

# 王杨

FDE 方向 · AI Coding 与复杂系统交付

男 | 10年工作经验 | 北京 | 在职 | 13520882907 | cowboy231bb@gmail.com

wangmouren.online | github.com/cowboy231

## 核心技能

**需求与业务建模**：从现场动作、业务规则与系统约束拆解实体、状态、不变量、异常路径和验收条件；具备临床标本库流程建模、基金业务规则与上下游字段口径统一实践。

**AI Coding与变更治理**：临床项目使用 Codex + MiniMax 辅助全栈实现、回归、排障和文档同步；通过小粒度变更、业务不变量和证据复核控制质量，由本人负责业务判断及真实写操作授权。

**全栈与系统集成**：Vue 3、Python / FastAPI、SQLite、RabbitMQ；统一上游数据契约并适配设备链路，区分业务记录、本地过程与物理结果；具备 Windows 内网交付、HTTP / WebSocket 与 Docker 工程实践。

**验证与异常恢复**：Mock、pytest、Robot Framework、SQL 数据核对与自动巡检；结合设备终态、任务快照、库存回读和实物核对验证流程，保留异常原因并设计受控恢复；按离线、现场及历史证据分别判断。

**技术表达与协作**：将复杂规则整理为方案、测试矩阵、Wiki 和操作指南；具备上下游口径协调、测试团队协作、新人培养经验，设计并迭代 17 周自动化课程，累计覆盖 300+ 学员。

## 工作经历与个人实践

临床标本库自动化系统 · AI Coding 全栈实践

个人项目实践 | 2026-07 - 至今

商汤科技 AI研究院

AI质量工程师 / AI驱动测试工程师 | 2025-12 - 至今

积家科技-外贸自研摄像头项目

高级测试工程师 | 2024-07 - 2025-10

九坤私募基金管理系统

高级测试工程师 | 2023-08 - 2024-06

小米 CRM / RMS / Dynamics 365

高级测试工程师 | 2022-08 - 2023-08

天眼查 OpenAPI / 专业版

高级测试开发工程师 / 测试组长 | 2021-02 - 2022-08

自动化测试讲师 ( 码同学 )

自动化测试讲师 ( 后期兼职收尾 ) | 2020-01 - 2021-03

马上消费金融

测试工程师 -> 测试组长 | 2016-11 - 2019-08

搜狗输入法

安卓测试工程师 | 2015-06 - 2016-08

2021-02 入职天眼查后，以兼职方式完成码同学剩余授课至 2021-03。

## 重点项目

---

### 临床标本库自动化系统 · AI Coding 全栈实践

个人项目 · 现场需求、业务建模与系统集成 | 2026-07 - 至今

- 从现场取放、扫码、核对与异常处理出发，梳理业务对象、动作顺序、上下游依赖及验收条件，将现场需求转为可执行流程。
- 建立样本实例、物理管、冻存盒、业务分区、任务与事件模型；区分上游业务记录、本地绑定与过程记录、设备物理结果，明确各层事实的来源。
- 使用 Codex + MiniMax 辅助实现 Vue 3 前端、FastAPI 服务、SQLite 数据模型与 RabbitMQ 适配；归一上游样本契约，连接喷码、整板扫描及自动化冰箱链路，提供 Windows 内网解压运行形态。
- 长任务以设备成功终态和可解析的实际结果判断完成；设备事件先持久化再确认，无法安全关联的事件留待核对，避免覆盖已有物理事实。
- 异常恢复保留原任务、原位置和设备原因，在同一业务分区内受控重新下发；通过 Mock、自动回归、库存回读与实物核对分层验证。
- AI 辅助实现、回归、日志排障和文档同步；本人负责业务判断、证据可信度、变更粒度及真实写操作授权。扫码器现场记录覆盖四个接口、95 管回包，发现六个协议漏写字段。
- 2026-08-24 离线快照：后端 1051 passed、7 skipped，前端 92 passed，Vite 构建成功；设备与组合流程按版本另行核验。

### 九坤私募基金 · 复杂业务建模与验证

高级测试工程师 · 业务规则、数据契约与场景化 Mock | 2023-08 - 2024-06

- 在真实生产数据不可用的保密约束下，梳理基金业务上下游依赖，将实体、字段和规则关系拆解为可执行的验证场景。
- 建立场景化 Mock 数据与 SQL 数据集，推动上下游统一字段定义和取值口径，减少业务规则与测试输入之间的歧义。
- 重构 Robot Framework 参数化与 BDD 脚本，将验证范围延伸至基金全生命周期的业务规则与数据一致性。

