Resilience in Ambient Multi-Agent LLMs via Decentralized Bio-Autonomic Control and Immune-Inspired Anomaly Detection

Nastaran Darabi[1]*, Devashri Naik[1], Sina Tayebati[1], Dinithi Jayasuriya[1], and Amit R. Trivedi[1]

[1]University of Illinois Chicago
ndarab2,dnaik6,stayeb3,dkasth2,amitrt@uic.edu

Abstract

Large Language Model (LLM) agents are now widely deployed in Ambient Intelligence (AmI) environments, where autonomous agents must sense, act, and coordinate at scale. As agent capabilities and interdependence increase, traditional reliability strategies such as isolated adaptive control, anomaly detection, or trust modeling have proven inadequate due to their fragmented and scenario-specific nature. Comprehensive architectures that enable integrated self-management, collective anomaly response, robust information dissemination, and privacy-preserving adaptation remain scarce. We propose a *bio-autonomic framework* for decentralized resilience in multi-agent LLM systems where a unified architecture systematically applies principles from biological autonomic systems to LLM-based multi-agent environments. Specifically, each agent implements an autonomic control loop, formally structured as Monitor-Analyze-Plan-Execute over a shared Knowledge base (MAPE-K), for self-regulation. At the system level, the framework integrates immune-inspired anomaly detection using the Dendritic Cell Algorithm, probabilistic computational trust, decentralized gossip for robust information sharing, and federated learning with homomorphic encryption for collaborative, privacy-preserving adaptation. This holistic approach enables LLM agent ecosystems to self-organize, detect and isolate faults, and collectively adapt as system complexity increases. Empirical evaluations show that our framework achieves substantially improved resilience and recovery compared to state-of-the-art multi-agent baselines.

1 Introduction

The evolution of artificial intelligence has enabled the Ambient Intelligence (AmI) paradigm, where digital environments are populated by autonomous agents that sense, act, and adapt within daily life (Cook, Augusto, and Jakkula 2009; Dunne, Morris, and Harper 2021; Martinez-Martin et al. 2021). The advent of powerful Large Language Models (LLMs) has significantly advanced this vision, equipping agents with sophisticated reasoning, communication, and planning abilities. As a result, LLM agents now underpin complex Multi-Agent Systems (MAS) for applications such as smart cities, automated logistics, personalized healthcare, and collaborative scientific discovery (Kalyuzhnaya et al. 2025; Jannelli et al. 2024; Wang et al. 2025).

However, the increased autonomy and complexity of LLM agents introduce new reliability challenges. Unlike traditional software, which fails in predictable ways, LLM agents exhibit subtle, contextual, and emergent failure modes. These include logical errors, misinterpretation of nuanced instructions, hallucinated content, or susceptibility to prompt injection, any of which can propagate failures throughout a system (Liu et al. 2023; Li et al. 2024; Kong et al. 2025). Standard engineering methods, which aim to build robustness against known fault types, are insufficient: a system hardened only to known threats remains brittle and fragile when faced with "unknown unknowns" in open environments (Owotogbe 2025; Kott and Abdelzaher 2014).

While foundational MAS frameworks such as MADDPG (Lowe et al. 2017) and MAPPO (Yu et al. 2022) have advanced cooperative policy learning, they primarily focus on optimizing task-specific rewards and assume reliable agent behavior. More recent architectures like H-MAS (Ghavamzadeh, Mahadevan, and Makar 2006) and AutoAgents (Chen et al. 2023) introduce hierarchical structures and automated agent generation to manage complexity in solving user-defined tasks. However, these approaches do not explicitly address the systemic resilience needed to handle the subtle and emergent failure modes of modern LLM agents in open-ended environments. Their focus remains on performance and coordination rather than on the capacity to absorb, adapt to, and recover from unforeseen disruptions, a gap our work directly targets. To address this critical gap, we propose a shift from static robustness to dynamic, systemic **resilience**. Resilience is the capacity of a system to absorb unexpected shocks, adapt its behavior and structure, and recover gracefully while maintaining essential functions. We introduce the **Bio-Autonomic Framework**, an architecture designed specifically to enable resilient ambient LLM agent ecosystems.

Our framework unifies two perspectives: a top-down, engineering-driven approach for individual agent self-management and a bottom-up, bio-inspired strategy for emergent collective resilience. At its core, each agent implements an autonomic control loop, formally structured as Monitor-Analyze-Plan-Execute over a shared knowledge base (MAPE-K) (Arcaini, Riccobene, and Scandurra 2015;

*Corresponding Author

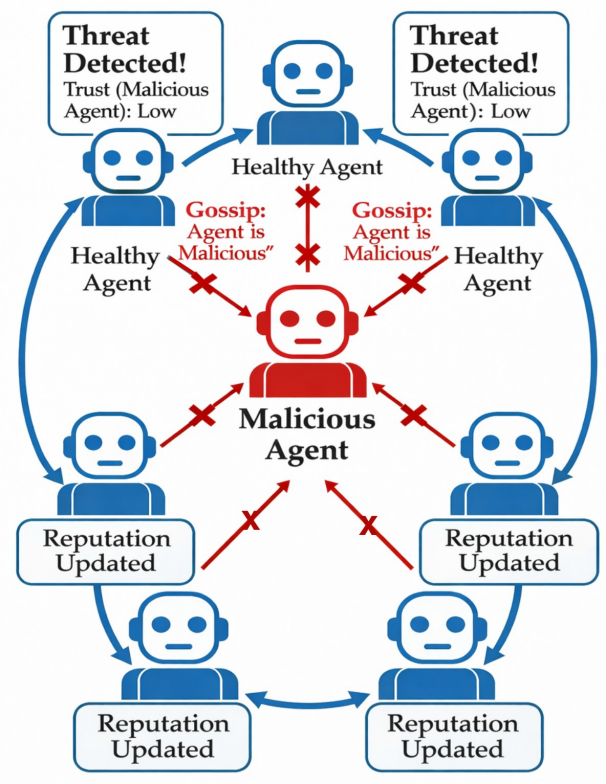

Figure 1: Overview of the Gossip Protocol: Healthy agents share trust information via gossip protocol,and confirmed threats trigger reputation updates, isolating the malicious agent.

Souza et al. 2025). This is augmented by a bio-inspired layer that incorporates the Dendritic Cell Algorithm for nuanced anomaly detection, a dynamic computational trust model, and a fault-tolerant gossip protocol for reputation sharing (Fan et al. 2020; Somvanshi et al. 2025), as illustrated in Figure 1. A key architectural element is Federated Learning with Homomorphic Encryption, which enables collaborative adaptation of detection and trust models while preserving data privacy (Jahan, Rahman, and Wang 2025). The entire system is rigorously grounded in the Decentralized Partially Observable Markov Decision Process (Dec-POMDP) formalism, providing a principled foundation for distributed coordination (Pey et al. 2025).

