

UFH CareLoop

睦诊续航

把患者在两次就诊之间发生的变化，整理成医生下次见面前可以核对、可以处理、也可以追溯的信息。

医生 AI Copilot 报名方案

个人参赛 · 医学生视角 · 报名设计版 v1.0

本材料使用公开资料、合成病例和模拟医院接口。文中的指标均为赛程内验收目标，不是已经取得的临床结果。

1. 我对题目的判断

我把 Copilot 的工作边界放在两次就诊之间。患者离开诊室后，症状变化、用药执行、检查待办和新的疑问会继续发生；医生再次见到患者时，需要知道这段时间里发生了什么，哪些问题已经处理，哪些问题还没有人接手。

现有工具各有价值，但通常只覆盖其中一段。

常见做法	能解决什么	仍然留下什么
电话随访	沟通直接，有温度	记录难统一，漏拨和交接依赖人工
患者门户或表单	方便触达和收集	收到回答不等于异常已经有人处理
通用 AI 问答	能解释和生成文字	容易脱离本次医嘱，也缺少责任人和处置时限

和睦家官网将其医疗模式概括为从预防延伸到诊断、治疗和康复的连续照护；护理介绍也强调循证、双语和跨文化沟通。AHRQ 对患者、家属和医护人员的焦点组研究则把沟通协同、出院流程复杂以及人力和时间不足列为从医院到家庭过渡中的常见障碍。这些公开材料支持我选择“连续照护”这个方向，但不能代替和睦家内部访谈，也不能据此断言医院现有流程存在某个具体缺陷。

CareLoop 要让每一条院外变化都能回答六个问题：

- 1

从哪里来
- 2

目前是什么状态
- 3

是否触发已知风险规则
- 4

由谁在什么时间前处理
- 5

医护最后怎样判断
- 6

怎样带回下一次就诊

我会先用成人日间手术合成病例做演示。它的时间节点紧、院外变化多，适合在几分钟内看清系统是否真的工作。72 小时随访只是这条旅程中的一个窗口；产品仍覆盖本次诊前、诊后、复诊和下一轮计划，并保留慢病、检查待办和康复路径包。

公开依据：和睦家连续照护模式、护理团队与日间手术材料；AHRQ Transitional Care Medication Safety。

2. 患者走过的一整轮

一条 `care_journey_id` 可以包含多次就诊。每次就诊有独立的 `episode_id`，结束后生成下一轮计划；一次术后随访不会被当成旅程终点。

D-7 本次诊前

汇总既往变化、待办和患者最想解决的问题

D0 诊中与离院

医生确认诊疗、离院和随访路径

D1 / D3 主动随访

收集结构化回答与中英文自由文本

异常时 升级与处置

规则先行，任务到人，超时再次升级

D7 复诊前后

生成有出处简报，确认结论并模拟写回

下一轮 继续照护

创建新的 `episode`，承接未完成问题

演示怎么跑

4 至 5 分钟的连续录屏只使用一位合成患者：

- 1. 系统把既往就诊、本次目标和未完成事项放进诊前准备卡。
- 2. 模拟医生确认本次离院与随访路径，系统按时间压缩方式创建 D1、D3 和 D7 任务。
- 3. D1 收到中英混合文本。多维表格 AI 提取症状、时间线、依从性和患者问题，同时保留可核对的原文片段。
- 4. D3 勾选一项结构化红旗，并在自由文本里加入否定句和提示注入。即使 AI 给出低风险候选，规则仍立即建立高优先级复核任务。
- 5. 模拟医护记录处置，修改并签发沟通草稿。比赛中的医生和护士节点由我切换合成角色，只验证权限、审核闸门和状态流转。
- 6. aily 在获准字段和限定知识范围内生成 D7 复诊前简报，列出来源、时间线、已处理事项和未解决问题。
- 7. 模拟医生确认复诊结论，系统写入 `simulated_success`，创建 `next_episode_id`。生产环境只有真实接口成功且有资质医护确认后，才允许进入生产结案状态。

同一条旅程还会快速回放三种输入：

场景	改变的输入	系统必须做到
常规恢复	未命中预设红旗	仍保留复核、沟通、简报和回写
红旗与超时	结构化红旗命中，任务到期未处理	立即升级，再次升级，留下处置记录
未响应、身份失败或注入	没有回答、身份校验失败，或文本要求绕过规则	不泄露、不把缺失当作正常，停止或转人工

3. 飞书 AI 在系统里具体做什么

我会在录屏中同时展示配置、输入、AI 原始输出、人工修订和运行日志。普通表格、自动化、任务和消息只负责承载流程；真正计入飞书 AI 使用证据的是多维表格 AI 与 aily 的实际运行。

