

Toward Autonomous AI Organizations: The Hierarchical Agentic AI Governance Architecture (HAAGA)

Benjamin Quito 

Conestoga College, Applied Computer Science and Information Technology, Kitchener, Canada

Email: bquito@conestogac.on.ca

How to cite this paper: Quito, B. (2026)
Toward Autonomous AI Organizations:
The Hierarchical Agentic AI Governance
Architecture (HAAGA). *Intelligent Control
and Automation*, 17, 147-174.
<https://doi.org/10.4236/ica.2026.174007>

Received: August 26, 2026

Accepted: October 11, 2026

Published: October 14, 2026

Copyright © 2026 by author(s) and
Scientific Research Publishing Inc.
This work is licensed under the Creative
Commons Attribution-NonCommercial
International License (CC BY-NC 4.0).
<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

As agentic artificial intelligence evolves from task-oriented automation to persistent organizational participation, existing multi-agent frameworks offer limited guidance on enterprise governance, institutional continuity, and accountability. This conceptual paper proposes the Hierarchical Agentic AI Governance Architecture (HAAGA), which organizes an Autonomous AI Organization across Governance, Administration, Operations, and Organizational Memory. An Autonomous AI Organization is distinguished from a persistent multi-agent system by an enduring mission and identity; durable departments, roles, and decision rights; formal escalation and final adjudication; institutional governance; controlled organizational memory; continuity across agent replacement; and auditable accountability. Decisions remain at the lowest level, with formally assigned authority and validated competence, while unresolved disagreements, insufficient authority, or requests for guidance trigger escalation from local to supervisory and executive levels. The Board of AI Governors and Executive AI are organizational roles that may be implemented by autonomous agents, human officials, or hybrid arrangements; the sponsoring legal organization and a designated human officer retain accountability for consequential decisions. HAAGA preserves departmental recommendations, evidence, reasoning, objections, and revisions in a controlled, versioned decision record that informs future cases. Preserved dissent improves reviewability but does not by itself create independent perspectives when departments share models, data, tools, or policy constraints. HAAGA therefore offers a theory-informed organizational specification whose effectiveness, scalability, and bias-mitigation properties require prototype implementation and empirical evaluation.

Keywords

Agentic AI, Autonomous AI Organizations, AI Governance,

1. Introduction

Artificial intelligence (AI) has undergone a remarkable transformation from systems designed to execute predefined computational tasks into increasingly autonomous entities capable of reasoning, planning, learning, and interacting with both humans and other intelligent systems. This transformation has been accelerated by advances in foundation models and large language models (LLMs), whose capabilities have enabled the emergence of agentic AI—autonomous systems that can perceive objectives, formulate plans, utilize external tools, collaborate with other agents, and adapt their behaviour to dynamic environments [1]–[3]. Unlike traditional AI applications that rely primarily on deterministic workflows, agentic AI systems exhibit goal-directed behaviour, allowing them to perform increasingly sophisticated cognitive and operational tasks with limited human intervention.

The increasing sophistication of these frameworks has prompted organizations to envision AI systems that function not merely as task-oriented assistants but as persistent organizational participants capable of supporting long-term operational and strategic activities. Rather than executing isolated workflows, future enterprise AI systems are expected to coordinate multiple functional responsibilities, preserve organizational knowledge, enforce institutional policies, support governance processes, and collaborate continuously across organizational boundaries. Consequently, the design challenges confronting agentic AI are evolving from questions of how autonomous agents collaborate toward broader questions concerning how autonomous AI organizations should be structured, governed, coordinated, and sustained.

This transition represents a significant challenge for both artificial intelligence and information systems research. While existing agentic AI frameworks have substantially advanced autonomous reasoning, collaboration, workflow orchestration, and task execution, comparatively less attention has been devoted to the organizational governance structures required to manage persistent AI organizations operating at enterprise scale. Organizational governance encompasses authority relationships, accountability mechanisms, policy enforcement, administrative coordination, institutional continuity, and organizational memory—capabilities long recognized as essential to effective organizational performance in the information systems and management literature [4]–[6]. As autonomous AI systems become increasingly embedded within organizational environments, these governance capabilities become equally important for ensuring reliable, accountable, and sustainable autonomous operation.

The central premise of this study is that the next evolution of agentic AI lies not

merely in developing increasingly capable autonomous agents, but in designing Autonomous AI Organizations governed through formal organizational structures. While existing research has primarily advanced the operational capabilities of autonomous agents, this study argues that the long-term success of enterprise-scale agentic AI will depend equally on organizational governance, institutional continuity, administrative coordination, and organizational memory. Consequently, the proposed framework shifts the focus of agentic AI research from orchestrating autonomous agents toward governing persistent Autonomous AI Organizations.

Despite the maturity of these organizational theories, their integration into contemporary agentic AI research remains limited. Existing multi-agent frameworks provide valuable mechanisms for communication, reasoning, planning, and task orchestration; however, they generally do not conceptualize autonomous AI systems as organizations composed of governance entities, administrative structures, operational units, institutional memory, and enduring organizational responsibilities. Consequently, current research provides comparatively limited guidance regarding how autonomous AI organizations should be governed throughout their lifecycle, how organizational authority should be distributed, how institutional knowledge should be preserved, how governance policies should be enforced, and how organizational accountability should be maintained as AI systems continue to evolve.

The present study addresses this research gap by developing the Hierarchical Agentic AI Governance Architecture (HAAGA). Rather than proposing another framework for coordinating autonomous agents, HAAGA introduces a theory-informed organizational governance architecture that conceptualizes autonomous AI as a persistent organization comprising governance entities, administrative structures, operational units, and organizational memory, all operating under established governance principles. In doing so, this study extends the focus of agentic AI research from coordinating autonomous agents to governing Autonomous AI Organizations.

Accordingly, this study asks: How can an enterprise-scale Autonomous AI Organization distribute decision authority across operational, administrative, and governance levels while preserving accountability, procedural pluralism, institutional memory, and organizational continuity?

Definition 1 (Autonomous AI Organization). An Autonomous AI Organization is a persistent organizational entity composed of autonomous AI agents occupying enduring governance, administrative, and operational roles, coordinated through formal organizational structures, institutional governance mechanisms, and shared organizational memory to achieve collective organizational objectives.

Minimum Properties of an Autonomous AI Organization

A persistent multi-agent system becomes an Autonomous AI Organization only when the following minimum set of properties operates together. Persistence

alone is insufficient because a technically long-running collection of agents may still lack an organizational identity, formal authority, institutional governance, or accountable continuity.

An enduring organizational identity and mission that continues beyond any single task or workflow.

Durable departments and organizational roles with formally assigned responsibilities and decision rights.

Bounded local, supervisory, and executive authority, including formal escalation and a defined final adjudicator.

A governance charter, institutional policies, and oversight mechanisms that persist independently of individual agent instances.

Controlled organizational memory that preserves policies, evidence, dissent, decisions, outcomes, and revisions with provenance and version history.

Continuity across agent replacement, model updates, restructuring, and other lifecycle changes.

Auditable accountability anchored to a sponsoring legal organization and a designated human accountable officer for consequential decisions.

2. Conceptual Foundations

2.1. Conceptual Framework Development Approach

This study uses a theory-driven integrative conceptual synthesis to develop HAAGA. The approach is appropriate because the research objective is to integrate constructs that are typically examined separately across agentic AI, information systems governance, organizational design, institutional continuity, coordination, and organizational memory, and translate them into a coherent enterprise architecture.

Source selection was purposive rather than systematic or exhaustive. The synthesis combines seminal works that define the selected organizational theories with contemporary surveys, peer-reviewed agentic AI frameworks, and official framework documentation needed to characterize the current technical landscape. Sources were retained when they directly informed authority, accountability, hierarchy, coordination, institutionalization, memory, persistence, or multi-agent execution.

The synthesis proceeded in four stages. First, recurrent theoretical constructs were extracted from the selected literature. Second, those constructs were used to identify four governance and organizational gaps in contemporary agentic AI. Third, the gaps were translated into theory-linked design principles. Fourth, the principles were mapped to architectural layers, decision pathways, and organizational memory requirements, and then compared with representative agentic AI frameworks.