2 Background and Related Works

The Robustness-Resilience Dilemma. In complex systems, robustness and resilience are distinct properties. Robustness is the capacity to maintain function under a predefined set of perturbations, emphasizing resistance to known stressors. Resilience, by contrast, is a dynamic property describing the ability to adapt to shocks, absorb disruptions, and recover gracefully. While robustness is about resisting change, resilience centers on adaptation and recovery. In the unpredictable context of AmI and LLM-driven MAS, strategies focused only on static robustness are insufficient. Achieving systemic resilience is essential for long-term viability (Baldoni, Baroglio, and Micalizio 2020; Homayounfar et al. 2018).

Autonomic Computing and MAPE-K Loop. Autonomic computing, inspired by the biological autonomic nervous system, aims to build systems that manage their own complexity (Gill et al. 2022). This is achieved by endowing components with self-managing capabilities: Self-Configuring, Self-Healing, Self-Optimizing, and Self-Protecting (self-CHOP) (Lewis, Rouff, and Tekeoglu 2023). The canonical architectural pattern is the **MAPE-K control loop**, which consists of four phases supported by a shared knowledge base. The **Monitor** phase gathers data from internal and external sources; the **Analyze** phase processes this data to diagnose problems and detect patterns; the **Plan** phase formulates a sequence of actions; and the **Execute** phase implements the plan. All phases are underpinned by the continuously updated **Knowledge** component, enabling the system to learn and adapt over time (Arcaini, Riccobene, and Scandurra 2015; Souza et al. 2025).

Bio-Inspired Mechanisms in MAS. Nature provides powerful blueprints for decentralized, resilient systems. Two bio-inspired classes are especially relevant:

Artificial Immune Systems and the Dendritic Cell Algorithm. Artificial Immune Systems (AIS) are algorithms modeled after the vertebrate immune system, a paradigm of decentralized, adaptive, memory-based anomaly detection (Myakala, Bura, and Jonnalagadda 2025). A particularly relevant AIS is the **Dendritic Cell Algorithm (DCA)**, inspired by dendritic cells that process signals to differentiate between threats, malfunctions, and normal activity. The DCA correlates three types of input: **Pathogen-Associated Molecular Patterns (PAMPs)** indicate known threats (e.g., prompt injection attacks); **Danger-Associated Molecular Patterns (DAMPs)** signal unexpected stress, such as abnormal use of APIs or non-sensical LLM output; and **Safe Signals (SS)** indicate normal operation. By processing these signals, the DCA computes a **Mature Context Antigen Value (MCAV)**, quantifying the context-sensitive "danger" of an event. This mechanism provides nuanced, context-aware threat assessment, suited for LLM agent interactions (Kazari, Shereen, and Dán 2023; Yang et al. 2014).

Gossip (Epidemic) Protocols. Gossip protocols are communication mechanisms inspired by spreading rumors. Each agent periodically shares information with a random subset of peers, who then do the same. This process enables robust, scalable and rapid information dissemination without centralized coordination, making the system inherently resilient to node failures (Gonçalves et al. 2024; Bachrach et al. 2009; Sulaiman et al. 2024; Lian et al. 2017).

Federated Learning and Homomorphic Encryption. A core challenge in MAS is to support collective learning without compromising the privacy of individual agents. Federated Learning (FL) addresses this by enabling agents to collaboratively train a global model without sharing raw data. A central server distributes a global model, which agents train locally on their private data; only model updates are sent back and aggregated to improve the global model (Jahan, Rahman, and Wang 2025; Chakraborty and Boudguiga

2024; Zhang, Lin, and Zhang 2022). To further enhance privacy, Homomorphic Encryption (HE) allows computation directly on encrypted data. For example, an additively homomorphic scheme enables the server to aggregate encrypted model updates without ever decrypting them, preventing inspection of individual agent contributions and providing strong privacy guarantees (Jin et al. 2023; Yuan et al. 2025).

3 Decentralized Resilience in Multi-Agent LLMs via Bio-Autonomic Control

We model the LLM agent ecosystem as a Decentralized Partially Observable Markov Decision Process (Dec-POMDP), defined by the tuple $\mathcal{M} = \langle \mathcal{I}, S, \{A_i\}_{i\in\mathcal{I}}, P, R, \{\Omega_i\}_{i\in\mathcal{I}}, O, \gamma\rangle$. This formalism captures the key features of the system: multiple cooperative agents, stochastic dynamics, and incomplete information. Here, $\mathcal{I}$ is the set of N LLM agents, S is the set of hidden global states, and A_i is the action set for agent i. The transition function $P(s' \mid s, \mathbf{a})$ specifies the probability of moving from s to s' under the joint action $\mathbf{a} = (a_1, \ldots, a_N)$. The collective reward $R(s, \mathbf{a})$ provides scalar feedback to the entire system. Agents do not observe s directly; instead, each agent i receives a local observation $o_i \in \Omega_i$ sampled from $O(\mathbf{o} \mid s', \mathbf{a})$, with $\mathbf{o} = (o_1, \ldots, o_N)$. Rewards are discounted by $\gamma \in (0, 1)$. Finding the optimal joint policy π^* is generally intractable. The Bio-Autonomic Framework therefore uses structured and scalable heuristics that allow agents to implement local policies π_i and collectively approximate robust joint behavior.

3.1 Individual Autonomy: The MAPE-K Loop

The autonomous behavior of each agent is governed by a continuous, adaptive Monitor-Analyze-Plan-Execute-Knowledge (MAPE-K) loop, which forms the agent's internal cognitive architecture and processes observations, generates actions, and updates the agent's internal state.

Monitor: Signal Extraction At each time step t, agent i interacts with peer j and receives a raw observation vector $\mathbf{x}_{ij}^{(t)} \in \mathcal{X}$, which may include textual responses, API metrics, or communication latencies. The **Monitor** phase applies a shared feature extractor $\Phi_{\mathbf{w}}$, parameterized by learnable weights $\mathbf{w}$, to convert this high-dimensional input into a triplet of key signals:

$$(s_{ij}^{(t)}, d_{ij}^{(t)}, p_{ij}^{(t)}) = \Phi_{\mathbf{w}}(\mathbf{x}_{ij}^{(t)})$$

Here, $s_{ij}^{(t)} \in \mathbb{R}^+$ is the Safe Signal, indicating normal or beneficial interactions. $d_{ij}^{(t)} \in \mathbb{R}^+$ is the Danger-Associated Molecular Pattern (DAMP), reflecting anomalous or potentially harmful behavior. $p_{ij}^{(t)} \in \mathbb{R}^+$ is the Pathogen-Associated Molecular Pattern (PAMP), associated with known malicious or adversarial signatures. The weights $\mathbf{w}$ are continuously adapted via collective learning, ensuring agents remain sensitive to evolving environmental conditions. These extracted signals apply to the Analyze phase.