环节	飞书能力	具体工作	输出边界
读懂患者回答	多维表格 AI	从中英文自由文本提取症状、时间、依从性、问题和原文片段	只生成候选字段，不诊断
整理沟通	多维表格 AI 生成文本	根据已审核字段起草双语沟通	未经医护确认不得外发
推进下一步	aily	读取已审核字段和规则结果，调用白名单工具创建或更新任务	不生成患者 ID，不关闭高风险事件
查找依据	aily 中显式配置的限定知识源	查询经批准的路径、材料和版本，生成带来源回答	无来源、冲突、过期或越权时拒答
准备复诊	aily + 多维表格 AI	形成纵向简报草稿，列出变化、处置和未解决问题	不补写无来源的诊断和用药
留下改进数据	多维表格	保存 AI 原稿、人工修改、最终版本和原因	每次修改都绑定运行版本



这四层的顺序不能互换。规则先处理已知红线，AI 可以提出更高风险的候选，但不能覆盖或降低规则结果；到了生产环境，诊断、用药、患者去向、外发内容和结案仍由有资质医护决定。

飞书官方资料显示，多维表格 AI 可用于翻译、分类、总结、信息提取和内容生成；aily 可以处理复杂任务、调用工具、基于知识回答、创建飞书任务并将信息写入多维表格。CareLoop 采用这些公开能力进行设计。赛事租户中的具体功能、额度和管理员权限仍要在第一天逐项实跑，官网介绍不算运行证据。

计划保存六张证据卡：中文提取、中英混合、困难错误及人工修订、规则不被 AI 降级、限定知识带来源回答、越权或无依据时拒答。每张卡都带 `run_id`、数据版本、配置版本和时间戳。

4. 系统怎样留下证据

原型使用七张逻辑表。它们分工明确，但通过同一组标识关联。

逻辑表	记录内容
Care_Episodes	本轮旅程、路径版本、同意、语言、复诊结论和下一计划
Followup_Events	触发、送达、回答原文、AI 候选和人工确认
Rule_Evaluations	规则结果、理由、版本和输入哈希
Review_Tasks	责任人、队列优先级、时限、处置和关闭原因
Communication_Drafts	模板、AI 原稿、人工编辑、签发和发送状态
Previsit_Briefs	时间线、来源、未解决问题、确认和写回
Audit_Events	权限、工具调用、状态变化、失败、重试和纠错

episode_id 是本轮就诊的唯一主键；care_journey_id 是可重复的纵向分组键。每条记录还带 scenario_id / run_id / source_id / version / artifact_hash。评委可以沿着一个测试场景，逐项核对原文、规则、任务、简报、指标和截图。

系统分别管理三种状态：

对象	状态示例	这样拆开的原因
Episode	本次诊前 → 路径已确认 → 随访中 → 复诊前 → 结论 → 下一计划 → 关闭	一次就诊包含多个随访事件
Event	已安排 → 等待回答 → 已收到 / 未响应 → 已复核 / 已升级 → 已处置 → 关闭	D1 和 D3 可以各自推进
Task	待分配 → 处理中 → 超时 → 再次升级 → 已完成 / 已取消	责任人和时限不与患者回答混在一起

一份版本清单驱动后续工作：

schema_manifest	path_manifest	case_manifest
冻结数据	七表装载与行数核对	五组运行
人工修订采集	聚类统计	证据卡与产物清单

case_manifest 只产生输入、空关联键和预期规则契约。AI 原始输出、实际任务、人工修订和审计只能由运行过程写入，避免把参考答案提前混入待测结果。

5. 安全先于自动化

这是医疗场景中的辅助系统，不提供面向患者的自动诊断。原型首先验证下面六条约束能否被执行：

- 1. 结构化红旗和已批准规则先运行；AI 不能降低它们给出的风险状态。
- 2. 身份未通过时，不读取或展示病情，也不调用敏感工具。
- 3. 患者撤回同意后，停止主动触达并记录原因。
- 4. 未响应是独立状态，不能被系统默认为“恢复正常”。
- 5. 知识没有来源、版本过期或来源互相冲突时，系统拒答或转人工。
- 6. 生产中的诊断、用药、患者去向、对外沟通和结案由有资质医护确认；比赛中的相应角色均标为 **SIMULATED**。

可能的失效	预防办法	演示中的检查
红旗漏掉	结构化问题与确定性规则	冻结集中的 100 个预设红旗场景
AI 把风险说低	风险单调约束	故意给出低风险候选，任务仍升级
错患者或越权	ID 不由模型生成，读写前校验	身份失败时不显示、不调用
草稿直接发给患者	草稿区和签发闸门	未签发状态不能发送
没回复却被关闭	<code>unreachable</code> 状态与人工队列	超时、重试和恢复对账都有日志
简报出现无依据内容	原文与知识来源锚点	P-final 中无依据临床陈述目标为 0

