

高密度技术履历证据审计

林知序（公开演示）

范围：教育 Title、Owner/0→1、Agent Infra、开源角色、产品指标、奖项与时间线 / 生成时间：2026-08-13T13:45:00+08:00 / 分析者：asu-resume-audit-skill

证据结论总览



这份高密度履历的教育标签、开源 Contributor/Maintainer 身份和多数技术产物存在明确锚点；主要风险集中在内部 Owner 任命缺少独立确认、团队训练收益向个人模块归因、直播搭建“0→1”与既有原型时间线冲突，以及产品用户指标只能由内部快照支持。

本报告只说明演示材料中的证据关系，不评价人格，也不构成法律认定。

已知限制

- 无法访问真实公司的代码所有权、绩效系统、实验平台和人事记录。
- 演示中的内部看板与设计文档只能证明材料之间一致，不能替代雇主独立确认。
- 示例域名和候选人均用于展示审计结构，不对应现实求职者。

审计流程

解析简历 文本 / 图片 / PDF	拆分主张 角色 / 范围 / 指标	检索证据 官方 / GitHub / 档案	交叉核验 支持 / 矛盾 / 缺口	生成报告 JSON / HTML / PDF
-----------------------	----------------------	--------------------------	----------------------	---------------------------

最高风险主张

编号	简历主张	结论	风险	核心分析
C-002	AgentScaling 轨迹合成模块 Owner，从 0→1 设计六阶段数据生产管线。	部分证实	高	责任矩阵和设计评审支持主要设计与接口责任，但证据来自候选人提供的内部材料，缺少主管确认或正式 Owner 名单。作用域必须保持在“轨迹合成模块”。
C-004	训练后模型任务通过率从 23.8% 提升至 31.6%。	部分证实	高	评测报告支持前后数值，数据看板支持训练数据来源；仍有模型版本、训练超参和其他数据变化等共同变量，不能把全部提升单独归因于候选人的模块。
C-008	直播间智能搭建 Workflow 策略模块 Owner，从 0→1 设计端到端链路。	存在直接矛盾	高	候选人确实负责新的 LLM Workflow 策略模块，但加入前已有规则原型。若“0→1”指新策略模块可成立，若指整个直播搭建产品则与提交时间线直接冲突。
C-011	Thumbnail AI 注册用户 2,080，付费用户 214，访问到付费转化率 10.3%。	部分证实	高	分子分母算术一致，数据库与事件日志支持数量；缺少第三方支付导出、退款和测试账号排除规则，因此只能部分证实。

完整主张与证据矩阵

C-001	学校具有列出的三项官方机构或建设标签。
状态	已证实
风险 / 置信度	低 / 高
证据	支持：S-001、S-002；反向：无
分析	学校官网、主管部门名单与学历材料共同支持学校身份和 Title；履历没有把合作关系写成第二学位。
下一步	无

C-002	候选人被正式或事实上授权为轨迹合成模块 Owner，并主要负责该模块从无到有建立。
状态	部分证实
风险 / 置信度	高 / 中
证据	支持：S-003；反向：无
分析	责任矩阵和设计评审支持主要设计与接口责任，但证据来自候选人提供的内部材料，缺少主管确认或正式 Owner 名单。作用域必须保持在“轨迹合成模块”。
下一步	- 补充模块责任人列表、设计评审参会人与代码所有权记录。 - 避免省略“轨迹合成模块”而外推为完整训练平台 Owner。

C-003	数据生产系统累计产出 222k 条符合履历定义的轨迹。
状态	部分证实
风险 / 置信度	中 / 中
证据	支持：S-004；反向：无
分析	看板支持数量，但需确认去重、失败样本、重采样和有效轨迹的统计口径；该规模是系统产出，不应暗示候选人手工完成全部轨迹。
下一步	- 提供 count SQL、去重键和有效轨迹定义。