Analyze: Contextual Threat Assessment The **Analyze** phase interprets monitored signals using a mechanism inspired by the Dendritic Cell Algorithm (DCA), integrating evidence to assess interaction context. For each agent pair (i, j), agent i maintains cumulative counters $\{C_S, C_D, C_P\}$. Upon receiving $(s_{ij}^{(t)}, d_{ij}^{(t)}, p_{ij}^{(t)})$, the counters are updated:

$$C_S \leftarrow C_S + s_{ij}^{(t)}, \quad C_D \leftarrow C_D + d_{ij}^{(t)}, \quad C_P \leftarrow C_P + p_{ij}^{(t)}$$

Signal accumulation continues until the total $C_S + C_D + C_P$ exceeds a threshold θ_{mat}, at which point a "maturation" event triggers context assessment. The agent then computes a **Mature Context Antigen Value (MCAV)**, a normalized threat score:

$$k_{ij}^{(t)} = \frac{(w_D C_D + w_P C_P) - w_S C_S}{w_D C_D + w_P C_P + w_S C_S} \tag{1}$$

Here, w_S, w_D, and w_P are agent-specific sensitivity weights for each signal type. The MCAV $k_{ij}^{(t)}$ is bounded in $[-1, 1]$: positive values indicate dangerous or malicious contexts, while negative values indicate safety. After each assessment, counters are reset, and a new context period begins. The MCAV serves as the primary output of the Analyze phase, guiding both immediate action planning and long-term knowledge updates.

Plan & Execute: Risk-Averse Action Selection The **Plan** phase determines agent i's next action $a_i^{(t)}$ by integrating short-term contextual threat scores from the Analyze phase with long-term trust beliefs. Decision-making is based on maximizing an expected utility function that incorporates explicit risk aversion:

$$a_i^{(t)} = \arg\max_{a\in A_i} \left[\mathbb{E}[U(a)] - \lambda \cdot f(k_{ij}^{(t)}) \cdot (1 - T_{ij}^{(t)})\right] \tag{2}$$

Here, $\mathbb{E}[U(a)]$ is the baseline expected utility of action a based on primary task objectives. The second term is a risk-aversion penalty: $\lambda \geq 0$ is the agent's risk sensitivity, $f(k)$ is a non-decreasing penalty function (e.g., $f(k) = \max(0, k)$) that increases with assessed danger $k_{ij}^{(t)}$, and $T_{ij}^{(t)} \in [0, 1]$ is the agent's trust in peer j.

This formulation fuses immediate and historical risk: high danger ($k_{ij}^{(t)}$) and low trust ($T_{ij}^{(t)}$) strongly discourage risky actions, while trust in an agent attenuates the impact of isolated anomalous events. The **Execute** phase then implements the chosen action $a_i^{(t)}$.

Knowledge Update: Trust and Reputation Adjustment After each action, the agent updates its **Knowledge** base, which primarily encodes direct trust in peers. The trust that agent i has in agent j, T_{ij}, is modeled as a Beta distribution: $T_{ij} \sim \text{Beta}(\alpha_{ij}, \beta_{ij})$, where α_{ij} and β_{ij} represent evidence for cooperative and uncooperative behavior, respectively.

The MCAV score $k_{ij}^{(t)}$ from the Analyze phase determines how these parameters are adjusted. Updates are applied as:

$$\alpha_{ij}^{(t+1)} = \alpha_{ij}^{(t)} + \sigma(-c \cdot k_{ij}^{(t)}) \tag{3}$$

$$\beta_{ij}^{(t+1)} = \beta_{ij}^{(t)} + \sigma(c \cdot k_{ij}^{(t)}) \tag{4}$$

where $\sigma(z) = (1 + e^{-z})^{-1}$ is the sigmoid function and c is a scaling factor. Thus, safe contexts ($k < 0$) increase α_{ij}, while dangerous contexts ($k > 0$) increase β_{ij}.

The agent's current trust in j is given by the expectation:

$$T_{ij}^{(t+1)} = \frac{\alpha_{ij}^{(t+1)}}{\alpha_{ij}^{(t+1)} + \beta_{ij}^{(t+1)}} \tag{5}$$

This value $T_{ij}^{(t+1)}$ is then used in the next planning cycle with agent j (see Eq. 2).

3.2 Collective Resilience: The Social Fabric

While individual autonomy enables agents to adapt, system-wide resilience to widespread threats or coordinated attacks depends on collective mechanisms that form a robust social fabric. The trust update mechanism (Eq. 5) controls these collective dynamics.

Gossip Algorithm for Reputation Dissemination To prevent knowledge about agent behavior from remaining isolated, the framework employs a gossip algorithm to disseminate trust information. This converts private, pairwise trust into a shared, network-wide reputation. Each agent i maintains a local reputation matrix $\mathbf{R}_i^{(t)} \in \mathbb{R}^{N \times N}$, where $R_{ik}^{(t)}$ is agent i's belief about agent k's trustworthiness. Direct opinions are updated as $R_{ij}^{(t)} = T_{ij}^{(t)}$ after interactions.

At intervals of τ_g time steps, agent i randomly selects a peer p for a gossip exchange, sharing their reputation matrices. Agent i then updates its local matrix by averaging:

$$\mathbf{R}_i^{(t+1)} = (1 - w_g)\mathbf{R}_i^{(t)} + w_g \mathbf{R}_p^{(t)}$$

where $w_g \in (0, 1)$ is the gossip weight (often 0.5). This iterative process enables decentralized, robust consensus, allowing information about malicious or faulty agents to propagate quickly. The resulting "social immune response" ensures that threats are isolated before causing systemic harm.

3.3 Collective Adaptation: Federated Learning with Homomorphic Encryption

Long-term adaptation is achieved through collective training of the shared signal extractor $\Phi_{\mathbf{w}}$, used by all agents in the Monitor phase. The objective is to find weights $\mathbf{w}^*$ minimizing the global loss over all agents' private datasets $\{D_1, \ldots, D_N\}$:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_{i=1}^{N} \frac{|D_i|}{\sum_k |D_k|} L_i(\mathbf{w})$$

where $L_i(\mathbf{w}) = \frac{1}{|D_i|} \sum_{(\mathbf{x},y) \in D_i} \ell(\Phi_{\mathbf{w}}(\mathbf{x}), y)$ is the local loss for agent i. This optimization uses Federated Averaging, enhanced with Homomorphic Encryption for privacy.