### 商汤 AI 测试资产 · Agent 执行与证据复核

AI 质量工程 · 将迭代需求转为可执行任务 | 2025-12 - 至今

- 在开发使用 AI 编码的快速迭代中，将测试目标、API 和数据库约束、预期结果及证据要求组织为 Agent 可执行任务。
- 接口测试和数据库核对以 Agent 执行为主；使用 AI 做首轮冒烟，本人复测异常与关键路径，核对请求、响应、数据及日志后形成业务结论。
- 将复核过的规则、方法、用例与执行记录沉淀到数字测试资产库，区分当前规则、历史观察与待裁决提案，为后续复跑和问题追溯保留依据。

## 更多项目实践

---

### 天眼查数据质量治理

高级测试开发工程师 / 测试组长 | 2021-02 - 2022-08

- 面向多实体 ToB 数据产品与 OpenAPI 梳理测试边界、异常分类及数据准确性风险，关注接口成功但业务数据不正确的问题。
- 建设自动巡检、数据准备与飞书告警套件，沉淀异常根因分类方法；承担新人培养、任务协调与测试团队协作。

### 企业席位测试建模与证据审计

商汤科技 · 身份、权限与权益一致性 | 2026-06 历史项目

- 梳理手机号、用户身份、企业、批次、席位、套餐与权益的关系，依据实际身份核对权限和权益归属，明确状态转换与异常边界。
- 以测试矩阵关联 API、数据库和页面证据；历史报告记录 111 条用例：81 PASS、18 WARN、8 FAIL、4 BLOCKED，按状态保留未通过原因。
- 对有状态操作记录执行前后快照及清理结果；无待执行项不代表全部通过，历史状态也不替代当前线上结论。

### 三账本计费顺序与流水核对

商汤科技 · 计费链路验证 | 2026-04 历史项目

- 围绕积分、代金券、现金三个账本核对扣费顺序与跨账本切换，结合 usage 与交易流水验证实际落账。
- 2026-04-29 nightly 报告包含 9 个有效计费样本、3 个无效或异常样本，覆盖积分优先、积分不足转券、券不足转现金等组合。
- 区分余额快照和消费流水：现金消费在 BALANCE / CONFIRMED 流水中可见，账户余额字段仍为充值快照；逐笔对账以交易流水为准。

## 更多项目实践

---

### LLM 容量矩阵与成本评估

商汤科技 · 性能与容量分析 | 2026-05 历史项目

- 围绕上下文长度 × 并发构建测试矩阵，在一致输入输出和流式口径下分析 E2E P95、TTFT P95、TPOT P95 与吞吐。
- 按 SLA 判断可用容量，区分请求成功与延迟达标；整理不同档位的厂商对比，保留测试配置、失败和高风险档位。
- 将容量数据组织为混合流量估算工具，展示输入占比、等效资源消耗和成本口径；用于方案评估与假设对比，按 SLA 和实测覆盖标注估算边界。

### 多语言 ASR 准确性 Benchmark

商汤科技 · AI 模型效果评测 | 2025-12 - 至今

- 将人工听测整理为数据制备、文本归一化、指标计算、语义诊断与人工抽检流程，复现公开评测框架并扩展中、日、英测试数据。
- 按语种设计指标：英文 casefolded CER/WER，中文 normalized CER，日文结合字符 CER 与 LLM 语义审核，识别表面差异与实际语义偏差。
- 将错误类型与 none/minor/medium/severe 四级严重度交叉归因；报告支持样本筛选、排序与定位，人工优先复核高风险样本。
- 2026-07-10 nightly 报告覆盖三语各 200 条样本，共 600 条；支持模型发布日生成报告。历史批次不外推为当前线上准确率。