C-004	候选人负责的轨迹合成工作导致固定 1,200 题上的任务通过率提升 7.8 个百分点。
状态	部分证实
风险 / 置信度	高 / 中
证据	支持：S-004、S-005；反向：无
分析	评测报告支持前后数值，数据看板支持训练数据来源；仍有模型版本、训练超参和其他数据变化等共同变量，不能把全部提升单独归因于候选人的模块。
下一步	- 补充仅替换轨迹数据的控制实验与训练配置 diff。 - 正文使用“训练后模型”而非“我将模型提升”。

C-005	在 9,600 条轨迹的指定抽检方案下，评测系统达到两个所述指标。
状态	已证实
风险 / 置信度	中 / 中
证据	支持：S-005；反向：无
分析	报告包含样本规模、Rubric 和抽检结果；公开版本仍应明确高风险失败定义和人工标注基准。
下一步	- 保留抽样脚本、混淆矩阵和标注一致性计算方式。

C-006	候选人参与 Context Engine 的 Compaction、Resume 和状态一致性子模块，但不是完整平台 Owner。
状态	已证实
风险 / 置信度	低 / 中
证据	支持：S-006；反向：无
分析	设计文档与分工记录支持共建者角色，履历主动限制了范围并否定整个平台 Owner，角色强度与证据匹配。
下一步	无

C-007	在 240 个指定长程任务上，系统达到所述上下文窗口和恢复成功率。
状态	部分证实
风险 / 置信度	中 / 中
证据	支持：S-006；反向：无
分析	内部回归表支持结果，但“可持续运行窗口”不是模型原生 context window，必须保持为 Agent 系统通过压缩和续跑得到的工程指标。
下一步	- 在公开文案中保留“通过 Compaction/Resume”限定。

C-008	候选人加入前不存在直播页面自动搭建能力，候选人从绝对零开始建立该产品。
状态	存在直接矛盾
风险 / 置信度	高 / 高
证据	支持：S-007；反向：S-008
分析	候选人确实负责新的 LLM Workflow 策略模块，但加入前已有规则原型。若“0→1”指新策略模块可成立，若指整个直播搭建产品则与提交时间线直接冲突。
下一步	- 改写为“从规则原型到 LLM Workflow 的架构重构”。 - 保留“Workflow 策略模块”而不使用产品级 0→1。

C-009	指定上线窗口内，Agent 生成覆盖 46% 目标页面并显著缩短首版搭建耗时。
状态	部分证实
风险 / 置信度	中 / 中
证据	支持：S-007；反向：无
分析	看板支持覆盖率和时长区间，但需确认覆盖分母、人工返工时间和页面复杂度分层；结果属于产品与团队，不是个人单独产出。
下一步	- 补充覆盖分母与复杂度分桶。 - 把个人贡献限定为 Workflow 策略模块。

C-010	EC-Monica 在指定月份达到所述调用量与报告时长变化。
状态	已证实
风险 / 置信度	中 / 中
证据	支持：S-009；反向：无
分析	服务报表支持数字，履历没有将调用次数等同于独立用户或收入；仍需注明“调用”而不是“用户”。
下一步	无

C-011	截至 2024-12，产品存在 2,080 个去重注册用户和 214 个真实付费用户，转化率为 10.3%。
状态	部分证实
风险 / 置信度	高 / 中
证据	支持：S-013；反向：无
分析	分子分母算术一致，数据库与事件日志支持数量；缺少第三方支付导出、退款和测试账号排除规则，因此只能部分证实。
下一步	- 提供支付平台订单导出并扣除退款、测试与赠送账号。 - 明确定义“访问到付费”中的访问分母为何使用注册用户。

C-012	候选人在该仓库以 Contributor 身份合入 6 个 PR，仓库当前约 7.6k Stars。
状态	已证实
风险 / 置信度	低 / 高
证据	支持：S-010；反向：无
分析	PR 与 Stars 均可由仓库记录核验；履历明确使用 Contributor，没有把仓库热度转化为核心作者身份。
下一步	无

C-013	候选人具有正式 Maintainer/CODEOWNER 身份，并参与四个版本发布。
状态	已证实
风险 / 置信度	低 / 高
证据	支持：S-011；反向：无
分析	MAINTAINERS、CODEOWNERS 和 Release 记录提供正式且可重复核验的角色证据。
下一步	无