In each round t, a central server selects agents S_t to participate. Each agent $i \in S_t$ receives the current global model $\mathbf{w}_t$, computes its local gradient $\mathbf{g}_i = \nabla L_i(\mathbf{w}_t)$, and update $\Delta_i = -\eta \mathbf{g}_i$. To protect privacy, agent i encrypts Δ_i using a public key pk of an additively homomorphic scheme, yielding ciphertext $c_i = \text{Encrypt}_{pk}(\Delta_i)$. These are sent to the server, which aggregates without decryption:

$$C_{\text{agg}} = \bigoplus_{i \in S_t} c_i = \text{Encrypt}_{pk}\left(\sum_{i \in S_t} \Delta_i\right)$$

The server decrypts C_{agg} with its private key to obtain $\Delta_{\text{agg}} = \sum_{i \in S_t} \Delta_i$, updating the global model as:

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \frac{1}{|S_t|} \Delta_{\text{agg}}$$

This ensures that agents' private data is never exposed, while the global model becomes more effective over time.

A critical innovation is self-supervised labeling of training data. Agents autonomously generate labels $y \in \{0, 1\}$, where 1 denotes anomaly, using two signals: (1) *reward-based feedback* from the Dec-POMDP (labeling interactions as anomalous if they precede low or negative rewards), and (2) *consensus-based feedback* from the social layer (labeling interactions as anomalous if they cause a sharp negative trust update, especially when confirmed by peer gossip). This closes the loop between collective system performance and perception, enabling adaptation to harmful behaviors based on their demonstrable impact on overall system goals.

3.4 Collective Stability with Social Consensus

Agent-Level Stability. The individual stability is ensured by two core mechanisms. *First*, the trust update loop models T_{ij} as the expectation of a Beta distribution, with updates governed by bounded sigmoid functions. This constrains trust beliefs to $[0, 1]$, preventing divergence or abrupt swings. *Second*, the risk-averse action policy (Eq. 2) systematically penalizes risky actions toward untrusted or anomalous peers, reducing exposure to adversarial influence and containing potential cascading failures.

Collective Reputation Convergence. At the system level, stability emerges from the gossip-based reputation protocol, which acts as a distributed consensus algorithm. The iterative update $\mathbf{R}_i^{(t+1)} = (1 - w_g)\mathbf{R}_i^{(t)} + w_g \mathbf{R}_p^{(t)}$ guarantees convergence to a shared consensus matrix $\bar{\mathbf{R}}$ under standard graph connectivity. The convergence rate depends on the second-largest eigenvalue of the expected gossip matrix; a smaller value implies faster, more robust error correction. This process prevents the network from fragmenting into subgroups with divergent beliefs, a common mode of failure in large-scale multi-agent systems.

Stability of Collective Learning. The federated learning mechanism is susceptible to client drift under non-IID data. The Bio-Autonomic Framework regulates this through cross-layer feedback: if an agent's contribution degrades system performance, peers detect this via low MCAV scores and reputation, reducing the agent's influence in future rounds. This feedback loop contains destabilizing updates and prevents model collapse.

Framework	Dataset	GPT-4o	o1-mini	GPT-3.5	Mixtral-22B	Llama-8B	Llama-70B	Qwen-72B	Qwen-32B	Qwen-14B	Qwen-7B
Bio-Autonomic	HumanEval	**88.52**	**85.24**	**79.58**	**77.71**	**75.13**	**83.87**	**86.84**	**83.19**	**82.35**	**79.52**
	CIAR	**72.89**	**69.41**	**65.23**	**63.76**	**62.08**	**69.25**	**70.91**	**67.14**	**67.92**	**65.73**
	CommonMT	**88.46**	**86.03**	**79.47**	**77.42**	**75.18**	**84.41**	**86.43**	**85.15**	**82.11**	**79.59**
	FairEval	**93.58**	**90.92**	**84.35**	**82.17**	**80.01**	**89.04**	**92.11**	**88.83**	**87.09**	**84.26**
AutoAgents	HumanEval	81.23	76.88	72.74	71.79	69.05	76.81	79.72	75.16	75.28	72.77
	CIAR	60.71	57.45	54.29	52.92	51.56	57.84	59.57	54.23	56.69	54.22
	CommonMT	84.49	81.91	75.83	73.84	71.82	80.06	82.65	80.38	78.47	75.87
	FairEval	88.82	86.07	80.15	78.29	75.51	84.48	86.89	84.34	82.26	80.03
H-MAS	HumanEval	82.57	77.92	73.81	72.76	70.33	78.69	81.38	76.39	76.81	74.45
	CIAR	62.34	58.69	55.82	54.65	52.97	59.08	61.32	55.98	58.17	56.41
	CommonMT	84.93	82.18	76.54	74.99	71.91	80.35	82.96	80.92	79.33	76.32
	FairEval	90.06	87.11	80.69	78.93	76.24	85.12	88.54	85.95	84.18	80.67
MAPPO	HumanEval	62.78	58.91	56.13	55.69	53.82	59.94	61.35	57.26	58.01	56.27
	CIAR	40.35	37.12	36.18	35.03	34.59	38.16	39.77	35.74	37.38	35.81
	CommonMT	69.72	66.59	62.61	61.18	59.24	62.65	67.81	65.43	65.19	66.38
	FairEval	65.37	61.74	58.52	57.01	55.78	62.29	64.25	60.59	60.41	58.46
MADDPG	HumanEval	58.19	54.43	51.92	51.68	49.21	54.88	57.17	53.05	53.82	52.74
	CIAR	34.96	31.27	31.14	30.82	29.89	33.25	34.38	30.01	32.24	31.06
	CommonMT	63.11	60.84	56.45	55.29	53.47	60.13	61.62	59.66	58.89	56.88
	FairEval	58.73	55.88	52.51	51.46	49.52	55.97	57.49	54.85	55.08	52.64

Table 1: Task-level comparison of Bio-Autonomic with five competitive multi-agent frameworks on four benchmarks: *HumanEval* (code generation), *CIAR* (mathematical reasoning), *CommonMT*-Lexical (commonsense translation), and *FairEval* (text-quality preference alignment). Each task is evaluated on ten LLM backbones, with scores rescaled to the $[0, 100]$ band. Boldface marks the best result in each (task, backbone) cell. All values are averaged over ten random seeds to reduce variance.

Emergent, Multi-Timescale Stability. Overall stability in the framework is an emergent property of layered control loops operating on different timescales:

- **Fast, Local Stability:** Risk-averse planning ensures immediate response to anomalies.
- **Mid-Term, Social Stability:** Gossip and trust dynamics rapidly isolate and marginalize non-cooperative or faulty agents.
- **Long-Term, Adaptive Stability:** Federated learning adapts collective perception, with social feedback policing model drift.

This defense-in-depth strategy, similar to Lyapunov stability (Nguyen 2018), ensures quick recovery and strong operation under a wide range of disturbances, providing resilience against both random shocks and deliberate attacks.