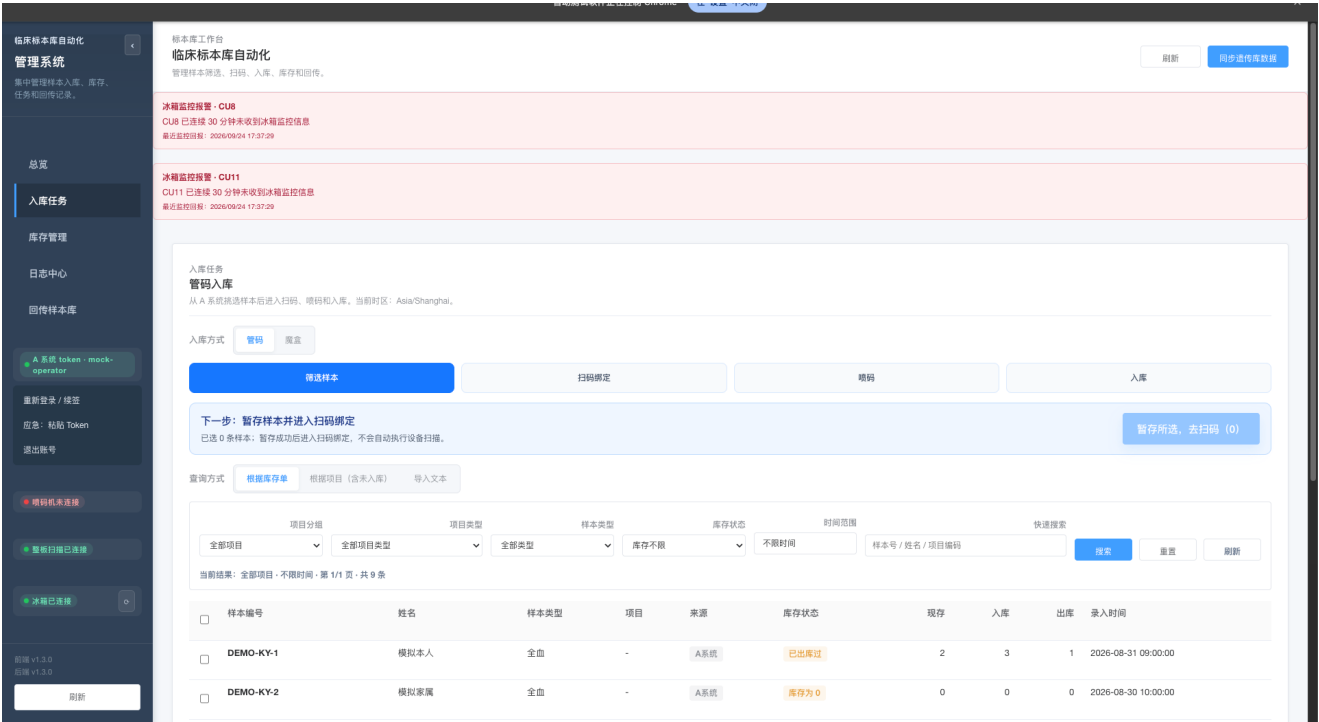
### 17 周自动化测试课程体系

码同学 · 课程设计与技术教学 | 2020-01 - 2021-03

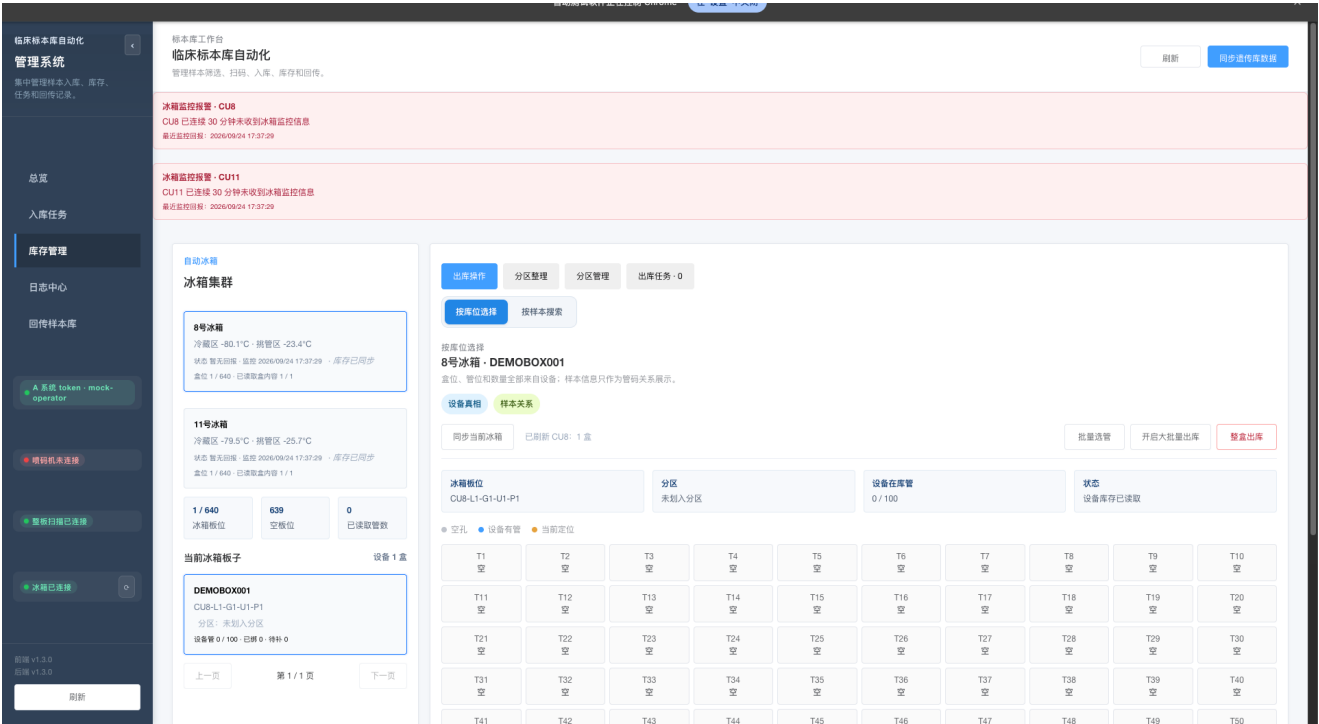
- 自主设计自动化测试课程，从 Python 基础、接口与 UI 自动化推进到持续集成，累计完成 3 次改版并覆盖 300+ 学员。
- 用业务规则设计编程练习，将条件分支、循环和用例设计联系起来；授课、答疑与录播内容并行，沉淀 Robot Framework、Airtest、Git 等专题。
- 2021-02 入职天眼查后，以兼职方式完成剩余授课周期。

