EHRFlow: A Large Language Model-Driven Iterative Multi-Agent Electronic Health Record Data Analysis Workflow

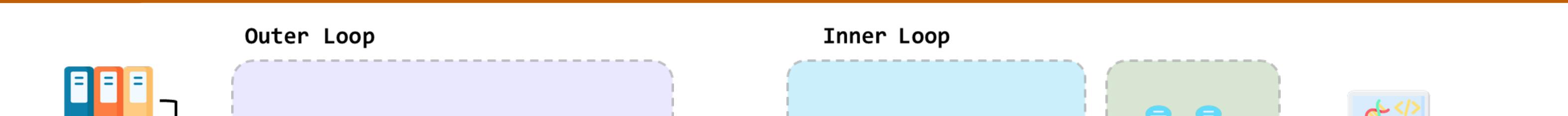
Hao Wu*, Yinghao Zhu*, Zixiang Wang, Xiaochen Zheng, Ling Wang, Wen Tang, Yasha Wang, Chengwei Pan, Ewen M. Harrison, Junyi Gao⁺, Liantao Ma⁺



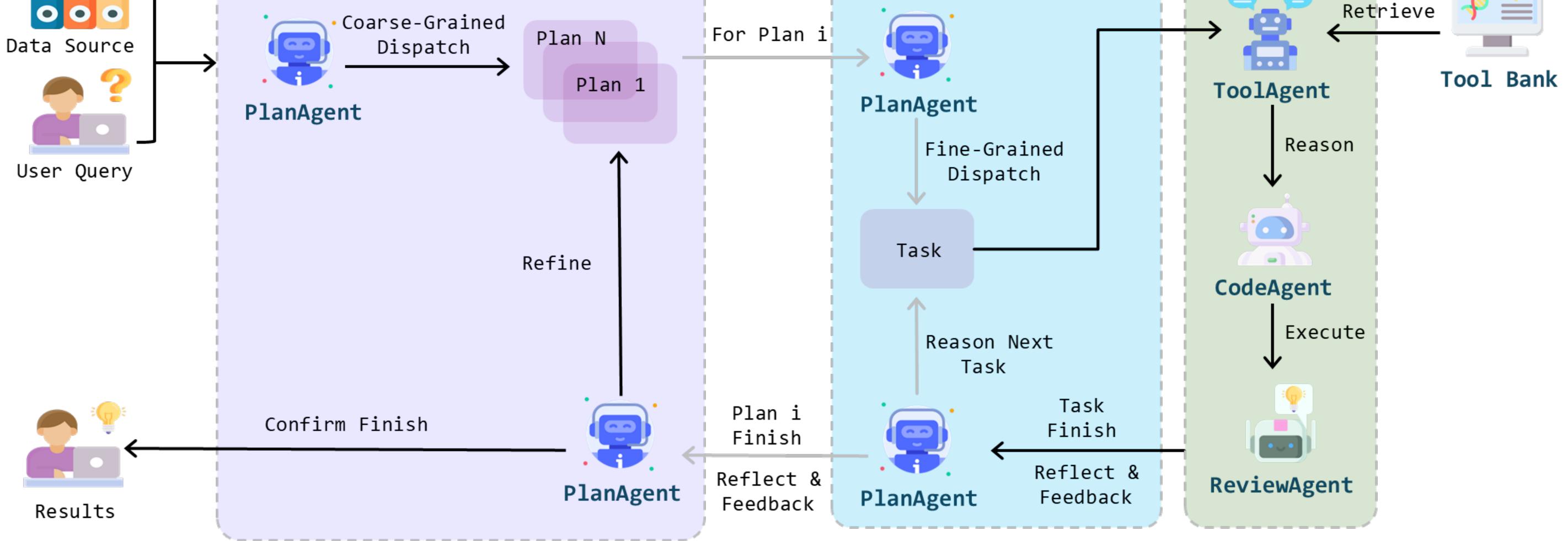
Boost EHR analysis pipeline for clinicians and medical researchers with LLMs

- Reduce the use barrier of data processing and deep learning for physicians and data scientists in EHR analysis.
- Avoid hallucinations and privacy concerns of applying LLMs in medical settings

