王杨

AI质量工程 / AI驱动测试 / 高级测试开发 SDET

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

wangmouren.online | github.com/cowboy231

核心技能

AI自动化测试：将测试目标、业务规则和数据约束整理为 Agent 可执行任务，覆盖接口、数据库核对、文档巡检及复杂业务 UI 测试；结合人工接管、关键路径复测和证据审查，形成可追溯的测试结论。

AI测评与Agent测试：中、日、英 ASR Benchmark，分语种 CER/WER、LLM 语义诊断与人工抽检。

自动化与协议测试：Python、pytest、Playwright、Robot Framework；HTTP、WebSocket、Mock、契约与集成测试；具备 Selenium、Appium 使用经验。

性能与数据质量：Locust、JMeter；吞吐、首包首块、实时率、P95、容量与 GPU 负载分析；MySQL、Redis、SQLite，数据对账、自动巡检与异常定位。

AI Coding与工程化：临床标本库使用 Codex + MiniMax 辅助全栈开发，覆盖 Vue 3、FastAPI、SQLite、RabbitMQ 与 Windows 内网交付；具备 Jenkins、GitLab CI、Docker 工程经验。

技术教学与团队协作：设计并迭代 17 周自动化测试课程，完成 3 次改版，累计覆盖 300+ 学员；具备录播交付、知识沉淀、新人培养及测试团队协作经验。

工作经历

全国采伐审批系统测试 · AI 自动化测试
AI 测试工程师（顾问合作） | 2026-10 - 至今

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。

项目经历

全国采伐审批系统 · AI 自动化测试

顾问合作 · 复杂业务流程与人机协作测试 | 2026-10 - 至今

- 业务探索：接触系统三四天后，与 AI 结对梳理审批流程、字段依赖和业务规则，在测试执行中持续核对、建立业务理解。
- AI 自动化执行：将测试目标、前置条件和预期结果拆成 Agent 可执行任务，由 AI 准备数据、填写表单、复核字段和重开业务记录；本人处理关键审批、异常接管和验收。
- 业务规则验证：排除身份校验等前置干扰，对目标阈值执行测试，并通过独立重开同一记录，补充问题值持久化的缺陷证据。
- 跨环节结果核对：人工完成审批，AI 查询并计算前后额度差异，验证该对账子项；将业务规则、样本、操作记录和证据沉淀为后续可复用的测试资产。

临床标本库自动化系统 AI Coding 全栈实践

个人项目 · 需求、开发、集成与分层验证 | 2026-07 - 至今

- 从现场操作、既有业务系统与设备约束中梳理实体、状态、不变量和异常路径，明确每条业务链路的验收条件。
- 使用 Codex + MiniMax 辅助实现 Vue 3 前端、FastAPI 服务、SQLite 数据模型与 RabbitMQ 适配；连接上游业务系统及喷码、整板扫描、自动化冰箱，形成可运行的内网系统。
- 将上游业务记录、本地绑定与过程记录、设备物理结果分层保存；长任务依据设备终态与实际结果判断完成，并设计异常恢复与库存回读。
- 本人把控业务判断、证据可信度、变更粒度与真实写操作；AI 辅助实现、回归和排障。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；现场链路按设备和版本另行核验。

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。

AI 测试资产与 Agent 执行体系

商汤科技 · AI 驱动测试 | 2025-12 - 至今

- 在开发使用 AI 编码的快速迭代中，将测试目标、API/数据库约束、预期结果和证据要求整理为可执行任务。
- 接口测试和数据库核对以 Agent 执行为主；使用 AI 完成首轮冒烟，本人复测异常与关键路径，审查请求、响应、数据及日志后形成业务结论。
- 将复核过的方法、用例、边界及执行记录沉淀到测试资产库，分别记录当前规则、历史观察与待裁决提案，支持后续复跑和追溯。

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。

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。

项目经历 续

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。

九坤私募基金测试体系

九坤私募 · 复杂业务建模与数据验证 | 2023-08 - 2024-06

- 在真实生产数据不可用的保密约束下，梳理上下游业务规则，建立场景化 Mock 数据和 SQL 数据集。
- 按实体、字段和规则关系拆解基金业务，重构 Robot Framework 参数化与 BDD 脚本，推动上下游统一字段定义和取值口径。
- 形成可执行的基金全生命周期验证方案，将测试范围从页面功能延伸至业务规则与数据一致性。

积家科技 外贸自研摄像头项目

高级测试工程师 · 海外智能硬件 | 2024-07 - 2025-10

- 参与面向海外市场的自研电池摄像头项目测试；公司同时提供面向合作方的 SaaS 系统。
- 项目主要测试约束包括高频迭代、文档不完整、系统耦合与基础设施薄弱，复杂场景构造和问题定位成本较高。

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。