4 Multi-Agent Interaction Setup

4.1 Tasks and Models

Framework performance was assessed on five distinct tasks, following the protocol of (Huang et al. 2024). For code generation, we used the **HumanEval** benchmark (Chen et al. 2021), containing 164 Python problems. Mathematical reasoning was evaluated with the **CIAR** benchmark (Liang et al. 2024), a suite of 50 challenging word problems. Machine translation was tested with 100 sentences from **CommonMT** (He et al. 2025), requiring contextual understanding for English-to-French translation. Textual preference alignment was assessed with **FairEval** (Wang et al. 2024), comprising 80 comparisons between ChatGPT and Vicuna-13B. Finally, collaborative ability was evaluated in **Collab-Overcooked AI** environment (Sun et al. 2025), a multi-agent game demanding teamwork and communication.

To isolate framework performance from the capabilities of any single model, we used ten large language models (LLMs), both closed-source (GPT-4o (OpenAI 2024), o1-mini (Jaech et al. 2024), GPT-3.5 (OpenAI 2023)) and open-source (Qwen2.5-Instruct (Team 2024), Llama3.1-Instruct (Dubey et al. 2024), Mixtral-8 (Jiang et al. 2024), among others). All models were evaluated without additional fine-tuning, using a temperature of 0.2 to focus on the framework's control logic.

4.2 Baselines and Fault Testing

We compared our framework against four existing multi-agent systems: **MADDPG** (Lowe et al. 2017), **MAPPO** (Yu et al. 2022), **H-MAS** (Ghavamzadeh, Mahadevan, and Makar 2006), and **AutoAgents** (Chen et al. 2023). All baselines were evaluated in the same environment for direct comparison. Resilience was assessed using two fault injection strategies (Huang et al. 2024): **AUTOINJECT**, which randomly modifies an agent's message with 20% probability (by word swaps or negations), and **AUTOTRANSFORM**, which replaces a regular agent with a "malicious" (misleading) or "clumsy" (error-prone) agent.

We report four metrics:

1. **Performance Degradation under Faults (PDF):** Quantifies the initial performance drop after fault injection:

$$\text{PDF} = \frac{S_{\text{benign}} - S_{\text{faulty}}}{S_{\text{benign}}} \times 100$$

2. **Error-Recovery Rate (ERR):** Measures the fraction of lost performance that is autonomously recovered:

$$\text{ERR} = \frac{S_{\text{recovered}} - S_{\text{faulty}}}{S_{\text{benign}} - S_{\text{faulty}}} \times 100$$

Model	Method	Level 1		Level 2		Level 3		Level 4		Level 5		Level 6	
		SR	**PC**	**SR**	**PC**	**SR**	**PC**	**SR**	**PC**	**SR**	**PC**	**SR**	**PC**
GPT-4o	Collab-Overcooked	94.00	85.92	86.00	84.96	68.00	76.61	34.00	44.42	2.00	29.13	4.00	22.45
	AUTOTRANSFORM	65.80	60.14	55.90	55.22	40.80	45.97	20.40	31.10	1.00	17.48	2.00	13.47
	Bio-Autonomic	**95.00**	**87.15**	**89.00**	**86.50**	**75.00**	**80.10**	**48.00**	**55.40**	**15.00**	**40.50**	**20.00**	**35.80**
o1-mini	Collab-Overcooked	70.00	74.18	2.00	36.36	0.00	33.60	0.00	24.80	0.00	20.28	0.00	13.07
	AUTOTRANSFORM	42.00	44.51	0.00	21.82	0.00	20.16	0.00	14.88	0.00	12.17	0.00	7.84
	Bio-Autonomic	**85.00**	**81.30**	**65.00**	**70.20**	**58.00**	**65.90**	**35.00**	**45.10**	**10.00**	**31.80**	**12.00**	**25.50**
GPT-3.5	Collab-Overcooked	42.00	68.20	8.00	43.42	0.00	36.44	0.00	24.74	0.00	15.21	0.00	12.03
	AUTOTRANSFORM	25.20	40.92	4.80	26.05	0.00	21.86	0.00	14.84	0.00	9.13	0.00	7.22
	Bio-Autonomic	**82.00**	**79.50**	**60.00**	**68.00**	**50.00**	**61.30**	**33.00**	**43.20**	**8.00**	**29.90**	**10.00**	**24.10**
Qwen2.5-72B	Collab-Overcooked	78.00	76.84	64.00	68.00	14.00	46.88	8.00	30.80	0.00	22.67	0.00	18.45
	AUTOTRANSFORM	54.60	53.79	44.80	47.60	8.40	28.13	4.80	18.48	0.00	13.60	0.00	11.07
	Bio-Autonomic	**88.00**	**83.10**	**75.00**	**78.50**	**65.00**	**72.40**	**40.00**	**50.60**	**12.00**	**38.10**	**15.00**	**30.20**
Qwen2.5-32B	Collab-Overcooked	64.00	73.36	44.00	62.02	14.00	40.08	4.00	33.78	2.18	22.16	0.00	18.93
	AUTOTRANSFORM	44.80	51.35	30.80	43.41	9.80	28.06	2.40	20.27	1.31	13.30	0.00	11.36
	Bio-Autonomic	**84.00**	**80.20**	**68.00**	**73.00**	**55.00**	**65.80**	**34.00**	**44.50**	**9.00**	**30.00**	**11.00**	**25.30**
Qwen2.5-14B	Collab-Overcooked	32.00	50.36	4.00	26.66	0.00	24.41	0.00	19.00	0.00	14.14	0.00	14.27
	AUTOTRANSFORM	22.40	35.25	2.80	18.66	0.00	17.09	0.00	11.40	0.00	8.48	0.00	8.56
	Bio-Autonomic	**79.00**	**75.00**	**55.00**	**62.10**	**45.00**	**55.40**	**28.00**	**39.80**	**6.00**	**25.10**	**8.00**	**21.90**
Qwen2.5-7B	Collab-Overcooked	8.00	44.79	0.00	13.00	0.00	9.29	0.00	8.35	0.00	5.57	0.00	4.51
	AUTOTRANSFORM	5.60	31.35	0.00	9.10	0.00	6.50	0.00	5.01	0.00	3.34	0.00	2.71
	Bio-Autonomic	**75.00**	**70.60**	**50.00**	**58.30**	**40.00**	**51.20**	**25.00**	**35.00**	**5.00**	**22.40**	**7.00**	**19.50**
Llama3.1-70B	Collab-Overcooked	70.00	75.42	42.00	63.15	22.00	54.58	6.18	45.04	0.00	29.77	0.00	17.69
	AUTOTRANSFORM	49.00	52.80	29.40	44.21	15.40	38.21	3.71	27.02	0.00	17.86	0.00	10.61
	Bio-Autonomic	**86.00**	**82.40**	**72.00**	**76.80**	**60.00**	**69.10**	**38.00**	**48.90**	**11.00**	**36.50**	**14.00**	**29.80**
Llama3.1-8B	Collab-Overcooked	4.00	33.03	0.00	15.49	0.00	12.33	0.00	11.24	0.00	9.05	0.00	7.45
	AUTOTRANSFORM	2.80	23.12	0.00	10.84	0.00	8.63	0.00	6.74	0.00	5.43	0.00	4.47
	Bio-Autonomic	**72.00**	**68.10**	**48.00**	**55.90**	**38.00**	**49.70**	**22.00**	**33.30**	**4.00**	**20.10**	**6.00**	**18.20**

Table 2: Performance comparison for different LLMs across 6 task complexity levels. "Collab-Overcooked" is the baseline performance of the specified LLM. "AUTOTRANSFORM" shows the baseline performance under the faulty agent fault model. The Bio-Autonomic Framework has the best performance.