Four gaps structure the analysis: a governance gap involving decision rights, accountability, and oversight; an organizational gap involving hierarchy, administration, and distributed decision-making; a knowledge-management gap involv-

ing institutional rather than temporary agent memory; and an institutionalization gap involving the preservation and revision of rules, roles, and organizational identity over time.

These gaps are translated into seven complementary principles: governance; organizational structure; subsidiarity and escalation; procedural pluralism and traceability; institutional continuity; coordination; and organizational memory. The principles provide the conceptual basis for the four HAAGA layers and for the formal movement of decisions and decision records between them.

The synthesis produced a theory-to-architecture traceability mapping, which is presented alongside the HAAGA design principles in Section 3.1. This mapping provides a conceptual justification for the proposed architecture; it does not constitute evidence that HAAGA has been implemented or empirically validated.

The resulting framework is conceptual and platform-independent. It establishes organizational and governance requirements rather than a software implementation. Prototype construction, simulation, comparison under controlled decision scenarios, and field validation are therefore reserved for later phases of the research program.

2.2. Evolution and Contemporary Agentic AI

Artificial intelligence has evolved from systems designed to automate narrowly defined computational tasks to increasingly autonomous and collaborative intelligent entities. This progression has been driven by advances in machine learning, knowledge representation, reasoning, planning, and, more recently, large language models (LLMs), which have substantially expanded the capabilities of intelligent systems. Rather than representing isolated technological milestones, these developments reflect a broader shift in how intelligence is conceptualized—from individual computational functions to autonomous systems capable of reasoning, collaboration, adaptation, and increasingly complex organizational behaviour. Understanding this evolution provides the conceptual foundation for positioning Autonomous AI Organizations as the next stage in the maturation of agentic artificial intelligence [7] [8].

Contemporary agentic AI frameworks demonstrate increasingly sophisticated collaboration and autonomous execution, but they remain primarily oriented toward bounded tasks and workflows. Formal authority, administrative oversight, institutional accountability, policy management, organizational memory, and long-term continuity are generally implicit or external to these systems. Enterprise-scale persistence, therefore, requires an organizational architecture that regulates authority and preserves governance independently of individual tasks or agent instances.

2.3. Organizational Governance

Organizational governance provides the formal structures, decision-making mechanisms, and accountability relationships through which organizations pur-

sue strategic objectives while ensuring effective coordination, control, and responsible use of resources. Unlike operational management, which focuses on executing organizational activities, governance establishes the policies, authority structures, oversight mechanisms, and decision rights that guide organizational behaviour over time. Within information systems research, governance has long been recognized as a critical capability for aligning organizational resources with strategic objectives while balancing autonomy, accountability, and organizational performance [5] [6].

The distinction between operational coordination and organizational governance is particularly important. Operational coordination focuses on how autonomous agents work together to accomplish specific objectives, whereas organizational governance determines how authority is distributed, responsibilities are assigned, decisions are supervised, organizational policies are enforced, and institutional continuity is preserved. Although complementary, these functions address fundamentally different organizational concerns. Coordination enables effective execution; governance ensures that execution remains aligned with organizational purpose and long-term strategic direction.

Consequently, as agentic AI systems evolve from temporary collaborative workflows toward persistent organizational entities, governance must become an architectural capability rather than an operational afterthought. Autonomous AI Organizations require governance structures that can define organizational hierarchies, allocate decision rights, enforce institutional policies, preserve organizational knowledge, and maintain administrative oversight across multiple autonomous agents operating over extended periods. These governance requirements establish the conceptual foundation for the Hierarchical Agentic AI Governance Architecture (HAAGA).

2.4. Organizational Theory

Organizations are not merely collections of individuals performing related tasks, but structured social systems designed to coordinate resources, responsibilities, and decision-making toward shared objectives. Organizational theory examines how organizations are designed, governed, and sustained, emphasizing the relationships among structure, authority, specialization, coordination, and organizational effectiveness. Rather than viewing organizational performance solely as the outcome of individual capabilities, organizational theory recognizes that effective organizations emerge from well-defined structures that enable diverse actors to function collectively as an integrated system [9].

Hierarchy represents another fundamental organizational principle. Hierarchical structures establish formal authority relationships that define reporting responsibilities, decision-making authority, accountability, and supervisory functions across multiple organizational levels. Rather than limiting autonomy, organizational hierarchies enable decentralized decision-making within clearly defined governance boundaries, allowing organizations to balance local operational flexi-

bility with centralized strategic direction. Hierarchical governance has therefore remained a defining characteristic of enduring organizations across both public and private sectors [10] [11].

Viewing autonomous agents as organizational actors rather than isolated computational entities fundamentally changes the architectural requirements of agentic AI. Autonomous AI Organizations require persistent organizational structures that define governance responsibilities, administrative authority, role specialization, reporting relationships, and decision rights, independent of individual workflows or task execution. Organizational theory, therefore, provides a foundational lens through which autonomous agents can be organized into coherent, governable organizational systems capable of sustaining long-term institutional objectives.

2.5. Institutional Theory

Institutional theory explains how organizations achieve stability, legitimacy, and long-term continuity by developing formal structures, shared norms, established practices, and governance mechanisms that persist beyond individual organizational members. Unlike organizational theory, which primarily focuses on internal structures and administrative design, institutional theory emphasizes the broader organizational environment and the processes through which organizations become enduring institutions. Organizations therefore survive not only because they perform efficiently, but because they establish governance structures and practices that are perceived as legitimate, consistent, and sustainable within their institutional contexts [12] [13].

As autonomous agents increasingly assume persistent organizational responsibilities within enterprise environments, institutional continuity becomes a critical architectural requirement. Autonomous AI Organizations must preserve governance policies, organizational structures, administrative procedures, and institutional knowledge independently of individual agent instances or temporary workflows. Organizational legitimacy likewise depends upon the consistent application of governance principles that remain stable despite changes in organizational composition, operational objectives, or technological infrastructure.

2.6. Coordination Theory

Coordination theory examines how interdependent actors organize, communicate, and synchronize their activities to achieve collective objectives. Originally developed to understand coordination within human organizations, the theory has been widely applied to information systems, distributed computing, workflow management, and collaborative technologies. At its core, coordination theory recognizes that organizational effectiveness depends not only on the capabilities of individual actors but also on the mechanisms that manage dependencies among their activities, resources, and decisions [4].

Despite these advances, coordination should not be interpreted as a substitute

for organizational governance. Coordination determines how specialized actors interact to accomplish operational objectives, whereas governance determines who possesses authority, what decisions may be made, which organizational policies apply, and how accountability is maintained across the organization. Coordination mechanisms optimize execution, while governance mechanisms establish the organizational context within which execution occurs. Consequently, coordination without governance may produce efficient workflows that nevertheless lack strategic alignment, institutional accountability, or policy compliance.

2.7. Organizational Memory

Organizations accumulate knowledge through experience, decision-making, policies, routines, and interactions that collectively shape future organizational behaviour. Organizational memory refers to the mechanisms by which organizations acquire, preserve, retrieve, and apply accumulated knowledge independently of individual members. Rather than residing solely within individuals, organizational memory is embedded in organizational repositories, formal procedures, institutional policies, historical decisions, operational routines, and shared knowledge structures, enabling organizations to maintain continuity despite personnel changes or evolving operational environments [14]-[16].

These principles have important implications for the evolution of agentic AI. Contemporary agentic AI frameworks commonly maintain temporary conversational histories, execution traces, planning states, or short-term contextual information to improve task performance. While these mechanisms enhance reasoning within individual workflows, they differ fundamentally from organizational memory. Existing frameworks generally preserve information only for the duration of a task, project, or interaction, providing limited support for long-term institutional knowledge, governance continuity, policy evolution, or organizational learning across successive organizational activities.

