# Supervisor MCP Agent

A comprehensive supervisor agent that acts as an inline firewall and auditor for other agents, providing quality control, monitoring, and intervention capabilities.

## Features

* **Inline Design**: Acts as firewall + auditor between user input and agent output
* **Multi-Agent Orchestration**: Tracks multiple workers in parallel
* **Framework Agnostic**: Compatible with MCP, LangChain, AutoGen, and custom scripts
* **Tiered Response System**: Warning → Correction → Escalation
* **Comprehensive Monitoring**: Task completion, instruction adherence, output quality, error tracking
* **Learning & Adaptability**: Pattern recognition and knowledge base building
* **Trust Boundaries**: Automatic correction vs. human escalation
* **Detailed Reporting**: Real-time alerts, audit trails, confidence scoring

## Installation

cd supervisor-mcp-agent  
uv sync

## Usage

### Starting the MCP Server (STDIO mode)

sh run.sh

### Available Tools

1. monitor\_agent - Start monitoring an agent/task
2. set\_monitoring\_rules - Configure monitoring parameters and thresholds
3. get\_supervision\_report - Generate comprehensive supervision reports
4. intervene\_task - Pause/resume/restart agent tasks
5. validate\_output - Check output quality and adherence
6. get\_audit\_log - Retrieve detailed audit trails
7. configure\_escalation - Set escalation rules and notification preferences
8. knowledge\_base\_update - Add patterns to the failure knowledge base
9. rollback\_state - Restore to last known good state
10. generate\_summary - Create periodic supervision summaries

## Configuration

The supervisor supports various monitoring rules, escalation thresholds, and intervention strategies. See the tool documentation for detailed configuration options.

## Architecture

The supervisor operates as an inline monitoring system that: 1. Intercepts and validates agent inputs/outputs 2. Applies monitoring rules and quality checks 3. Implements tiered responses based on severity 4. Maintains audit trails and learning patterns 5. Provides real-time intervention capabilities

## License

MIT License