# 作品集 临床标本库业务工作台

个人 AI Coding 实践。截图使用演示数据，展示业务流程与库存视图；不作为全系统设备验收证明。



入库工作台：以业务阶段组织筛选、暂存、扫码绑定、喷码与入库操作。



库存工作台：展示位置与状态。界面监控回报时间不等于回归测试快照日期。

# 作品集 企业席位建模与覆盖审计

2026-06 历史证据中枢。先明确实体、状态与权限，再按用例关联证据。

测试状态

测试范围

测试矩阵

执行证据

事实模型

执行资产

资源清单

业务动作

问题汇总

Agent 执行协议

分类口径：本区只放确定性事实、从 DB/API/页面归纳出的关系链、耦合关系和显式标注的推测规则。动态测试动作放到“业务动作模型”。

▼ 核心实体模型

| 实体              | 业务含义                                                     | 主要证据                                              |
|-----------------|----------------------------------------------------------|---------------------------------------------------|
| 手机号 Phone       | 登录入口，不等于唯一业务身份。一个手机号可对应个人账户号、企业管理员账号、企业成员账号等多个 users.id。 | users.phone: /user/login: /user/own_users         |
| 用户身份 User       | 真正参与权限、套餐、席位绑定的主体。多账号模型下必须用 users.id 判断，不可只用手机号。         | users.id, users.role, users.org_id                |
| 个人账户 Personal   | role=100，拥有个人套餐/个人额度，不应管理企业席位。                           | users.role=100, org_id 为空                         |
| 企业管理员 Org Admin | role=101，企业管理身份：新逻辑允许管理员给自己分配席位，但不能因此降级为成员。              | users.role=101, organizations.owner_user_id       |
| 企业成员 Org User   | role=102，被分配席位后产生/绑定的企业子账号，享有企业套餐权益。                     | users.role=102, org_id 指向企业                       |
| 企业 Organization | 组织容器，拥有 owner/admin、成员和企业席位池。                            | organizations.id/name/owner_user_id/status        |
| 企业批次 Batch      | 企业购买/兑换套餐产生的席位池，包含总席位、有效期、所属 owner。                      | enterprise_subscription_batches                   |
| 席位 Seat         | 批次下的最小可分配单位。分配时绑定 assignee_user_id 和 plan_apply_id。      | enterprise_subscription_seats                     |
| 套餐申请 Plan Apply | 套餐权益落地记录。席位分配成功后应为被分配身份创建有效 apply。                       | subscription_plan_applies                         |
| 权益/额度 Rights    | 页面/API 最终可用额度，必须与身份隔离。                                   | subscription_plan_right_applies / user usage APIs |