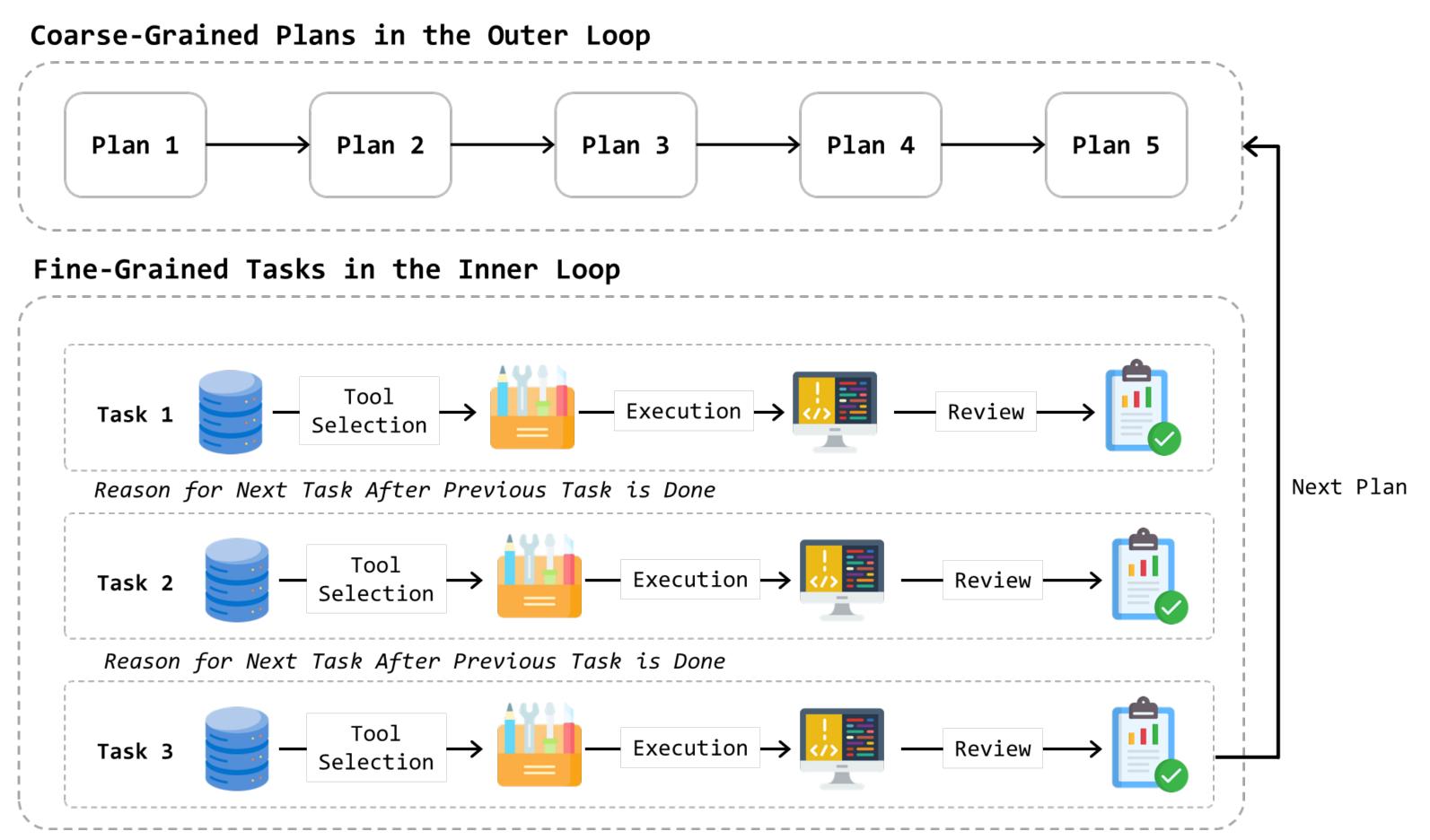
Objective: Complete complex EHR data analysis works using natural language. Clinicians set the targets, and the platform understand the targets, make plans and write codes, without human intervention.