For Autonomous AI Organizations, organizational memory must function as a persistent institutional capability rather than as a temporary computational context. Governance policies, administrative procedures, organizational roles, historical decisions, operational experiences, compliance records, and organizational knowledge should remain available independently of individual agent instances or specific workflows. Such persistent memory enables autonomous organizations to learn from past experiences, maintain consistent governance, support institutional continuity, and preserve organizational identity as agents are replaced, updated, or reassigned over time.

2.8. Synthesis

The preceding review demonstrates that no single theoretical perspective provides a sufficient foundation for designing persistent Autonomous AI Organizations. While each theory explains an essential organizational capability, each also addresses only a specific dimension of organizational functioning. Consequently,

designing Autonomous AI Organizations requires an integrated theoretical perspective that combines organizational governance, organizational structure, institutional continuity, coordination mechanisms, and organizational memory into a unified organizational framework.

3. Hierarchical Agentic AI Governance Architecture

3.1. Design Principles

The preceding literature review established that existing agentic AI research has made significant progress in autonomous reasoning, planning, collaboration, and workflow orchestration, but provides limited guidance on governing persistent AI organizations that operate as enduring organizational entities. While contemporary agentic AI frameworks successfully coordinate autonomous agents to accomplish complex tasks, they generally emphasize computational execution rather than organizational governance. Consequently, existing approaches provide limited support for organizational authority, institutional continuity, administrative management, long-term organizational learning, and organizational accountability.

To address these limitations, this paper proposes the Hierarchical Agentic AI Governance Architecture (HAAGA), a theory-informed conceptual architecture for governing Autonomous AI Organizations. Rather than viewing autonomous agents as independent computational entities that collaborate temporarily to solve tasks, HAAGA conceptualizes them as organizational actors operating within a persistent organizational structure governed by formal authority, administrative coordination, institutional governance, and organizational memory. The objective is not merely to improve autonomous task execution but to enable AI organizations to sustain coordinated operations over extended periods while remaining aligned with organizational objectives, governance policies, and institutional constraints.

The design of HAAGA is guided by seven complementary principles derived from the theoretical perspectives synthesized in Section 2.

Governance Principle. Autonomous AI Organizations require formal governance structures that define organizational authority, decision rights, accountability mechanisms, and policy enforcement. Governance ensures that organizational decisions remain aligned with strategic objectives while providing mechanisms for oversight, compliance, and organizational control.

Organizational Structure Principle. Autonomous AI Organizations should organize autonomous agents within hierarchical organizational structures that clearly define roles, responsibilities, reporting relationships, and functional specialization. Hierarchical organization promotes scalability, administrative clarity, and effective coordination among specialized organizational units.

Subsidiarity and Escalation Principle. A decision should remain at the lowest organizational level possessing the authority and competence to resolve it. Escalation is triggered by unresolved departmental disagreement, insufficient author-

ity, or a documented request for guidance from the next level.

Procedural Pluralism and Traceability Principle. Departments may produce different recommendations from their specialized perspectives. Each recommendation, its supporting evidence, and its reasoning must remain distinct and accompany an escalated matter so that the deciding authority can evaluate disagreement without erasing minority or dissenting positions.

Institutional Continuity Principle. Organizational governance must persist independently of individual autonomous agents or specific operational workflows. Governance policies, organizational roles, institutional rules, and organizational objectives should remain stable across organizational evolution, enabling continuity despite changes in personnel, technologies, or operational requirements.

Coordination Principle. Specialized organizational units must collaborate through formal coordination mechanisms that regulate communication, resource allocation, dependency management, and collective decision-making. Effective coordination enables distributed autonomous agents to function as coherent organizational systems rather than isolated computational components.

Organizational Memory Principle. Autonomous AI Organizations require persistent organizational memory capable of preserving governance policies, administrative procedures, institutional knowledge, historical decisions, and accumulated organizational experience. Organizational memory enables continuous learning, policy refinement, and consistent governance across successive organizational activities.

Collectively, these principles embed governance within the organization rather than treating it as an external constraint. HAAGA does not claim to eliminate model, data, or human bias. Instead, its departmental separation, bounded decision rights, preserved dissent, and reviewable decision records are intended to mitigate procedural bias arising from concentrated authority, omitted perspectives, or undocumented reasoning.

Preserved dissent improves reviewability by revealing alternative recommendations and the reasons they were rejected, but it does not guarantee independent perspectives. Departments that share foundation models, training data, retrieval sources, tools, prompts, or policy constraints may reproduce correlated errors or common-mode bias. Implementations should therefore supplement procedural pluralism with independent evidence channels where feasible, model and data diversity, conflict-of-interest checks, periodic bias audits, and human review for consequential decisions.

HAAGA operationalizes these principles through four interdependent layers: Governance, Administration, Operations, and Organizational Memory. The operational, administrative, and governance layers also constitute the local, supervisory, and executive decision-making levels. Organizational Memory supports all three by preserving policies, cases, recommendations, reasons, outcomes, and subsequent revisions. The relationship between HAAGA's theoretical foundations, architectural design principles, and organizational components is summarized in **Table 1**.

Table 1. Theory-to-architecture traceability matrix.

Theory	Architectural Design Principle	HAAGA Architectural Component
Governance Theory	Decision rights, accountability, policy enforcement	Governance Layer
Organizational Theory	Hierarchical structure, specialization, administrative control	Administrative Layer
Institutional Theory	Organizational continuity, legitimacy, lifecycle management	Governance Policies and Organizational Lifecycle
Coordination Theory	Communication, dependency management, collaborative execution	Operational Layer
Organizational Memory	Knowledge preservation, organizational learning, institutional knowledge	Organizational Memory Layer
Governance and Organizational Theory	Subsidiarity, bounded decision rights, formal escalation	Local, Supervisory, and Executive Decision Authorities
Coordination and Organizational Memory	Independent recommendations, traceable reasoning, decision precedent	Departmental Review and Decision Record

Source: Author's own work.

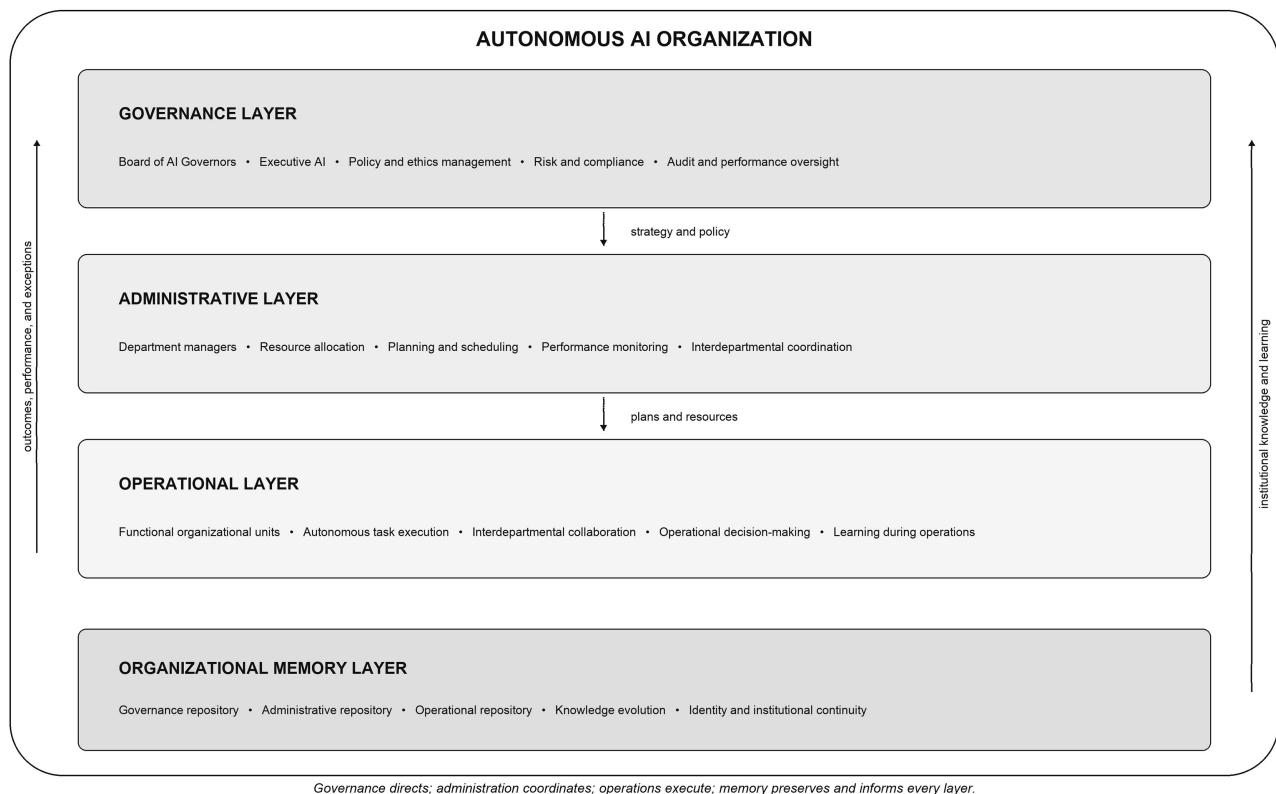
As **Table 1** shows, the foundational theories contribute distinct but complementary elements to HAAGA. Governance Theory informs organizational authority and accountability; Organizational Theory supports hierarchy and administrative specialization; Institutional Theory establishes continuity and lifecycle governance; Coordination Theory informs operational collaboration; and Organizational Memory Theory supports knowledge preservation and organizational learning. The final two rows demonstrate how these perspectives operate together to support HAAGA's cross-layer mechanisms, particularly bounded decision rights, formal escalation, independent departmental recommendations, traceable reasoning, and decision precedent.