▼ 状态机与关键状态

| 对象   | 状态         | 含义                         | 判定证据                                                                   |
|------|------------|----------------------------|------------------------------------------------------------------------|
| 账号身份 | 未登录        | 手机号可登录但未选择具体 users.id。     | /user/login 返回 token: /user/own_users 返回身份列表                           |
| 账号身份 | 个人身份       | role=100：只能看个人套餐，不可操作企业席位。 | /user/self role=100                                                    |
| 账号身份 | 企业管理员身份    | role=101：可查看/分配企业席位。       | /user/self role=101: org_id 非空                                         |
| 账号身份 | 企业成员身份     | role=102：可使用企业权益，但不可分配席位。  | /user/self role=102                                                    |
| 席位   | unassigned | 席位空闲，可分配。                  | enterprise_subscription_seats.status=unassigned, assignee_user_id null |

事实模型：区分手机号与用户身份，整理企业、席位、套餐与权益关系。

| 测试状态                                                                              | 测试范围                            | 测试矩阵                                                                           | 执行证据                                                                                        | 事实模型                                                                     | 执行资产 | 资源清单 | 业务动作 | 问题汇总 | Agent 执行协议 |
|-----------------------------------------------------------------------------------|---------------------------------|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|------|------|------|------|------------|
| 测试矩阵和覆盖审计                                                                         |                                 |                                                                                |                                                                                             |                                                                          |      |      |      |      |            |
| 搜索 ID / 模块 / 测试点 / 证据    全部状态    全部模块    非 PASS    清空筛选                           |                                 |                                                                                |                                                                                             |                                                                          |      |      |      |      |            |
| 导出状态与证据    导入文本框 JSON    清空本地状态    导出的 JSON 会显示在这里；也可以把 JSON 粘贴到这里后点击“导入文本框 JSON” |                                 |                                                                                |                                                                                             |                                                                          |      |      |      |      |            |
| PASS 81 / 111    FAIL 8 / 111    WARN 18 / 111    BLOCKED 4 / 111    TODO 0 / 111 |                                 |                                                                                |                                                                                             |                                                                          |      |      |      |      |            |
| 设置状态                                                                              | ID / 模块                         | 测试点与预期                                                                         | 证据记录                                                                                        | 断言与旧证据                                                                   |      |      |      |      |            |
| PASS<br>PASS                                                                      | MODEL-001<br>PO - 模型基线<br>风险：只读 | 同手机号多身份识别<br>预期：登录同一手机号后 /user/own_users 返回个人/企业管理员/企业成员身份；每个身份有独立 users.id。   | 证据记录：写实际结果、失败原因、截图/API/DB/脚本 summary 的完整路径；执行过的 Case 必须留下可复核证据。<br><br>主证据    已记录路径    复制路径 | 查看断言                                                                     |      |      |      |      |            |
| PASS<br>PASS                                                                      | MODEL-002<br>PO - 模型基线<br>风险：只读 | 身份切换稳定性<br>预期：切换到企业管理员后 /user/self 必须 role=101；切换到成员后 role=102；切换个人后 role=100。 | 证据记录：写实际结果、失败原因、截图/API/DB/脚本 summary 的完整路径；执行过的 Case 必须留下可复核证据。<br><br>主证据    已记录路径    复制路径 | 查看断言<br>API: /user/switch + /user/self<br>旧证据: results/enterprise_seat_s |      |      |      |      |            |
| PASS<br>PASS                                                                      | POOL-001<br>PO - 席位池<br>风险：只读   | 批次与席位汇总一致<br>预期：holdings/summary/batch_by_level 与 DB seats 计数一致。               | 证据记录：写实际结果、失败原因、截图/API/DB/脚本 summary 的完整路径；执行过的 Case 必须留下可复核证据。<br><br>主证据    已记录路径    复制路径 | 查看断言                                                                     |      |      |      |      |            |
| PASS                                                                              | POOL-002<br>PO - 席位池            | 企业成员列表 API 与 DB 对齐<br>预期：/org/users 返回本企业 role=102 成员，状态与 DB 一致。               | 证据记录：写实际结果、失败原因、截图/API/DB/脚本 summary 的完整路径；执行过的 Case 必须留下可复核证据。<br><br>主证据    已记录路径    复制路径 | 查看断言                                                                     |      |      |      |      |            |