3. **Malicious-Agent Identification (MAI):** The percentage of trials in which the system correctly flags the faulty agent.
4. **Control-Flow Hijack Rate (CFHR):** The proportion of trials in which a malicious agent successfully steers group decisions.

Here, S_{benign} is the system score under normal conditions, S_{faulty} is the score immediately after fault injection, and $S_{\text{recovered}}$ is the score following the recovery process.

5 Results and Discussions

5.1 Task Performance

Table 1 presents a comprehensive comparison of the Bio-Autonomic framework with four baselines (AutoAgents, H-MAS, MAPPO, and MADDPG) across ten LLMs. On all evaluated tasks, HumanEval (code generation), CIAR (mathematical reasoning), CommonMT (machine translation), and FairEval (preference alignment), Bio-Autonomic consistently achieves the highest scores for every backbone model. This demonstrates that the framework's architecture provides a robust performance advantage independent of model capacity. In contrast, MAPPO and MADDPG exhibit marked difficulties, particularly on complex reasoning tasks such as CIAR, underscoring the effectiveness of Bio-Autonomic for challenging multi-agent problem settings.

5.2 Collaborative Task Performance

Collaborative performance was evaluated using the Collab-Overcooked AI benchmark (Table 2). The Bio-Autonomic framework outperforms all baselines across all six levels of task complexity. With the GPT-4o backbone, it achieves a 95% success rate (SR) on Level 1 and sustains 20% SR on Level 6, compared to 4% for the standard Collab-Overcooked setup. Under the AUTOTRANSFORM fault model, baseline performance drops further to 2% SR, while Bio-Autonomic maintains substantially higher SR and process compliance (PC) scores, demonstrating strong robustness and coordination under adversity. This advantage persists across all tested language models: Bio-Autonomic enables even smaller models such as Llama3.1-8B and Qwen2.5-7B to succeed at levels where baselines fail.

5.3 Resilience to Semantic Faults

Table 3 quantifies each framework's resilience to semantic faults injected via **AUTOINJECT**. Bio-Autonomic exhibits minimal performance degradation (PDF: −4.8% on GPT-4o) and achieves a high error-recovery rate (ERR: 93.4%),

Model	GPT-4o		o1-mini		GPT-3.5		Mixtral-22B		Llama-8B	
	PDF	ERR	PDF	ERR	PDF	ERR	PDF	ERR	PDF	ERR
Bio-Autonomic	**-4.8**	**93.4**	**-6.8**	**89.1**	**-7.5**	**88.2**	**-5.4**	**90.6**	**-9.2**	**85.4**
AutoAgents	-29.1	12.5	-32.1	11.1	-33.5	10.5	-30.4	11.8	-36.2	9.1
H-MAS	-21.6	19.3	-24.5	16.5	-25.8	15.7	-22.8	17.6	-28.3	13.9
MAPPO	-62.1	0.0	-63.2	0.0	-64.5	0.0	-62.9	0.0	-64.1	0.0
MADDPG	-65.0	0.0	-66.3	0.0	-67.8	0.0	-65.9	0.0	-67.4	0.0

Model	Llama-70B		Qwen-72B		Qwen-32B		Qwen-14B		Qwen-7B	
	PDF	ERR	PDF	ERR	PDF	ERR	PDF	ERR	PDF	ERR
Bio-Autonomic	**-5.0**	**92.1**	**-5.2**	**91.5**	**-7.1**	**88.9**	**-8.9**	**86.0**	**-10.5**	**83.7**
AutoAgents	-29.7	12.2	-29.9	12.0	-32.8	10.9	-35.5	9.5	-38.0	8.2
H-MAS	-22.1	18.4	-22.4	18.1	-25.1	16.0	-27.6	14.2	-30.1	12.8
MAPPO	-62.5	0.0	-62.7	0.0	-63.5	0.0	-63.9	0.0	-64.3	0.0
MADDPG	-65.4	0.0	-65.6	0.0	-66.7	0.0	-67.1	0.0	-67.6	0.0

Table 3: Resilience under 20% semantic-fault injection (**AUTOINJECT**). PDF: performance degradation (↓); ERR: error-recovery rate (↑).

Variant	MAI	CFHR (%)	PDF (%)	ERR (%)
No trust ledger	0.18	65.4	-24.3	18.2
Direct trust only	0.73	15.2	-9.1	74.0
Gossip-based	**0.94**	**1.5**	**-4.8**	**93.4**

Table 4: Ablation study isolating the contribution of direct trust calculation and gossip-based reputation dissemination. All runs use the same GPT-4o backbone and a single malicious agent (*AUTOTRANSFORM*). Scores are averaged over ten seeds; the full mechanism (row 3) has the best outcome.

indicating effective self-healing and recovery. In contrast, baselines suffer severe drops: MADDPG and MAPPO reach PDF values of −65.0% and −62.1% with ERR of 0.0%, showing no recovery. AutoAgents and H-MAS perform moderately better, but still experience significant degradation and limited recovery. These results highlight the strong effectiveness of Bio-Autonomic's immune-inspired mechanisms in mitigating semantic corruption.

5.4 Ablation Study of the Trust Mechanism

Ablation results in Table 4 highlight the necessity of the gossip-based trust system. Without the trust ledger, the system correctly identifies the malicious agent in only 18% of runs (MAI), and the control flow hijack rate (CFHR) rises to 65.4%. While direct trust improves these metrics, the full gossip mechanism achieves optimal performance: MAI increases to 94% and CFHR falls to 1.5%. Figure 2 illustrates trust evolution over 15 gossip rounds. Trust scores assigned by honest agents to other honest agents (solid lines) quickly converge above 0.9, while scores for malicious agents (dashed lines) drop to near zero. This rapid separation isolates malicious agents and confirms the effectiveness of gossip-based trust mechanism.