The resulting four-layer architecture is presented in **Figure 1**.

3.2. Governance Layer

The Governance Layer represents the executive level of organizational authority within HAAGA. It establishes strategic direction, organizational objectives, decision rights, accountability mechanisms, and governance policies. It also resolves matters that cannot be decided at the operational or administrative levels while ensuring alignment with institutional objectives, ethical principles, regulatory requirements, and organizational policies.

The Governance Layer consists of five complementary governance functions that collectively establish organizational control: Board of AI Governors, Executive AI, Policy and Ethics Management, Risk and Compliance Management, and Audit and Performance Oversight. Together, these functions provide the organizational mechanisms necessary to govern persistent AI organizations while maintaining accountability, transparency, and strategic alignment.



Source: Author's own work.

Figure 1. Overall architecture of the Hierarchical Agentic AI Governance Architecture (HAAGA).

3.2.1. Governance-Role Composition and Legal Accountability

HAAGA defines the Board of AI Governors and Executive AI as organizational roles rather than prescribing a single actor type. Depending on the deployment context, either role may be performed by autonomous agents, authorized human officials, or a hybrid arrangement in which AI components analyze and recommend while human officials authorize specified classes of action. For consequential decisions, the sponsoring legal organization and a designated human accountable officer retain legal and regulatory responsibility for deployment, compliance, and resulting organizational effects. That external accountability does not render the human officer a hidden participant in the internal hierarchy: the officer may halt execution, withdraw delegated authority, or require a revision to governance, but any change to an executive adjudication must remain explicit in the decision record.

3.2.2. Board-Directed Review and Executive Reconsideration

The Board of AI Governors establishes the organization's charter, strategic vision, governance objectives, institutional policies, decision mandates, and long-term priorities. A Board-requested review is mandatory. The review directive must state its basis and any binding charter, policy, legal, ethical, risk, or resource constraints. For a consequential matter, the Board or designated human accountable officer may suspend implementation while the review is pending. The Board does

not silently rewrite the executive decision; it returns the matter to the Executive AI as a recorded review event.

The Executive AI is the single final adjudicating authority within HAAGA's internal decision hierarchy. Operating within the charter and binding directives established by the Board of AI Governors, it resolves matters that remain undecided after supervisory review. No operational department, administrative authority, or other governance function may alter an executive adjudication. In a Board-directed review, the Executive AI must reconsider the matter within the stated constraints and then affirm, modify, or overturn its prior decision through a formal, reasoned, and auditable decision. The record must preserve the original decision, the review directive, the reconsidered decision, and the reasons for any change.

The Governance Layer does not directly perform routine operational work. It establishes decision boundaries, resolves escalated matters, and issues directives that the Administrative Layer converts into coordinated plans. Operational units implement authorized decisions, while Organizational Memory preserves the complete decision record and makes prior decisions available as precedents for future cases.

3.3. Administrative Layer

The Administrative Layer serves as the organizational management function of the Hierarchical Agentic AI Governance Architecture (HAAGA). Positioned between the Governance Layer and the Operational Layer, it translates strategic governance decisions into coordinated organizational plans, allocates organizational resources, manages interdepartmental dependencies, and monitors organizational performance. Whereas the Governance Layer establishes organizational direction and policies, the Administrative Layer ensures these directives are effectively implemented across the Autonomous AI Organization.

The Administrative Layer consists of five complementary organizational management functions: Department Managers, Resource Allocation, Planning and Scheduling, Performance Monitoring, and Interdepartmental Coordination. Together, these functions transform governance objectives into coordinated organizational activities while maintaining consistency across operational units.

Department Managers represent the leadership of specialized organizational domains within the Autonomous AI Organization. Similar to functional managers in human organizations, they oversee organizational units responsible for specific business capabilities such as finance, human resources, engineering, customer service, research, or marketing.

Planning and Scheduling translates governance objectives into actionable organizational plans by defining priorities, sequencing activities, assigning responsibilities, and coordinating organizational timelines. Rather than generating isolated task plans for individual agents, administrative planning develops coordinated organizational work plans that account for dependencies across multiple organizational units.

The Administrative Layer serves as the supervisory bridge between governance and operations. Its designated Supervisory Decision Authority considers matters that operational units cannot resolve, determines whether the administrative level has sufficient authority, and either issues a supervisory decision or escalates the complete case to the Executive AI. It also coordinates implementation, communicates outcomes, and preserves planning and decision records within Organizational Memory.

3.4. Operational Layer

The Operational Layer represents the execution core of the Hierarchical Agentic AI Governance Architecture (HAAGA). It is responsible for performing the day-to-day organizational activities necessary to achieve the strategic objectives established by the Governance Layer and coordinated by the Administrative Layer. While governance defines organizational direction and administration manages organizational execution, the Operational Layer performs the specialized organizational functions that generate value for the Autonomous AI Organization.

The Operational Layer, therefore, consists of specialized organizational departments composed of autonomous agents possessing domain-specific knowledge, capabilities, and responsibilities. Each department performs a distinct organizational function while remaining accountable to the Administrative Layer and operating within the governance policies established by the Governance Layer. This organizational specialization enables scalability, modularity, and coordinated execution across complex organizational activities.

Operational activities are carried out by specialized organizational units that represent core organizational functions. Depending on the Autonomous AI Organization's mission, these units may include Finance, Human Resources, Marketing, Engineering, Customer Service, Legal Affairs, Research and Development, Procurement, or other functional departments.

Operational units execute activities independently within authorized boundaries. Each operational level has a designated Local Decision Authority responsible for resolving disagreements among local departments when possible. Operational work includes planning local activities, using organizational resources, collaborating across departments, making authorized decisions, and adapting to changing conditions.

Local authority is intentionally bounded. A matter is escalated when departments remain in disagreement, the Local Decision Authority lacks authority, or the local level seeks guidance. The complete case then moves to the Supervisory Decision Authority. Decisions affecting strategy, institutional policy, regulatory compliance, ethical governance, or organizational restructuring typically fall outside local authority and must be escalated.

Rather than allowing these experiences to disappear after task completion, HAAGA captures operational knowledge and transfers it to the Organizational Memory Layer, where it becomes available for future governance decisions, ad-

ministrative planning, organizational learning, and the preservation of institutional knowledge. This continuous knowledge transfer enables Autonomous AI Organizations to improve organizational performance over time by leveraging accumulated organizational experience rather than repeated trial and error.