C-014	候选人的三个指定 PR 已合入且包含所述代码范围。
状态	已证实
风险 / 置信度	低 / 高
证据	支持：S-012；反向：无
分析	合并状态、文件 diff 与 benchmark 提交均支持，角色仍限定为 Contributor。
下一步	无

C-015	产品快照中的用户、付费与生成量均代表去重真实用户和成功任务。
状态	暂不可核验
风险 / 置信度	中 / 低
证据	支持：S-014；反向：无
分析	当前只有候选人提供的单份快照，无法判断测试账号、重复账号、退款和失败任务口径，需要原始表或第三方支付数据才能核验。
下一步	- 补充用户去重 SQL、成功任务状态定义与支付平台原始导出。

C-016	候选人获得列出的三项奖项或荣誉。
状态	已证实
风险 / 置信度	低 / 高
证据	支持：S-015；反向：无
分析	官方名单支持奖项名称与时间；履历没有增加“最年轻”“唯一”或排名第一等无法穷举的最高级。
下一步	无

识别到的包装与夸大模式

模块 Owner 向平台 Owner 外扩风险

多个模块级 Owner 连续出现时，读者容易把局部闭环理解为整个平台领导职责。

观察信号：轨迹合成模块 Owner；评测诊断模块 Owner；Workflow 策略模块 Owner

反向证据/合理解释：履历多数标题保留了模块限定，Context Engine 还主动写为共建者，因此不是无边界的全局 Owner 叙事。

置信度：高

0→1 起点重定义

把新算法模块或架构重写描述为整个产品从零开始。

观察信号：从 0→1；加入前已有规则原型；新 Workflow 与旧产品并存

反向证据/合理解释：新的 LLM Workflow 策略模块确实可被视为模块级 0→1。

置信度：高

团队训练收益向个人模块归因

模型任务通过率同时受训练超参、基座版本、数据组合和推理配置影响，不能只归因于轨迹合成负责人。

观察信号：前后通过率；缺少单变量控制实验；个人职责与团队训练结果并列

反向证据/合理解释：数据生产看板和训练报告建立了候选人模块与训练数据之间的直接链路。

置信度：中

项目热度品牌转移风险

仓库 Stars 可以说明项目影响力，但不能证明个人是项目核心作者。

观察信号：7.6k Stars；个人贡献与项目热度同一行

反向证据/合理解释：履历明确写 Contributor，并单独列出 6 个合并 PR，没有使用核心作者措辞。

置信度：中

产品指标伪精确风险

内部数据库快照提供精确数字，但退款、测试账号、重复用户和时间窗口可能改变含义。

观察信号：2,080 注册用户；214 付费用户；10.3% 转化率；缺少第三方支付导出

反向证据/合理解释：Thumbnail AI 的分子分母算术一致，并存在数据库、事件和支付报表三类内部材料。

置信度：中

系统指标与模型原生能力混淆

通过 Compaction 和 Resume 获得的可持续运行 token 不等于基座模型原生上下文窗口。

观察信号：430k token；Context Engine；Checkpoint Resume

反向证据/合理解释：履历将该指标放在 Context Continuation 项目下，没有直接称为模型 context length。

置信度：中

时间线

日期	事件	详情	来源
2024.06	Thumbnail AI 开始开发	个人项目提交记录显示前端、生成服务和支付链路开始建立。	S-013
2024.08	软件创新赛全国一等奖	官方获奖名单记录候选人与参赛项目。	S-015
2024.12	直播搭建规则原型存在	候选人加入前，旧仓库已存在规则配置和页面生成原型。	S-008
2025.01	加入直播与电商 Agent 团队	开始负责新的 LLM Workflow 策略模块和 EC-Monica 分析编排。	S-007、S-009
2025.05	直播 Workflow 上线	看板开始记录页面覆盖率、搭建时间与组件注册数量。	S-007
2025.07	加入 AI4SE Agent Scaling	责任矩阵记录轨迹合成、评测诊断与 Context Engine 三类工作。	S-003、S-006
2025.11	eval-trace-core 0.4 发布	Maintainer 与 CODEOWNER 身份进入公开仓库记录。	S-011
2026.03	AgentScaling 轨迹生产扩容	数据看板开始按 Scaffold、失败类型和有效性门禁分层统计。	S-004
2026.05	Rubric Eval 大规模抽检	在 9,600 条轨迹上完成模型评审与人工复核。	S-005
2026.06	SFT 固定回归评测	1,200 题回归集记录任务通过率、工具调用和成本变化。	S-004、S-005