覆盖审计：111 条中 81 PASS、18 WARN、8 FAIL、4 BLOCKED；TODO 为零不代表全通过。

# 作品集 需求指标与测试证据对照

2026-07-03 历史报告示例。将需求、测试结论、证据及风险放在同一矩阵，保留 PARTIAL 与待复核项。

## 官方要求满足矩阵

根据征集文档，语音与音频类核心要求包括首包时延、单账号 QPM、中文/英文、热词/VAD/语义修正、语音翻译准确度与端到端时延。下表先给商务可用结论，再链接证据页。

| 官方关注指标                                                     | 状态      | 当前结论                                                        | 证据页                                 | 边界/风险                        |
|------------------------------------------------------------|---------|-------------------------------------------------------------|-------------------------------------|------------------------------|
| <b>ASR 首包时延</b><br>语音识别：首包时延≤2s                            | PASS    | Pro 流式首服务端事件约 103ms，首转写约 1885ms                             | <a href="#">JIRA-ASR-STREAM-001</a> | 首转写口径接近 2s，建议按评审脚本复测确认       |
| <b>ASR 单账号吞吐</b><br>语音识别：单账号≥300 QPM                       | PASS    | 同一 online API key 下，Pro 1.1s/10.56s 两档 300 QPM 均 300/300 成功 | <a href="#">JIRA-ASR-PERF-001</a>   | 当前为短窗口压测；正式提交建议延长窗口          |
| <b>中英文识别与音频时长</b><br>支持中文（含方言）/英文，覆盖短音频与 10 秒级音频           | PASS    | Pro 已覆盖中文 1.1s/10.56s/15s 在线转写；英文样本在翻译与多语种包中保留              | <a href="#">JIRA-ASR-PERF-001</a>   | 方言专项样本需按评审语料补充               |
| <b>热词/VAD/语义修正</b><br>具备热词/VAD/语义修正能力                      | PARTIAL | 已完成 20 条运营商务语链验证，19/20 命中                                   | <a href="#">JIRA-HOTWORD-001</a>    | 当前是术语识别链路，不是热词开关 A/B 增益证明    |
| <b>语音翻译准确度</b><br>语音翻译：准确度≥90%                             | PASS    | 20 条小样本链路成功，语义初审 19/20，可支撑初版 95% 口径                         | <a href="#">JIRA-TRANS-ACC-001</a>  | 正式外发前需要人工 final review       |
| <b>语音翻译端到端时延</b><br>端到端延迟≤4s（流式处理），单账号≥50 QPM              | PARTIAL | 已按正常语速补 50 QPM 专项压测，60/60 成功；平均约 6.28s，时延仍需优化/口径确认          | <a href="#">JIRA-TRANS-ACC-001</a>  | 当前不能对外承诺 50 QPM 下端到端≤4s 全量满足 |
| <b>接口/SLA/账号交付</b><br>标准公网 HTTPS、Token 鉴权、API 调用账号/密钥、流量管控 | PASS    | 本轮所有 Pro 证据均通过 online HTTPS + 单个 API key 执行，可证明单账号调用链路      | <a href="#">JIRA-CAPACITY-001</a>   | 生产 SLA 仍需商务/平台承诺文件补充         |

## 推荐评审路径

从 ASR 单账号 300 QPM 与稳定性开始，再看流式首包、翻译和热词。Lite HTTP 页面仅作为备选模型附录，不放在商务主结论里。

## 生成信息

生成时间：2026-07-03 11:22:01 CST  
主线模型：senseaudio-asr-pro-1.5-260319  
执行口径：online HTTPS 接口、Token 鉴权、同一 API 调用账号/密钥。