The Operational Layer serves as the primary value-creation component of HAAGA while remaining tightly integrated with the other architectural layers. Governance establishes organizational objectives and decision boundaries; administration coordinates organizational activities and allocates resources; and organizational memory preserves operational knowledge for future learning.

3.5. Decision Authority, Escalation, and Traceability

Authority and Competence Mandates

Authority is the formally delegated right to decide or act within a defined mandate. Each mandate is recorded in a decision-rights registry that identifies the permitted subject domain, action types, financial and resource limits, risk class, legal and regulatory implications, affected stakeholders, and conditions requiring escalation. Competence is the validated capability to exercise that authority in the circumstances of a particular case. It is assessed against domain knowledge, access to required evidence and tools, evaluated performance, uncertainty or confidence thresholds, recency of validation, and any conflict or independence requirements. Possessing competence does not create authority, and possessing authority does not establish competence; both are required for a decision to remain at a level.

The Board of AI Governors sets and revises executive and organization-wide mandates through the governance charter, subject to the sponsoring legal organization's obligations and approval by the designated human accountable officer where law or policy requires it. Executive authority translates those mandates into supervisory delegations, and supervisory authorities define bounded local delegations. Risk and Compliance Management and Audit and Performance Oversight monitor the registry and validation evidence. Mandates are reviewed periodically and after material incidents, model or data changes, regulatory changes, sustained performance deterioration, or documented conflicts between precedent and current obligations.

HAAGA applies a lowest-competent-level rule. A matter originating in an operational department is first considered locally. If multiple local departments disagree, their separate recommendations are submitted to the designated Local Decision Authority, which either resolves the matter within its mandate or escalates it to the Administrative Layer.

At the Administrative Layer, the designated Supervisory Decision Authority evaluates the full case and may issue a decision when the matter falls within its mandate. It escalates the case when supervisory departments remain unable to resolve it, when the required authority exceeds the administrative mandate, or when documented guidance is required from the Governance Layer.

At the Governance Layer, executive departments may also produce different

recommendations. If those departments cannot reach a resolution, the Executive AI adjudicates the matter. Its decision is final within the architecture and cannot be modified by a lower level or another department. Only the Executive AI may reconsider its own decision, and any revised decision must identify the earlier decision, explain the basis for the change, and preserve both versions in Organizational Memory.

Escalation transfers the entire decision record rather than only a summary or preferred recommendation. The record contains each participating department's recommendation, supporting evidence, reasoning, unresolved objections, the authority basis for escalation, and the decision issued at each level. This paper trail allows later reviewers to reconstruct how a decision was reached and which alternatives were considered.

Authorized decisions become precedents for future cases. A prior decision informs later decisions involving materially similar facts, but it does not operate as an unchangeable rule when circumstances, evidence, policies, or legal obligations differ. A departure from precedent requires an explicit explanation and is itself added to Organizational Memory.

HAAGA may also generate or revise proposed operational code and procedures from approved architectural principles, new organizational information, and accumulated decision records. Generation does not confer authority to deploy. Each proposal must enter the same review hierarchy and receive authorization from the competent level; unresolved or out-of-scope proposals are escalated with their supporting record.

3.6. Organizational Memory Layer

The Organizational Memory Layer serves as the institutional knowledge repository of the Hierarchical Agentic AI Governance Architecture (HAAGA). Its primary responsibility is to preserve organizational knowledge, governance policies, administrative procedures, historical decisions, operational experiences, and institutional learning throughout the lifecycle of the Autonomous AI Organization. While the Governance Layer establishes organizational direction, the Administrative Layer coordinates organizational activities, and the Operational Layer executes organizational functions, the Organizational Memory Layer ensures that organizational knowledge persists beyond individual operational activities, enabling long-term organizational continuity and continuous organizational improvement.

Minimum Controls for Organizational Memory

Organizational Memory requires minimum controls so that a stored decision can function as reliable institutional evidence rather than an unverified recollection. These controls apply to governance policies, cases, recommendations, evidence, decisions, outcomes, and precedents.

Provenance: assign a unique record identifier and preserve the source, timestamp, contributing department or actor, model and tool versions where ap-

plicable, supporting evidence, and complete decision chain.

Versioning and integrity: use an append-only version history with integrity checks and an audit log, ensuring that prior states remain recoverable and that silent overwriting is prohibited.

Access permissions: apply role- or attribute-based controls, least-privilege access, confidentiality classification, and separate permissions for reading, appending, approving, correcting, or retiring records.

Retention: apply documented retention schedules, legal holds, archival requirements, and authorized disposal procedures that reflect regulatory, contractual, evidentiary, and privacy obligations.

Correction: correct an inaccurate record through a linked amendment that identifies the reason, approving authority, date, and affected versions; retain the original rather than deleting or rewriting it.

Conflicting and superseded precedents: flag conflicts, identify the governing authority and factual scope of each precedent, link superseding decisions, and prevent a superseded record from being treated as controlling without an explicit explanation.

Unlike conventional agentic AI frameworks that primarily maintain temporary conversational histories, execution traces, or workflow context, the Organizational Memory Layer maintains institutional knowledge that continuously evolves throughout the organization's existence. Organizational memory, therefore, represents a strategic organizational asset rather than a temporary computational mechanism, enabling Autonomous AI Organizations to improve governance quality, administrative effectiveness, and operational performance over time.

The Governance Knowledge Repository preserves strategic objectives, institutional policies, governance rules, ethical guidelines, regulatory requirements, and every finalized decision record. For escalated cases, it retains each department's recommendation, evidence, reasoning, objections, the authority exercised at each level, and any later reconsideration. This enables consistent use of precedent without concealing disagreement or revision.

Operational activities continuously generate organizational knowledge through task execution, collaborative problem-solving, resource utilization, customer interactions, and organizational outcomes. The Operational Knowledge Repository captures these experiences to support organizational learning and future decision-making.

Organizational memory extends beyond passive storage by supporting continuous organizational learning. Prior decisions and their outcomes inform later cases, policy refinement, administrative practice, and proposed procedural changes. When a later decision departs from a recorded precedent, the changed circumstances and reasons for departure become part of the new institutional record.

Persistent organizational memory contributes directly to organizational identity. The accumulated governance decisions, institutional policies, historical experiences, organizational objectives, and operational practices collectively define

how the Autonomous AI Organization behaves across time.

This continuity distinguishes persistent organizations from temporary multi-agent collaborations. Even as autonomous agents are upgraded, replaced, or reorganized, the organization's accumulated knowledge, governance philosophy, institutional rules, and operational experience remain intact. Organizational identity, therefore, resides within institutional knowledge rather than within any individual autonomous agent.

Accordingly, organizational memory functions as the knowledge infrastructure that integrates governance, administration, and operations into a persistent organizational system. Rather than serving as a passive repository, it enables Autonomous AI Organizations to preserve institutional continuity, learn from experience, and adapt organizational behaviour while maintaining alignment with strategic objectives and governance principles.

