# **Technical Design Specification: Tri-Phasic Teacher-Student Agent**

Version: 3.2 (Revised)

Date: 15 June 2025

Status: Proposed (Based on Technical Research Report v17)

### 1. Architectural Overview

This specification details a simplified, robust, tri-phasic agentic system based on a Teacher-Student model. The architecture abandons the multi-agent committee in favor of two distinct, single agents: a Student (Gemini) for generation and a Teacher (DeepSeek) for evaluation. The system executes the Plan -> Synthesize -> Evaluate workflow with mandatory Human-in-the-Loop (HITL) gates. The entire process is governed by a unified constitutional\_prompt that defines the collaborative roles.

* Workflow: Planning (Student) -> HITL Plan Approval -> Synthesis (Student) -> Evaluation (Teacher) -> HITL Evaluation Review & Optional Restart.
* Supervisor (supervisor.py): The master orchestrator. Manages the sequential invocation of the three graphs. It will initialize two LLM clients (one for Gemini, one for DeepSeek) and pass the appropriate client to each graph-building function. The complex debate logic is removed.
* Agent Core (agent\_core.py): The code is simplified. It still provides three factory functions (build\_planning\_graph, build\_synthesis\_graph, build\_evaluation\_graph), but each graph is now a simple, single-agent workflow. The internal "self-correction" loops within each graph are retained.
* Configuration (prompts.yaml): The prompt file is updated, primarily with a new, comprehensive constitutional\_prompt.

### 2. LLM and Agent Configuration

* Student Agent:
  + Model: gemini-1.5-flash (or similar)
  + Temperature: 0.2 (low, for deterministic and factual generation)
  + Responsibility: Executes the Planning and Synthesis graphs.
* Teacher/Auditor Agent:
  + Model: deepseek-coder (or other appropriate DeepSeek model)
  + Temperature: 0.1 (very low, for precise and objective auditing)
  + Responsibility: Executes the Evaluation graph.

### 3. Constitutional System Prompt

The constitutional\_prompt in prompts.yaml will be updated to define this new collaborative dynamic. This prompt will be passed as the system prompt to both LLMs.

constitutional\_prompt: >

You are an expert AI assistant participating in a structured, three-phase research and generation workflow: Plan, Synthesize, and Evaluate. Your role is governed by the following principles:

## Overall Workflow:

1. \*\*Planning Phase:\*\* A 'Student' agent (Gemini) will generate a detailed plan based on the user's request and a knowledge base.

2. \*\*Synthesis Phase:\*\* After human approval of the plan, the 'Student' agent will generate a primary artifact with verifiable in-line citations.

3. \*\*Evaluation Phase:\*\* An independent 'Teacher' agent (DeepSeek) will formally audit the student's work for logical consistency, factual accuracy via citations, and alignment with the user's goals.

## Your Role and Collaborative Stance:

- This is a \*\*collaborative, not adversarial\*\*, Teacher-Student process. The goal is to produce the best possible final artifact for the user.

- The \*\*Student's\*\* task is to produce clear, well-structured, and honestly-cited work that is easy for the Teacher to audit.

- The \*\*Teacher's\*\* task is to provide precise, objective, and constructive feedback to help the Student improve in the next iteration.

- Both agents must adhere strictly to the provided knowledge base and the user's explicit instructions.

### 4. agent\_core.py Simplification

* Each build\_\*\_graph function will now construct a simple, linear graph for a single agent.
* The internal self-correction loops (e.g., holistic\_review -> refine\_draft) are kept, as they are still valuable for the Student to improve its own draft before submitting it for audit.
* The num\_revisions and critique fields in the GraphState will now refer to this internal self-correction, not a peer debate.

### 5. supervisor.py Simplification

* The run\_workflow function will be drastically simplified.
* Planning: It will invoke the planning\_graph once with the Student (Gemini) LLM. It will wait for user approval.
* Synthesis: It will invoke the synthesis\_graph once with the Student (Gemini) LLM.
* Evaluation: It will invoke the evaluation\_graph once with the Teacher (DeepSeek) LLM.
* The complex loops for managing constitutional debates are removed entirely. The only top-level loop is the optional full-cycle restart initiated by the user.

This revised design represents a significant simplification and strengthening of the system. It is more robust, efficient, and aligns with best practices for creating accountable and verifiable AI systems.