定的真实任务和相同 Tests 比较版本：10 个 Golden Tasks 覆盖登录、支付、权限、超时和异常处理；假设 V1 完成 6 个、V2 完成 8 个，成功率从 60% 到 80%，提升 20 个百分点，同时比较质量、耗时、Token 和失败原因。数字仅为示例，需重复运行并控制模型、工具、数据和环境。评分可用程序、人工或 LLM。例如，参考包正常任务完成后，另开评审 Chat，提供 evals.json 的场景 1、已确认答案、实际日报及工具执行记录，请它逐项给出结论和证据。日报为 65，但没有实际调用日志时，调用项应为 UNKNOWN；日报为 39 时，数字项为 FAIL。只让 LLM 核对已有材料不会重跑任务，当前没有自动评估程序。Trigger 看正确触发，Production 持续观察真实质量、成本与失败。

继续阅读与实践：点击打开在线深入阅读

开始实践与问题反馈

Skill Engineering 自学课件

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 选任务并自评

规则清楚，先确认验收依据

02 生成并运行检查

共用样例，保留实际结果与记录

03 换人复用并迭代

同版本再试，把新问题补进用例

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。

出品：AI 经营交付效能专项组

10 / 10



本页讲解

从一件重复工作开始。已有 Skill 时，先用只读自评定位缺口；没有 Skill 时先完成前面的四步。准备完整 Skill、输入样例和已确认答案，让 AI 先列检查计划，逐案核对后再确认。随后生成适配自己脚本接口与输出的 tests 和 evals，两者共用包内样例与答案。按使用说明实际运行，保留命令、退出码、结果和失败记录，复制完整包到新目录也应能检查。检查程序同样可能写错，显示 PASS 后仍要确认它读了正确标准。Agent 任务另开 Chat 执行并留证，再由核对者评审。未确认口径保持待确认，不能为通过检查改答案。分享时请同事用同版本复测，把脱敏输入、预期与实际和执行记录交给张坚、李琦。

继续阅读与实践：点击打开在线深入阅读

出品：AI 经营交付效能专项组

10 / 10