3.7. Organizational Lifecycle

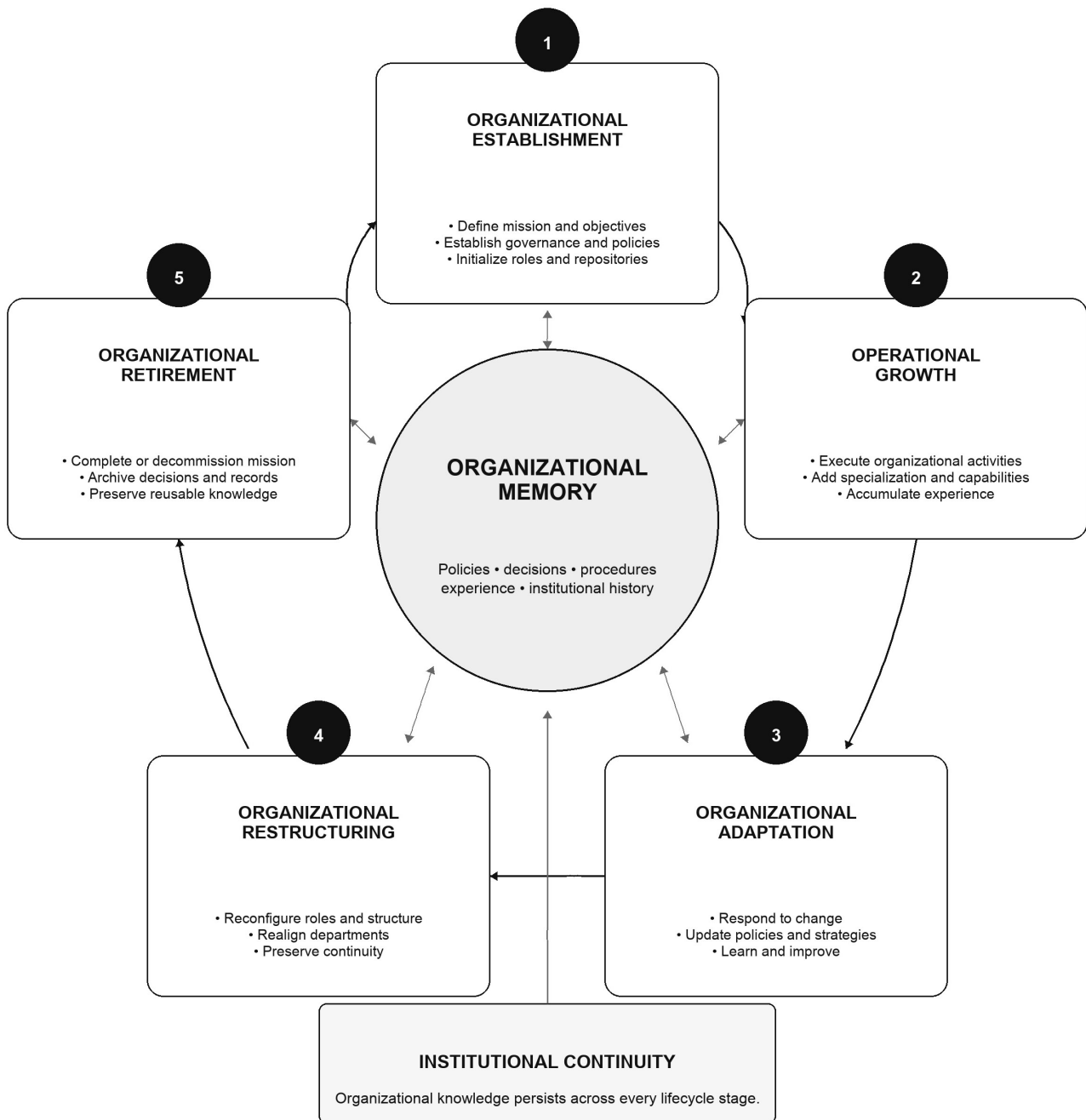
While the four architectural layers define HAAGA's organizational structure, Autonomous AI Organizations are not static entities. Like human organizations, they evolve over time in response to changing objectives, environmental conditions, technological advancements, governance decisions, and accumulated experience. Consequently, HAAGA conceptualizes Autonomous AI Organizations as dynamic organizational systems that continuously adapt while preserving institutional continuity through governance and organizational memory.

HAAGA conceptualizes the organizational lifecycle as consisting of five inter-related stages: organizational establishment, operational growth, organizational adaptation, organizational restructuring, and organizational retirement. Together, these stages describe the continuous evolution of Autonomous AI Organizations while preserving governance consistency and institutional knowledge. **Figure 2** presents these five stages and their relationship to continuous organizational evolution.

The lifecycle begins with the creation of an Autonomous AI Organization. During this stage, governance structures are established, organizational objectives are defined, administrative departments are created, governance policies are approved, and organizational memory repositories are initialized.

Organizations must continuously respond to internal and external changes, including new business objectives, regulatory requirements, technological innovations, environmental uncertainties, and stakeholder expectations. Organizational adaptation enables Autonomous AI Organizations to modify governance policies, administrative procedures, operational strategies, and organizational structures without compromising institutional continuity.

Over time, organizational priorities may require significant structural changes. Departments may be merged or divided, governance responsibilities reassigned, organizational hierarchies evolve, and autonomous agents replaced or upgraded.



Source: Author's own work.

Figure 2. Organizational lifecycle of an autonomous AI organization.

Before retirement, governance decisions, institutional knowledge, operational experiences, compliance records, and organizational history are preserved to ensure accountability, regulatory compliance, knowledge reuse, and organizational traceability. Organizational memory thus enables historical organizational knowledge to remain available for future organizations, audits, or governance review even after operational activities have ceased.

Throughout every lifecycle stage, governance remains responsible for maintain-

ing organizational accountability; administrative coordination ensures organizational stability; operational units perform organizational functions; and organizational memory preserves institutional continuity. Rather than representing discrete phases with rigid boundaries, these lifecycle stages often overlap as organizations continuously evolve in response to changing internal and external conditions.

4. Discussion

The Hierarchical Agentic AI Governance Architecture (HAAGA) extends contemporary research on agentic artificial intelligence by shifting the focus from autonomous agents performing individual tasks to Autonomous AI Organizations operating as persistent organizational entities. While recent agentic AI frameworks have significantly advanced autonomous reasoning, planning, collaboration, and tool utilization, they primarily emphasize the capabilities of individual agents or collections of collaborating agents. HAAGA argues that organizational intelligence requires capabilities that extend beyond autonomous execution to include governance, administrative coordination, institutional continuity, and organizational memory.

Grounded in Governance Theory, Organizational Theory, Institutional Theory, Coordination Theory, and Organizational Memory Theory, HAAGA conceptualizes Autonomous AI Organizations as enduring organizational systems capable of strategic governance, coordinated management, operational execution, and continuous organizational learning. This organizational perspective expands the conceptual boundary of agentic AI by introducing governance structures and institutional mechanisms traditionally associated with human organizations into autonomous artificial intelligence systems.

4.1. Theoretical Contributions

A primary theoretical contribution of HAAGA is the shift from agent-centric intelligence to organization-centric intelligence. Contemporary agentic AI research has largely concentrated on enhancing the reasoning, planning, memory, and collaborative capabilities of autonomous agents. Although these advances have significantly improved autonomous task execution, they generally treat organizations as collections of collaborating agents rather than as enduring organizational entities possessing governance structures, institutional rules, and organizational identity.

HAAGA proposes that organizational intelligence emerges from structured governance relationships rather than solely from improvements in individual agent capabilities. Consequently, the Autonomous AI Organization becomes the primary unit of analysis, extending agentic AI research beyond autonomous task execution toward persistent organizational systems capable of long-term institutional continuity.

A second contribution is the formal conceptualization of the Autonomous AI

Organization (AAO). Existing agentic AI frameworks typically instantiate groups of autonomous agents to accomplish specific objectives and then terminate upon task completion. In contrast, HAAGA introduces AAOs as persistent organizational entities composed of autonomous agents who occupy enduring governance, administrative, and operational roles and share institutional knowledge through organizational memory.

HAAGA also contributes by integrating multiple organizational theories into a unified architectural framework. Governance Theory provides mechanisms for strategic oversight and accountability. Organizational Theory explains hierarchical structures and roles within organizations. Institutional Theory supports organizational legitimacy and policy continuity. Coordination Theory explains interdepartmental collaboration and organizational synchronization, while Organizational Memory Theory provides mechanisms for preserving institutional knowledge across time.

Another theoretical contribution lies in positioning governance as a foundational capability of autonomous AI organizations rather than as an external control mechanism. Existing AI governance research frequently treats governance as a regulatory process imposed upon AI systems by human organizations. HAAGA instead embeds governance directly within the organizational architecture through dedicated governance structures responsible for strategic planning, policy development, risk management, compliance monitoring, and organizational oversight.

HAAGA further distinguishes organizational memory from the computational memory mechanisms commonly employed in contemporary agentic AI frameworks. Existing systems generally maintain conversational context, execution traces, or temporary workflow histories to support ongoing task execution. While valuable for short-term reasoning, these memory mechanisms do not preserve institutional knowledge across successive organizational activities.