下一步核验

- 1. 将所有 Owner 表述持续限定在轨迹合成、评测诊断或 Workflow 策略模块。
- 2. 把直播搭建“0→1”改写为从规则原型到 LLM Workflow 的架构重构。
- 3. 对训练收益补充模型版本、训练超参、数据配方与单变量控制实验。

4. 为 222k 轨迹规模提供去重键、有效性定义和统计 SQL。
5. 为产品用户与付费指标补第三方支付导出、退款和测试账号排除规则。
6. 保留 Contributor/Maintainer 的正式角色证据，不用项目 Stars 替代个人贡献。
7. 明确 430k token 是 Agent 系统续跑窗口，不是基座模型原生上下文。
8. 对每个公开指标附样本量、分母、时间窗口和计算脚本。

来源与证据等级

S-001 / 南方理工大学学校概况与教育主管部门名单

official_record / 学校官网与教育主管部门 / 2026-08-13 / 等级 A

支持省部共建、卓越工程师教育培养计划与新工科项目等机构标签。

S-002 / 匿名学历与成绩材料

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持学校、专业、学位授予方、修读课程和就读时间。

S-003 / AgentScaling 模块责任矩阵与设计评审

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持轨迹合成模块接口、交付责任和主要设计者，无法独立证明正式 Owner 任命。

S-004 / AgentScaling 数据生产与成本看板

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

记录 Micro/Long-horizon 轨迹数量、失败类别和标准化成本。

S-005 / SFT 与 Rubric Eval 固定回归报告

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

含 1,200 题训练评测和 9,600 条轨迹 Rubric 抽检结果，未提供可公开复跑环境。

S-006 / Context Engine 设计文档与分工记录

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持 Context Engine 共建、Compaction 和 Resume 子模块职责，不支持整个平台 Owner。

S-007 / 直播搭建平台责任分工与上线看板

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持 Workflow 策略模块责任、组件数量、页面覆盖率和搭建时间。

S-008 / 直播搭建旧版规则原型提交记录

official_record / 内部代码仓库镜像 / 2026-08-13 / 等级 A

显示候选人加入前已存在基于规则的页面搭建原型；可支持架构重写，但反驳产品从绝对零开始。

S-009 / EC-Monica 服务调用与任务分类报表

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持月调用量、任务分类和报告生成时长，不支持业务收入归因。

S-010 / deer-harness PR、Stars 与贡献者记录

repository_record / GitHub 项目仓库 / 2026-08-13 / 等级 A

可核验 9 个 PR、6 个合入和仓库 7.6k Stars；未显示核心作者或 Maintainer 身份。

S-011 / eval-trace-core MAINTAINERS、CODEOWNERS 与 Release

repository_record / GitHub 项目仓库 / 2026-08-13 / 等级 A

正式文件记录 Maintainer 与代码所有权，并显示四个版本的发布活动。

S-012 / fast-transport 合并 PR 与 Benchmark 提交

repository_record / GitHub 项目仓库 / 2026-08-13 / 等级 A

支持三个合并 PR、Python Binding 修复和跨语言 benchmark。

S-013 / Thumbnail AI 数据库快照、支付报表与事件日志

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持注册、付费、月活、生成量与首版耗时；缺少第三方支付平台原始导出。

S-014 / Krene-Art 产品快照与生成日志

user_provided_document / 匿名候选人 / 2026-08-13 / 等级 B

支持用户、付费、生成量和平均耗时，无法独立确认测试账号清理口径。

S-015 / 赛事获奖名单与开源结项公示

official_record / 赛事组委会与开源活动官网 / 2026-08-13 / 等级 A

支持全国一等奖和优秀结项；校级毕业设计由学校公示单独确认。