Overview framework of EHRFlow



Motivation

Four main agents:

- PlanAgent, ToolAgent, CodeAgent, and ReviewAgent **Dual-loop mechanism:**
- Outer loop: Task decomposition into course-grained plans
- Inner loop: For each plan, fine-grained tasks dispatch, execution, and reflection

Privacy protection:

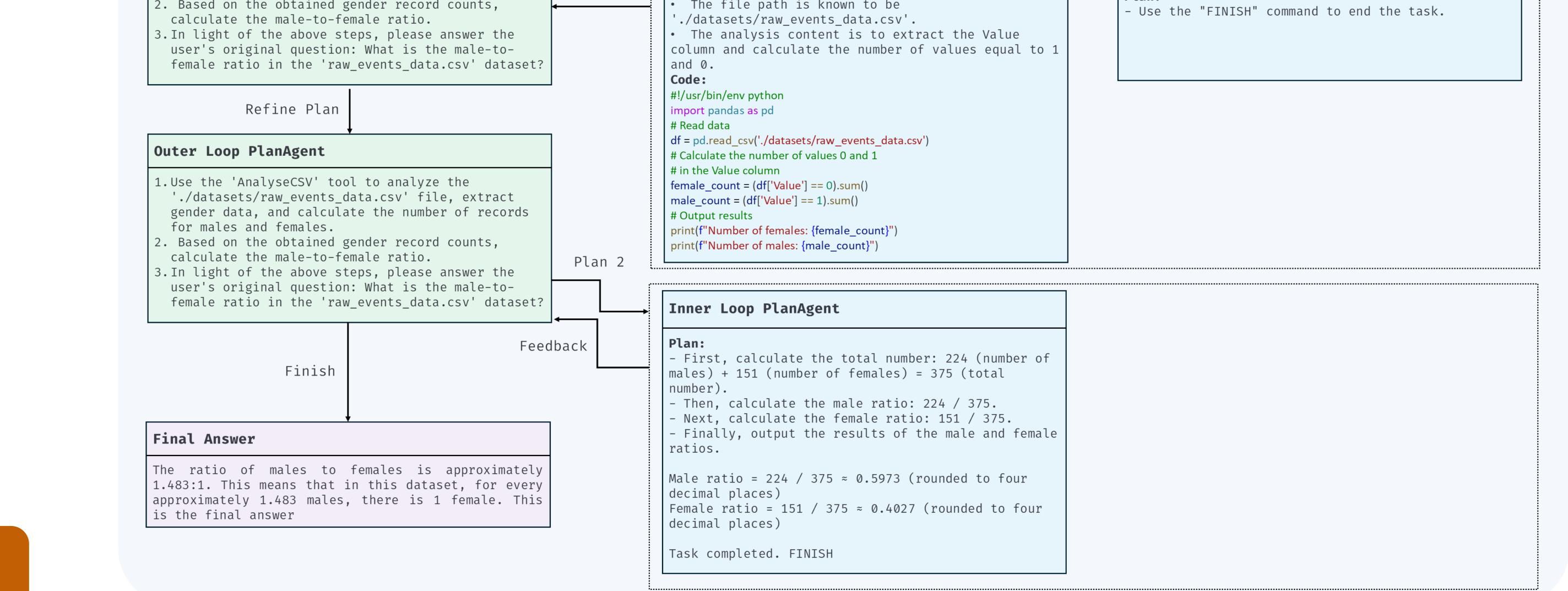
Only data schema provided, sensitive operations executed locally

Plan and task decomposition and execution pipeline

Pipeline

🔄 :Hello, I am EHRFlow, an intelligent assistant proficient in processing structured medical data. 🍯 After submitting the EHR data, you can freely ask me any questions or processing requirements about EHR medical data, and I will help you process the data according to your needs and return the corresponding results to assist you in better medical data analysis. Is there anything I can help you with now? 😇 :What is the male-to-female ratio in the raw_events_data.csv dataset?

	1			
Outer Loop PlanAgent		Inner Loop PlanAgent	Inner Loop PlanAgent	
1.Use the 'AnalyseCSV' tool to analyze the	Plan 1	Plan:		
'./datasets/raw_events_data.csv' file, extract		 Execute the AnalyseCSV tool to analyze the 		
gender data, and calculate the number of records		'./datasets/raw_events_data.csv' file.		
for males and females.	Feedback	 Select 'data analysis' as the analysis type. 	 Dlane	
2. Based on the obtained gender record counts,	<u> </u>	 The file path is known to be 	Plan:	



Case Study



GitHub Code