The Organizational Memory Layer introduced in HAAGA serves as a persistent organizational repository for governance decisions, administrative knowledge, operational experiences, organizational policies, and institutional history. By conceptualizing memory as an organizational rather than computational capability, HAAGA enables continuous organizational learning, institutional continuity, and long-term organizational evolution.

Finally, HAAGA introduces an explicit organizational lifecycle that conceptualizes Autonomous AI Organizations as evolving organizational systems rather than static software architectures. The lifecycle encompasses organizational establishment, operational growth, organizational adaptation, organizational restructuring, and organizational retirement while preserving institutional continuity through organizational memory.

Collectively, these contributions position HAAGA as an organizational governance framework that extends contemporary agentic AI beyond autonomous task execution toward persistent, governed, and continuously evolving organizational

systems. By integrating organizational governance, administrative coordination, institutional continuity, and organizational memory into a unified architecture, HAAGA establishes a theoretical foundation for future research on Autonomous AI Organizations and offers a new perspective on how autonomous intelligence can be organized, governed, and sustained over time.

4.2. Comparison with Existing Agentic AI Frameworks

Recent advances in agentic artificial intelligence have produced numerous frameworks that enable autonomous reasoning, planning, collaboration, and tool utilization among multiple AI agents. Representative frameworks such as AutoGen, CrewAI, LangGraph, CAMEL, and Microsoft Magentic-One have significantly advanced the state of the art in autonomous multi-agent systems by improving agent communication, workflow orchestration, collaborative reasoning, and task execution [17]-[21]. These frameworks demonstrate that groups of autonomous agents can effectively cooperate to accomplish complex computational tasks.

Despite these advances, existing agentic AI frameworks primarily emphasize task-oriented collaboration rather than organizational governance. Their architectural focus remains centred on coordinating autonomous agents during the execution of individual workflows, with relatively limited attention given to governance structures, organizational management, institutional continuity, or long-term organizational evolution. Consequently, these frameworks typically instantiate temporary collections of collaborating agents that dissolve once the assigned objective has been achieved.

In contrast, HAAGA adopts an organization-centric perspective, conceptualizing Autonomous AI Organizations as persistent organizational entities governed by formal organizational structures. Rather than focusing exclusively on agent collaboration, HAAGA introduces dedicated governance, administrative, operational, and organizational-memory layers that collectively support strategic decision-making, institutional continuity, organizational learning, and long-term adaptation.

LATTICE is a governance-first architecture for deterministic authorization of individual agent actions through policy-as-code, gated execution, confidence-based escalation, and cryptographic audit trails [22]. HAAGA neither incorporates nor depends on LATTICE. LATTICE's primary concern is runtime authorization of agent actions, whereas HAAGA's unit of analysis is the enterprise organization, including departments, bounded decision rights, cross-level escalation, administrative coordination, institutional memory, and organizational continuity. LATTICE is therefore treated as an external comparator rather than as a component or enforcement mechanism within HAAGA.

Table 2 compares the organizational capabilities specified by HAAGA with the documented capabilities of representative task-oriented agentic AI frameworks.

Table 2. Comparison of representative task-oriented agentic AI frameworks and the HAAGA conceptual specification.

Architectural Capability	AutoGen	CrewAI	LangGraph	CAMEL	Magentic-One	HAAGA (conceptual)
Multi-agent collaboration	Yes	Yes	Yes	Yes	Yes	Specified
Autonomous reasoning	Yes	Yes	Yes	Yes	Yes	Specified
Tool integration	Yes	Yes	Yes	Yes	Yes	Specified
Workflow orchestration	Yes	Yes	Yes	Yes	Yes	Specified
Hierarchical governance	No	Limited	Limited	No	Limited	Specified
Administrative management	No	No	No	No	No	Specified
Organizational departments	No	No	No	No	No	Specified
Organizational memory	Limited	Limited	Limited	Limited	Limited	Specified
Institutional continuity	No	No	No	No	No	Specified
Organizational lifecycle	No	No	No	No	No	Specified
Persistent AI organization	No	No	No	No	No	Specified

Note. The comparison is bound to the cited descriptions of AutoGen [21], CrewAI [17], LangGraph [19], CAMEL [20], and Magentic-One [18]. Yes denotes a documented platform capability; Limited denotes partial, configurable, or task-scoped support; No denotes that the capability is not specified as native to the cited platform. Specified denotes a requirement in HAAGA's conceptual architecture, not a claim of technical implementation or empirical validation. LATTICE is discussed separately because it is an authorization architecture rather than a task-oriented agent framework. Source: Author's own work.

Table 2 distinguishes the currently documented capabilities of representative task-oriented frameworks from the capabilities specified by HAAGA as conceptual requirements. The table does not establish technical superiority or empirical performance. Instead, it shows that HAAGA addresses a different architectural level: the persistence, governance, coordination, and institutional continuity of an Autonomous AI Organization.

The comparison demonstrates that HAAGA should not be interpreted as a replacement for existing agentic AI frameworks. Rather, it complements these technologies by providing an organizational governance architecture that operates above existing orchestration and collaboration mechanisms. Frameworks such as AutoGen, CrewAI, LangGraph, CAMEL, and Magentic-One may therefore be viewed as implementation platforms supporting operational agent collaboration

within the broader organizational governance architecture proposed by HAAGA.

Accordingly, HAAGA expands the conceptual scope of agentic AI by integrating organizational governance with autonomous multi-agent systems. Rather than replacing advances in autonomous reasoning or collaboration, the proposed framework provides the organizational context within which these capabilities can operate as components of enduring Autonomous AI Organizations.

4.3. Practical Implications

The proposed Hierarchical Agentic AI Governance Architecture has implications for organizations seeking to deploy autonomous artificial intelligence beyond isolated task automation toward persistent organizational capabilities. By integrating governance, administrative coordination, operational execution, and organizational memory within a unified architecture, HAAGA provides a conceptual foundation for designing AI organizations capable of sustained decision-making, institutional continuity, and organizational adaptation.

Unlike existing agentic AI frameworks that primarily support autonomous workflow execution, HAAGA addresses organizational capabilities required for long-term enterprise operations, including strategic governance, organizational accountability, policy management, resource coordination, institutional learning, and organizational evolution. These capabilities have potential applications across multiple sectors in which AI systems are expected to operate continuously while remaining aligned with organizational objectives, regulatory requirements, and ethical principles.

HAAGA's dedicated governance and organizational memory layers provide mechanisms to document governance decisions, preserve audit trails, maintain institutional knowledge, and support regulatory accountability. These capabilities may facilitate the responsible deployment of autonomous organizational systems in highly regulated environments.

Rather than replacing human organizations, Autonomous AI Organizations may increasingly operate alongside human decision-makers within hybrid organizational environments. Human executives may establish strategic objectives while autonomous organizational units perform administrative coordination and operational execution under established governance policies.

The practical significance of HAAGA lies not in replacing existing agentic AI frameworks but in providing an organizational governance architecture that supports their deployment in persistent organizational environments. By integrating governance, administration, operations, and organizational memory, HAAGA offers a conceptual foundation for organizations seeking to transition from isolated autonomous workflows toward governed Autonomous AI Organizations capable of sustained organizational operation, institutional continuity, and continuous organizational learning.

4.4. Limitations and Future Research

First, HAAGA is a conceptual architecture rather than a fully implemented tech-

nical framework. It has not been validated through prototype development, simulation, or operational deployment. Its escalation, final-authority, auditability, precedent, and procedural-bias-mitigation mechanisms must therefore be treated as design propositions whose effectiveness remains to be tested.

Second, the framework emphasizes organizational governance rather than low-level implementation details. Existing agentic AI platforms differ substantially in their communication protocols, orchestration mechanisms, memory architectures, and execution models. HAAGA intentionally abstracts away from these technical implementation differences to provide a platform-independent organizational governance architecture. Future research should therefore investigate how HAAGA can be instantiated using existing agentic AI frameworks and evaluate the proposed architecture's compatibility with different implementation environments.