比赛只使用合成数据和假 ID。真实身份映射、完整病历、正式医嘱和真实 EHR 接口留在医院系统；飞书仅处理经医院批准的最小字段。是否符合个人信息保护、医院制度和适用伦理要求，必须由医院临床、法务、信息安全和隐私团队共同评估，不能由一份参赛方案自行宣布。

我作为医学生可以做的是：把症状时间线、否定表达、用药数字、排除人群和安全提示设计得更接近临床沟通，并用 OSCE、FMEA 和循证检索的方式检查样本。临床阈值、处方、诊断和患者去向不属于我的执业权限。

6. 我会怎样验证它

测试单位统一为 200 个 `scenario_bundle`。它们不对应 200 位真实患者；每个场景包含最小的 episode 背景、一个受控扰动、预期规则或任务，以及预期的简报增量。

类别	数量	组成
结构化红旗	100	10 个母模板 × 10 个受控变体
常规恢复	40	4 个母模板 × 10 个受控变体
困难与系统故障	60	6 个母模板 × 10 个受控变体

变体覆盖语言、否定与并列、身份与同意、送达与未响应、提示注入、越权、规则版本和系统故障。标签和预期状态在运行前冻结；改标会生成新的 `dataset_version`，不会覆盖旧版。另设 30 题限定知识检索集，其中 15 题有答案、5 题无答案、4 题来源冲突、3 题知识过期、3 题越权请求，用来检查来源准确、范围拒答和版本冲突。

五组对照

组	要回答的问题
B0 纯人工	原始处理时间和遗漏情况怎样
B1 表单 + 规则	不使用生成式 AI 已经能解决多少
B2 通用 AI 摘要	缺少来源、权限和流程约束时会犯什么错
P-draft	CareLoop 的 AI 原稿质量怎样
P-final	加入人工确认后的总成本与可用性怎样

B1、B2 和 P-draft 共 600 次由流水线批量运行；B0 和 P-final 共 400 次由我完成计时复核，系统自动保存起止、修改和异常。五组使用同一数据版本和任务定义。统计按母模板聚类，避免把同一模板的 10 个变体误当成 10 个完全独立病例。

验收目标	适用对象
预设结构化红旗升级漏检 0	B1 与 CareLoop workflow
正常运行时红旗任务恰好创建一次，100%	B1 与 CareLoop workflow
故障被发现并转人工，恢复后补建或对账无遗失，各 100%	CareLoop workflow
关键字段 Macro-F1 $\geq$ 0.90	P-draft，本人冻结的参考标注
无依据临床陈述 0	P-final；B2 只作负对照
每场景人工分钟比 B0 下降 $\geq$ 40%	P-final，探索性单评测者结果

这些数字是合成测试的准入目标。它们不代表临床有效性，也不外推患者体验、再入院、不良事件或真实医护工作量。企业价值先用可核对的公式估算：

年度可回收工时

= 年度纳入 episode 数 × (B0 人工分钟 - P-final 人工分钟) ÷ 60

真实 episode 数、小时成本和患者结局由和睦家数据与获批影子试点填写。

7. 一个人的四周工作安排

我会用约 25 个专注人日完成这套原型，包含周末工作；批量运行安排在夜间，所有任务由版本清单和运行日志对账。范围不因一人参赛而删减，但每一项都要区分设计、实跑和临床验证。

工作	人日
租户能力检查、依赖和版本基线	1
七表、关联字段和三类状态机	3
多维表格 AI、规则、任务与沟通草稿	4
aily、限定知识与工具白名单	3
主病例、完整旅程和三种安全场景	3
200 个场景包、30 题知识集、冻结与复标	4
五组评测、失败分析和重跑	4
录屏、报告、读者测试与提交	3

成人日间手术会完成全链路演示和全量测试。慢病长期管理、检查结果与复诊待办、康复三类路径各交付一份 `path_manifest`、字段映射、状态约束和 `schema dry-run`；在医院确认前不填写临床阈值，也不声称经过临床验证。

报名时的真实状态

- 已完成：问题分析、产品设计、数据结构、安全边界、测试设计和单人执行方案。
- 赛程开始后实跑：赛事租户能力检查、七表原型、200 个场景包、五组评测、六张飞书 AI 证据卡和连续 Demo。
- 需要和睦家确认：首条真实路径与排除人群、红旗与处置时限、患者入口、最小数据集、责任链、EHR 写回方式和影子试点指标。

最终交付要能被逐项检查：患者回答留下原文，规则及时守住已知风险，任务有责任人，医护判断被记录，复诊前信息有出处，结论还能进入下一轮。

公开资料

1. 和睦家医疗：Continuity of Care Model

2. 和睦家医疗：Nursing Team

3. 和睦家日间手术公开材料

4. AHRQ: Transitional Care Medication Safety
5. 飞书：多维表格 AI 功能介绍

6. 飞书：aily 智能体定制能力

7. 飞书：使用知识问答

8. WHO: Ethics and governance of artificial intelligence for health