# 智能体系统项目任务计划

## 一、整体路线图：从实验模型 → 智能体原型

本项目目标是将 Qwen-0.5B 模型训练为一个具备流程编排、意图识别、报告生成与患者沟通能力的医疗智能体。

|  |  |  |  |
| --- | --- | --- | --- |
| 阶段 | 目标 | 关键产物 | 周期 |
| 阶段1：协议冻结与数据体系统一 | 明确四个子任务（Planner、Intent、Report、Comms）的 Schema、奖励指标和输入输出规范。 | Schema文档、final\_state字段清单、奖励项说明书 | 1周 |
| 阶段2：最小可运行环境（MVP环境） | 实现最小版 Environment：Executor + Validator + FSM + JSON Enforcer。 | Python环境包结构、完整诊断episode样例 | 1–2周 |
| 阶段3：数据生成与离线强化训练 | 构造小规模高质量数据，运行DPO/ORPO离线强化训练。 | 训练集（每类≥2k条）、DPO/ORPO checkpoint | 2–3周 |
| 阶段4：在线交互强化 | 嵌入环境进行在线滚动学习，提升闭环行为。 | Online RL pipeline、奖励曲线收敛 | 2–4周 |
| 阶段5：端到端Agent原型 | 完整病例输入→系统自动完成诊断、报告、沟通。 | CLI/Web demo、报告验证与沟通 | 2周 |
| 阶段6：评测与优化 | 量化各子任务指标，调整奖励与数据。 | 报告通过率>90%、流程错误率<10% | 持续 |

## 二、近期重点任务（未来2–3周）

目标：实现可执行的最小环境（MVP），让模型在其中跑通一条完整闭环诊断。

### 第1周：协议与环境初始化

• 确认四个子任务的JSON Schema字段与奖励项（生成schema\_final.yaml）。  
• 搭建环境目录结构：engine/{executor,validator,fsm}，编写函数stub。  
• 完成Observation builder与Output enforcer。  
• 构造3条完整episode样例用于验证。

交付成果：schema\_final.yaml、最小环境骨架、demo病例样例。

### 第2–3周：MVP环境与初始训练

• 实现Executor与FSM的最小逻辑，支持execute/rollback/clarify。  
• 完成Validator模块，返回pass/errors/warnings。  
• 实现Reward Aggregator，用于计算Planner、Report、Comms的奖励。  
• 利用大模型生成小规模离线数据：Planner 100条，Intent 500条，Report与Comms各200条。  
• 运行DPO/ORPO离线训练，产出checkpoint。  
• 在环境中执行5条病例验证闭环执行与报告生成。

交付成果：可运行MVP环境、离线RL数据、首个训练checkpoint、自动报告样例。

## 三、阶段目标总结

在3周内应完成：  
✅ Schema与奖励定义文档；  
✅ MVP环境运行成功；  
✅ 离线RL数据准备完成；  
✅ 模型首次DPO/ORPO训练收敛；  
✅ 自动生成报告样例通过验证。

## 四、后续展望

• 接入真实影像接口（echocare/DICOM）。  
• 扩展至多病种（thyroid→liver→breast）。  
• 增加可视化界面展示轨迹、节点与报告。  
• 引入人工偏好反馈进行RL精调。