指标与证据对应：首包、吞吐、识别与翻译分开评价；短窗口结果不升级为长期 SLA。

报告表达先给出适用范围，再说明执行配置和证据；未达项保留限制条件，避免仅凭请求成功判断验收。

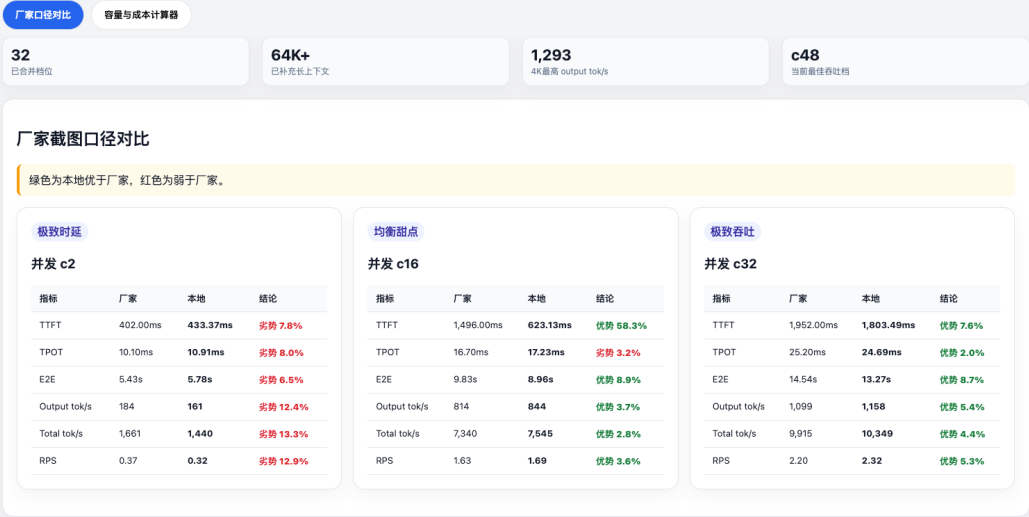
这组作品展示历史方法与产出，不代表当前线上表现、商务承诺或已完成正式采购。

# 作品集 LLM 容量分析与估算方法

2026-05-21 历史产物。按相同测试口径比较延迟与吞吐，成本工具仅作为假设下的估算参考。

## Qwen3.6-35B-A3B 厂家对比与采购决策工具

本地主矩阵 + 服务器长上下文补测 + c48/c64补充档 | stream=true | enable\_thinking=false | max\_tokens=500  
生成时间: 2026-05-21 15:07:32



作品集 ASR 三语评测与样本定位

2026-07-10 nightly 历史报告。三语各 200 条样本；调用成功和转写准确性分别统计。

Generated 2026-07-10

ASR Lite Nightly 三语精度测试综合结论

1. 测试结论

本轮在 nightly 环境完成 senseaudio-asr-lite-1.5-260319 的中/日/英 ASR 精度测试，每个语种 200 条样本，三语接口调用均为 200/200 成功。综合来看：

- 英文表现最好：按英文 casefolded 归一化后，CER 为 0.5713%，WER 为 1.6128%。
- 中文短音频表现较好：normalized CER 为 4.1360%。
- 日文 Reazon 子集字符级 CER 为 18.8758%；进一步用 senseaudio-s2-lite 做语义审核后，none/minor 语义可接受率为 76.50%，允许 medium 局部偏差时语义可用率为 79.00%。
- 本轮结果只代表 nightly 环境，不代表 staging 或 online；staging HTTP ASR 入口本轮曾返回 HTTP 500，不能作为精度结论环境。

3. 指标汇总

| 语种 | 样本数 | 成功数 | 推荐解读指标                  | CER                   | WER           | 平均延迟    |
|----|-----|-----|-------------------------|-----------------------|---------------|---------|
| 英文 | 200 | 200 | casefolded CER/WER      | 0.5713%               | 1.6128%       | 909.3ms |
| 日文 | 200 | 200 | normalized CER + 模型语义审核 | 18.8758%；语义可接受 76.50% | 不建议按空格 WER 解读 | 804.5ms |
| 中文 | 200 | 200 | normalized CER          | 4.1360%               | 不建议按空格 WER 解读 | 667.9ms |