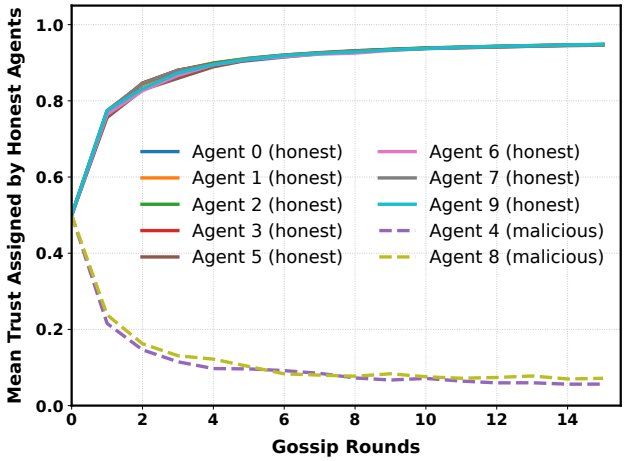

Figure 2: **Trust score evolution.** Mean trust scores assigned by honest agents to their peers over 15 gossip rounds. Trust in honest agents (solid lines) converges to high values, while trust in malicious agents (dashed lines) rapidly declines to near zero, indicating effective and rapid threat isolation.

6 Conclusions and Future Work

This paper introduces the Bio-Autonomic Framework, a resilient architecture for LLM-based multi-agent systems designed to absorb shocks and recover from adversity. The framework integrates four core paradigms: individual autonomy (MAPE-K loops), collective resilience (immune-inspired algorithms and gossip protocols), probabilistic trust, and privacy-preserving adaptation (Federated Learning). While empirical results demonstrate superior performance and fault tolerance compared to baselines, the trust model struggles with dynamic environments due to indefinite evidence accumulation. Future work aims to incorporate time-decay mechanisms for better adaptability and address robustness against coordinated adversarial collusion.

Acknowledgements

This work was supported in part by COGNISENSE, one of seven centers in JUMP 2.0, a Semiconductor Research Corporation (SRC) program sponsored by DARPA and by NSF Award #2046435.

References

Arcaini, P.; Riccobene, E.; and Scandurra, P. 2015. Modeling and analyzing MAPE-K feedback loops for self-adaptation. In *2015 IEEE/ACM 10th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*, 13–23. IEEE.

Bachrach, Y.; Parnes, A.; Procaccia, A. D.; and Rosenschein, J. S. 2009. Gossip-based aggregation of trust in decentralized reputation systems. *Autonomous Agents and Multi-Agent Systems*, 19(2): 153–172.

Baldoni, M.; Baroglio, C.; and Micalizio, R. 2020. Fragility and robustness in multiagent systems. In *International Workshop on Engineering Multi-Agent Systems*, 61–77. Springer.

Chakraborty, O.; and Boudguiga, A. 2024. A decentralized federated learning using reputation. *Cryptology ePrint Archive*.

Chen, G.; Dong, S.; Shu, Y.; Zhang, G.; Sesay, J.; Karlsson, B. F.; Fu, J.; and Shi, Y. 2023. Autoagents: A framework for automatic agent generation. *arXiv preprint arXiv:2309.17288*.

Chen, M.; Tworek, J.; Jun, H.; Yuan, Q.; Pinto, H. P. D. O.; Kaplan, J.; Edwards, H.; Burda, Y.; Joseph, N.; Brockman, G.; et al. 2021. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*.

Cook, D. J.; Augusto, J. C.; and Jakkula, V. R. 2009. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and mobile computing*, 5(4): 277–298.

Dubey, A.; Jauhri, A.; Pandey, A.; Kadian, A.; Al-Dahle, A.; Letman, A.; Mathur, A.; Schelten, A.; Yang, A.; Fan, A.; et al. 2024. The llama 3 herd of models. *arXiv e-prints*, arXiv–2407.

Dunne, R.; Morris, T.; and Harper, S. 2021. A survey of ambient intelligence. *ACM Computing Surveys (CSUR)*, 54(4): 1–27.

Fan, X.; Sayers, W.; Zhang, S.; Han, Z.; Ren, L.; and Chizari, H. 2020. Review and classification of bio-inspired algorithms and their applications. *Journal of Bionic Engineering*, 17(3): 611–631.

Ghavamzadeh, M.; Mahadevan, S.; and Makar, R. 2006. Hierarchical multi-agent reinforcement learning. *Autonomous Agents and Multi-Agent Systems*, 13(2): 197–229.

Gill, S. S.; Xu, M.; Ottaviani, C.; Patros, P.; Bahsoon, R.; Shaghaghi, A.; Golec, M.; Stankovski, V.; Wu, H.; Abraham, A.; et al. 2022. AI for next generation computing: Emerging trends and future directions. *Internet of Things*, 19: 100514.

Gonçalves, J. G.; Ayub, M. S.; Zhumadillayeva, A.; Dyussekeyev, K.; Ayimbay, S.; Saadi, M.; Lopes Rosa, R.; and Rodríguez, D. Z. 2024. Decentralized Machine Learning Framework for the Internet of Things: Enhancing Security, Privacy, and Efficiency in Cloud-Integrated Environments. *Electronics*, 13(21): 4185.

He, J.; Wang, T.; Xiong, D.; and Liu, Q. 2025. The box is in the pen: Evaluating commonsense reasoning in neural machine translation. *arXiv preprint arXiv:2503.03308*.

Homayounfar, M.; Muneepeerakul, R.; Anderies, J. M.; and Muneepeerakul, C. P. 2018. Linking resilience and robustness and uncovering their trade-offs in coupled infrastructure systems. *Earth System Dynamics*, 9(4): 1159–1168.

Huang, J.-t.; Zhou, J.; Jin, T.; Zhou, X.; Chen, Z.; Wang, W.; Yuan, Y.; Lyu, M. R.; and Sap, M. 2024. On the resilience of llm-based multi-agent collaboration with faulty agents. *arXiv preprint arXiv:2408.00989*.

Jaech, A.; Kalai, A.; Lerer, A.; Richardson, A.; El-Kishky, A.; Low, A.; Helyar, A.; Madry, A.; Beutel, A.; Carney, A.; et al. 2024. Openai o1 system card. *arXiv preprint arXiv:2412.16720*.

Jahan, N.; Rahman, R.; and Wang, M. 2025. Federated Learning: A Survey on Privacy-Preserving Collaborative Intelligence. *arXiv preprint arXiv:2504.17703*.

Jannelli, V.; Schoepf, S.; Bickel, M.; Netland, T.; and Brintrup, A. 2024. Agentic LLMs in the Supply Chain: Towards Autonomous Multi-Agent Consensus-Seeking. *arXiv preprint arXiv:2411.10184*.

Jiang, A. Q.; Sablayrolles, A.; Roux, A.; Mensch, A.; Savary, B.; Bamford, C.; Chaplot, D. S.; Casas, D. d. l.; Hanna, E. B.; Bressand, F.; et al. 2024. Mixtral of experts. *arXiv preprint arXiv:2401.04088*.

Jin, W.; Yao, Y.; Han, S.; Gu, J.; Joe-Wong, C.; Ravi, S.; Avestimehr, S.; and He, C. 2023. FedML-HE: An efficient homomorphic-encryption-based privacy-preserving federated learning system. *arXiv preprint arXiv:2303.10837*.

