

PAPER #2

MODSIM World 2026 | Autonomy

May 22, 2026

Title

Society of Agents: A Constitutional Governance Architecture for LLM-Spawned Multi-Agent Simulation

Authors and Affiliations

Charles F. Day — Charles F. Day & Associates LLC (CFDay) / QWISE LLC

Projected Track

Autonomy

Abstract (455 words)

Recent work proposes Minimum Viable Coherence and Cohesion (MVC²) as a standardizable composite metric for determining whether LLM-spawned multi-agent systems remain within governed bounds. This paper presents the architecture that operationalizes that metric: the Society of Agents (SoA), a constitutional governance framework for multi-agent M&S in which agents are dynamically instantiated under constitutional governance and operated under SoA constitutional constraints within the defined test framework.

SoA enforces governance structurally, not through monitoring after the fact. The system is deterministically constrained within defined governance bounds, specifying role authorities, admissible actions, evidence requirements, and escalation conditions across human-in-the-loop (HITL) and human-on-the-loop (HOTL) modes. A governance-conditioned resolution gate ensures that every agent output passes through constitutional validation before inclusion in system results, producing auditable, reproducible decision pathways. A commercial LLM, operating as an adversarial stakeholder, probes constraint violations and surfaces latent assumptions through governed convergence—an iterative validation cycle that measures error modes and directs correction under human oversight.

The architecture extends to quantum-inspired simulation stacks, where quantum-inspired structural analogies applied to governance and coordination, implemented in classical systems, represent convergence behavior consistent with quantum-inspired annealing models. In this mode, a quantum-inspired portfolio optimization is supported by three layers of agentic AI with HITL governance and optional HOTL oversight, providing transparency, explainability, and agentic controls across the optimization lifecycle.

SoA operationalizes MVC² as a governance trigger. When MVC² is satisfied, the system operates with increasing AI operational responsibility within bounded constitutional controls. When MVC² falls below threshold, the architecture enforces automatic escalation: increased human oversight, constrained agent scope, or full reversion to human-directed operation. This makes

MVC² not merely a reporting metric but an enforceable governance boundary with deterministic consequences.

Within a defined problem scope and agent configuration appropriate to the task, we report two validation events demonstrating governed convergence under constitutional constraints. First, an adversarial validation in which a commercial LLM generated 88 structured test problems across 10 mathematical fields. Adversarial LLM evaluation indicated approximately 60 percent system accuracy. Initial performance confirmed that assessment. Human-on-the-loop observation identified a systematic error pattern; human-in-the-loop directed the system to self-assess and generate its own corrective instructions, which were incorporated as a new governance iteration. Post-correction accuracy reached 97.7 percent, with facet-level governance quality improving from 94.3 to 98.8 percent. The 40-percentage-point improvement quantifies the value of human-governed constitutional AI—accuracy achieved through governance intervention, not increased model capability. Second, in compliance evidence generation, HOTL workflows processed approximately 200 research-intensive questions at approximately 3 minutes per question, reduced from a 45-minute baseline (approximately 93 percent reduction), achieving 110 controls assessed with all 320 assessment objectives satisfied and zero deficiencies in independent C3PAO external audit. These results demonstrate that constitutional governance architecture enables trustworthy acceleration in the new epoch of agentic M&S.