结论与指标：按语种选择 CER/WER 等解读口径，并声明环境边界。

样本明细（全量可筛选）

筛选后 200 / 总计 200

支持通过快速标签切换到原始高误差、规范化高误差或空结果样本视图。

全部样本原始高误差样本规范化高误差样本空结果样本疑似吐Prompt样本语义不一致异常输出

关键词筛选

状态筛选

空结果

吐Prompt筛选

语义诊断

错误类型筛选

最小原始CER

样本ID / 参考文本 / 识别文本

全部

全部

全部

全部

全部

0.00

最小规范化CER

排序方式

分页大小

0.00

样本ID 正序

25

上一页

第 1 / 8 页

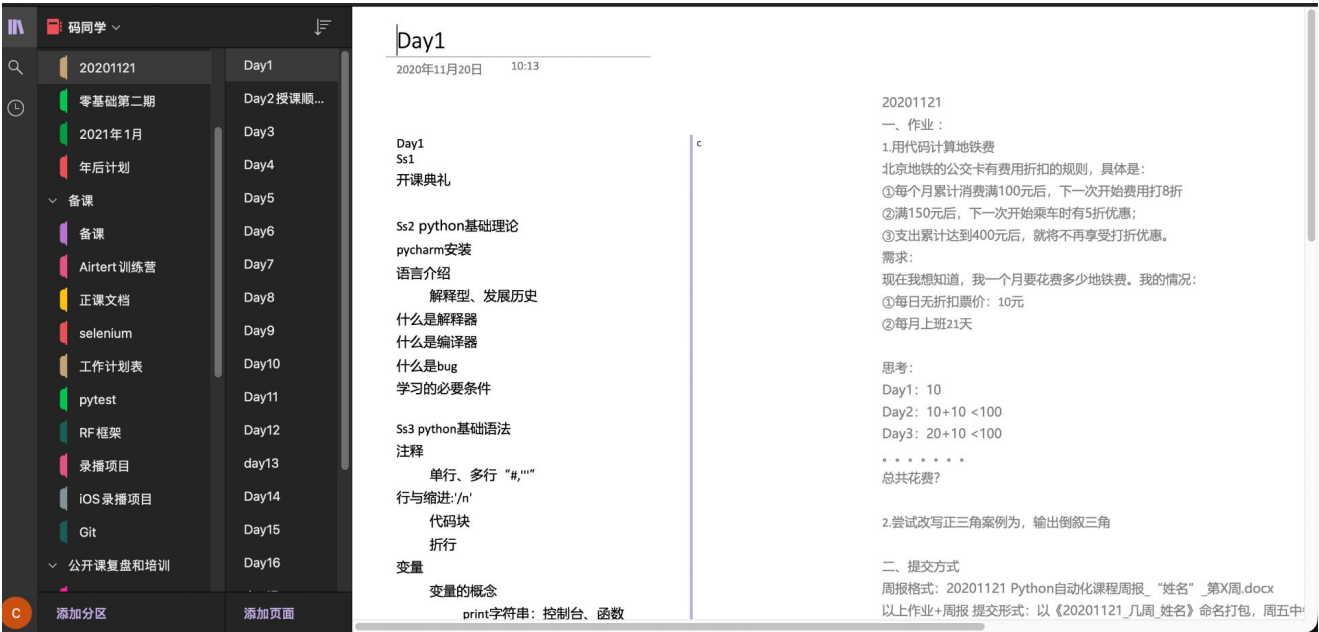
下一页

| 样本ID          | 状态      | 原始CER  | 规范化CER | 简体对齐CER | 延迟(ms)  | 疑似吐Prompt | 语义诊断 | 严重程度 | 错误类型                                              | 参考文本                     | 识别文本                    | 诊断摘要                     | 打分错误 |
|---------------|---------|--------|--------|---------|---------|-----------|------|------|---------------------------------------------------|--------------------------|-------------------------|--------------------------|------|
| ja_reazon_000 | success | 11.11% | 0.00%  | 0.00%   | 524.00  | 否         | 语义一致 | 无    | surface_for m_difference                          | これまたジミーさん                | これまたジミーさん。              | 识别结果与参考文本语义完全一致，仅标点符号差异。 |      |
| ja_reazon_001 | success | 32.56% | 30.95% | 30.95%  | 1286.00 | 否         | 部分一致 | 轻微   | homophone_error, insertion, word_order_difference | 今も相手にロンバルドのほうに肩口で握られてもすぐ | 今も相手にロンバルドの方ですね、肩越しで握ら  | 存在同音错误（人名）及口语填充词插入，核心语义保 |      |
| ja_reazon_002 | success | 23.81% | 25.00% | 25.00%  | 694.00  | 否         | 语义一致 | 无    | word_order_difference                             | 僕はタクシーのことに關してはあんまりこ      | あんまり僕はタクシーのことに關して、あんまりこ | 语序微调，语义无偏差。              |      |
| ja_reazon_0   | success | 42.11% | 42.11% | 42.11%  | 636.00  | 否         | 部分一致 | 轻微   | omission                                          | [バーミヤ                    | オンエア獲                   | 核心语                      |      |

样本明细：通过语义诊断、错误类型及 CER 筛选样本，支撑人工抽检。

# 作品集 技术教学与计费验证方法

展示技术表达和业务核对能力。计费方法表依据 2026-04-29 历史报告重述，不含账号或请求标识。



课程资料：按授课阶段组织 Python 与自动化专题，以业务规则设计练习。

| 计费场景     | 核对重点                           |
|----------|--------------------------------|
| 只有代金券    | 请求使用记录与代金券交易流水对应               |
| 积分优先     | 确认积分账优先消费，未触发的账本不扣费            |
| 积分不足转券   | 校验积分耗尽后的剩余费用进入代金券账             |
| 代金券不足转现金 | 分别对账代金券与现金流水，核对跨账本分配           |
| 券耗尽后现金扣费 | 以 BALANCE / CONFIRMED 流水核对现金消费 |
| 异常样本     | 与有效计费样本分开统计，不把样本数写成路径数         |

历史报告：9 个有效计费样本、3 个无效或异常样本。余额充值快照不替代消费流水。