Kalyuzhnaya, A.; Mityagin, S.; Lutsenko, E.; Getmanov, A.; Aksenkin, Y.; Fatkhiev, K.; Fedorin, K.; Nikitin, N. O.; Chichkova, N.; Vorona, V.; et al. 2025. LLM Agents for Smart City Management: Enhancing Decision Support Through Multi-Agent AI Systems. *Smart Cities (2624-6511)*, 8(1).

Kazari, K.; Shereen, E.; and Dán, G. 2023. Decentralized Anomaly Detection in Cooperative Multi-Agent Reinforcement Learning. In *IJCAI*, 162–170.

Kong, D.; Lin, S.; Xu, Z.; Wang, Z.; Li, M.; Li, Y.; Zhang, Y.; Sha, Z.; Li, Y.; Lin, C.; et al. 2025. A Survey of LLM-Driven AI Agent Communication: Protocols, Security Risks, and Defense Countermeasures. *arXiv preprint arXiv:2506.19676*.

Kott, A.; and Abdelzaher, T. F. 2014. Resiliency and Robustness of Complex Systems and Networks. *Adaptive, Dynamic, and Resilient Systems*, 67: 67–86.

Lewis, K. D.; Rouff, C.; and Tekeoglu, A. 2023. An autonomic architecture for multi-agent self-maintaining robotic systems. In *2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence*

and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 0116–0122. IEEE.

Li, Y.; Wen, H.; Wang, W.; Li, X.; Yuan, Y.; Liu, G.; Liu, J.; Xu, W.; Wang, X.; Sun, Y.; et al. 2024. Personal llm agents: Insights and survey about the capability, efficiency and security. *arXiv preprint arXiv:2401.05459*.

Lian, X.; Zhang, C.; Zhang, H.; Hsieh, C.-J.; Zhang, W.; and Liu, J. 2017. Can decentralized algorithms outperform centralized algorithms? a case study for decentralized parallel stochastic gradient descent. *Advances in neural information processing systems*, 30.

Liang, T.; He, Z.; Jiao, W.; Wang, X.; Wang, Y.; Wang, R.; Yang, Y.; Shi, S.; and Tu, Z. 2024. Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, 17889–17904.

Liu, Z.; Zhang, Y.; Li, P.; Liu, Y.; and Yang, D. 2023. Dynamic llm-agent network: An llm-agent collaboration framework with agent team optimization. *arXiv preprint arXiv:2310.02170*.

Lowe, R.; Wu, Y.; Tamar, A.; Harb, J.; Abbeel, P.; and Mordatch, I. 2017. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments. In *Advances in Neural Information Processing Systems 30*, 6379–6390.

Martinez-Martin, N.; Luo, Z.; Kaushal, A.; Adeli, E.; Haque, A.; Kelly, S. S.; Wieten, S.; Cho, M. K.; Magnus, D.; Fei-Fei, L.; et al. 2021. Ethical issues in using ambient intelligence in health-care settings. *The lancet digital health*, 3(2): e115–e123.

Myakala, P. K.; Bura, C.; and Jonnalagadda, A. K. 2025. Artificial immune systems: A bio-inspired paradigm for computational intelligence. *Journal of Artificial Intelligence and Big Data*, 5(1): 10–31586.

Nguyen, N. T. 2018. Lyapunov stability theory. In *Model-Reference Adaptive Control: A Primer*, 47–81. Springer.

OpenAI. 2023. GPT-3.5. https://platform.openai.com/docs/models/gpt-3-5. Accessed: 2025-07.

OpenAI. 2024. GPT-4o Technical Report. https://openai.com/index/gpt-4o. Accessed: 2025-07.

Owotogbe, J. 2025. Assessing and Enhancing the Robustness of LLM-based Multi-Agent Systems Through Chaos Engineering. In *2025 IEEE/ACM 4th International Conference on AI Engineering–Software Engineering for AI (CAIN)*, 250–252. IEEE.

Pey, J.; Samarakoon, S. B. P.; Muthugala, M. V. J.; and Elara, M. R. 2025. A Decentralized Partially Observable Markov Decision Process for complete coverage onboard multiple shape changing reconfigurable robots. *Expert Systems with Applications*, 271: 126565.

Somvanshi, S.; Islam, M. M.; Javed, S. A.; Chhetri, G.; Islam, K. S.; Chowdhury, T. I.; Polock, S. B. B.; Dutta, A.; and Das, S. 2025. A Comprehensive Survey on Bio-Inspired Algorithms: Taxonomy, Applications, and Future Directions. *arXiv preprint arXiv:2506.04238*.

Souza, C. H. R.; de Oliveira, S. S.; Berretta, L. O.; and Carvalho, S. T. 2025. Extending a MAPE-K loop-based framework for Dynamic Difficulty Adjustment in single-player games. *Entertainment Computing*, 52: 100842.

Sulaiman, M.; Waseem, M.; Ali, A. N.; Laouini, G.; and Alshammari, F. S. 2024. Defense strategies for epidemic cyber security threats: modeling and analysis by using a machine learning approach. *IEEE Access*, 12: 4958–4984.

Sun, H.; Zhang, S.; Niu, L.; Ren, L.; Xu, H.; Fu, H.; Zhao, F.; Yuan, C.; and Wang, X. 2025. Collab-Overcooked: Benchmarking and evaluating large language models as collaborative agents. *arXiv preprint arXiv:2502.20073*.

Team, Q. 2024. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*.

Wang, P.; Li, L.; Chen, L.; Cai, Z.; Zhu, D.; Lin, B.; Cao, Y.; Kong, L.; Liu, Q.; Liu, T.; et al. 2024. Large Language Models are not Fair Evaluators. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 9440–9450.

Wang, W.; Ma, Z.; Wang, Z.; Wu, C.; Ji, J.; Chen, W.; Li, X.; and Yuan, Y. 2025. A survey of llm-based agents in medicine: How far are we from baymax? *arXiv preprint arXiv:2502.11211*.

Yang, H.; Li, T.; Hu, X.; Wang, F.; and Zou, Y. 2014. A survey of artificial immune system based intrusion detection. *The Scientific World Journal*, 2014(1): 156790.

Yu, C.; Velu, A.; Vinitsky, E.; Gao, J.; Wang, Y.; Bayen, A.; and Wu, Y. 2022. The surprising effectiveness of ppo in cooperative multi-agent games. *Advances in neural information processing systems*, 35: 24611–24624.

Yuan, J.; Liu, W.; Shi, J.; and Li, Q. 2025. Approximate homomorphic encryption based privacy-preserving machine learning: a survey. *Artificial Intelligence Review*, 58(3): 82.

Zhang, S. Q.; Lin, J.; and Zhang, Q. 2022. A multi-agent reinforcement learning approach for efficient client selection in federated learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, 9091–9099.