Third, the minimum Organizational Memory controls specify governance requirements rather than a complete technical implementation. Future work must still determine appropriate knowledge representations, storage architectures, retrieval and ranking methods, cryptographic or other integrity mechanisms, privacy-preserving access controls, retention automation, conflict detection, and empirical validation of precedent use at enterprise scale.

Fourth, HAAGA's preservation of departmental dissent improves traceability and reviewability, but does not demonstrate that departments provide independent perspectives. Shared models, data, tools, or policy constraints can create correlated errors and common-mode bias. Future evaluations should therefore measure viewpoint independence, source diversity, correlated failure, and the incremental effect of human review and external evidence on consequential decisions.

Fifth, HAAGA assumes that autonomous organizational entities can operate under governance structures analogous to those found in human organizations. While this assumption is grounded in established organizational and governance theories, its implications for organizational performance, decision quality, accountability, trust, and regulatory compliance remain largely unexplored within the emerging field of agentic AI. Empirical investigation will therefore be essential to determine whether the proposed organizational governance mechanisms produce measurable improvements in organizational effectiveness.

One important direction is to implement and empirically evaluate HAAGA. Future studies should develop prototype Autonomous AI Organizations using existing agentic AI frameworks and evaluate the effectiveness of the proposed governance architecture in realistic organizational environments. Such investigations could examine organizational performance, governance effectiveness, decision quality, coordination efficiency, and organizational resilience under different operational conditions.

A second research direction concerns the measurement of organizational governance within Autonomous AI Organizations. While HAAGA introduces governance as a foundational architectural capability, future work should develop

quantitative metrics for evaluating governance quality, accountability, transparency, policy compliance, organizational adaptability, and institutional continuity. Establishing standardized governance metrics would facilitate comparative evaluation across different organizational architectures and implementation platforms.

Another promising research direction involves human-AI hybrid organizations, in which autonomous organizational units operate alongside human managers and employees. Investigating governance structures, delegation of authority, accountability relationships, trust mechanisms, and collaborative decision-making between humans and Autonomous AI Organizations may provide valuable insights into the future of organizational design and digital transformation.

Finally, future research should examine the evolution of Autonomous AI Organizations over extended periods. Longitudinal studies could investigate how governance structures adapt, how organizational memory accumulates institutional knowledge, how organizational identities evolve, and how persistent autonomous organizations respond to changing strategic objectives, technological innovations, and external environmental conditions. These investigations would provide empirical evidence regarding the organizational lifecycle proposed within HAAGA.

5. Conclusions

This paper addressed the identified governance gap by proposing the Hierarchical Agentic AI Governance Architecture (HAAGA), a theory-informed conceptual framework for Autonomous AI Organizations (AAOs). HAAGA integrates Governance Theory, Organizational Theory, Institutional Theory, Coordination Theory, and Organizational Memory Theory through four interdependent layers: Governance, Administration, Operations, and Organizational Memory. Across these layers, decisions remain at the lowest competent level, unresolved matters escalate from local to supervisory and executive authority, and the full record of departmental recommendations and reasoning is preserved for accountability and future learning.

The theoretical contribution lies in repositioning the organization, rather than the autonomous agent, as the primary unit of analysis. HAAGA specifies bounded decision rights, a single final executive authority, formal reconsideration, institutional precedent, and organizational memory as architectural requirements for persistent autonomous organizations. Its procedurally plural design is intended to mitigate decision bias arising from concentrated authority or omitted perspectives, but implementation and empirical evaluation are required before claims of effectiveness can be made.

As agentic artificial intelligence continues to evolve, future progress will depend not only on improving the intelligence of autonomous agents but also on understanding how autonomous intelligence can be organized, governed, and sustained over time. By introducing the concept of Autonomous AI Organizations and proposing the Hierarchical Agentic AI Governance Architecture, this paper contrib-

utes to the emerging discourse on organizational governance in agentic AI and provides a foundation for future research at the intersection of artificial intelligence, organizational theory, and information systems.

Acknowledgements

I would like to thank my family for supporting me in writing this paper. To Brother Eduardo V. Manalo for his spiritual guidance. And above all else, to our Almighty God, who made this journey possible and attainable.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Bommasani, R., *et al.* (2021) On the Opportunities and Risks of Foundation Models. arXiv: 2108.07258. <https://arxiv.org/abs/2108.07258>
- [2] Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., *et al.* (2024) A Survey on Large Language Model Based Autonomous Agents. *Frontiers of Computer Science*, **18**, Article No. 186345. <https://doi.org/10.1007/s11704-024-40231-1>
- [3] Zhang, M., Yang, Y., Xie, R., Dhingra, B., Zhou, S. and Pei, J. (2026) Generalizability of Large Language Model-Based Agents: A Comprehensive Survey. *ACM Computing Surveys*, **58**, 1-44. <https://doi.org/10.1145/3794858>
- [4] Malone, T.W. and Crowston, K. (1994) The Interdisciplinary Study of Coordination. *ACM Computing Surveys*, **26**, 87-119. <https://doi.org/10.1145/174666.174668>
- [5] Weill, P. and Ross, J.W. (2004) IT Governance: How Top Performers Manage IT Decision Rights for Superior Results. Harvard Business School Press. <https://store.hbr.org/product/it-governance-how-top-performers-manage-it-decision-rights-for-superior-results/2535>
- [6] Brown, A.E. and Grant, G.G. (2005) Framing the Frameworks: A Review of IT Governance Research. *Communications of the Association for Information Systems*, **15**, 696-712. <https://doi.org/10.17705/1cais.01538>
- [7] Russell, S. and Norvig, P. (2021) Artificial Intelligence: A Modern Approach. 4th Edition, Pearson. <https://aima.cs.berkeley.edu/>
- [8] Wooldridge, M. (2009) An Introduction to Multi Agent Systems. 2nd Edition, Wiley. <https://www.wiley.com/en-us/An+Introduction+to+MultiAgent+Systems%2C+2nd+Edition-p-9780470519462>
- [9] Scott, W.R. and Davis, G.F. (2015) Organizations and Organizing: Rational, Natural, and Open Systems Perspectives. 2nd Edition, Routledge.
- [10] Mintzberg, H. (1979) The Structuring of Organizations. Prentice-Hall.
- [11] Fayol, H. (1949) General and Industrial Management (Storrs, C. Trans). Pitman Publishing. (Original Work Published 1916).
- [12] Meyer, J.W. and Rowan, B. (1977) Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, **83**, 340-363. <https://doi.org/10.1086/226550>
- [13] DiMaggio, P.J. and Powell, W.W. (1983) The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological*

- Review*, **48**, 147-160. <https://doi.org/10.2307/2095101>
- [14] Walsh, J.P. and Ungson, G.R. (1991) Organizational Memory. *The Academy of Management Review*, **16**, 57-91. <https://doi.org/10.2307/258607>
 - [15] Nonaka, I. (1994) A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, **5**, 14-37. <https://doi.org/10.1287/orsc.5.1.14>
 - [16] Nonaka, I. and Takeuchi, H. (1995) *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press.
 - [17] CrewAI (2026) CrewAI Documentation. <https://docs.crewai.com/>
 - [18] Fourney, A., *et al.* (2024) Magentic-One: A Generalist Multi-Agent System for Solving Complex Tasks. arXiv: 2411.04468.
 - [19] LangChain (2026) LangGraph Overview. <https://docs.langchain.com/oss/python/langgraph/overview>
 - [20] Li, G., *et al.* (2023) CAMEL: Communicative Agents for ‘Mind’ Exploration of Large Language Model Society. https://papers.nips.cc/paper_files/paper/2023/hash/a3621ee907def47c1b952ade25c67698-Abstract-Conference.html
 - [21] Wu, Q., *et al.* (2024) AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversations. *Proceedings of the First Conference on Language Modeling (COLM 2024)*, Philadelphia, 7-9 October 2024, 1-46. <https://openreview.net/pdf/db83d7ffc58f8ea86ddd2bcc9314d43c2bdaa911.pdf>
 - [22] Calboreanu, E. (2026) LATTICE: A Governance-First Architecture for Authorized Autonomous AI Operations. *Frontiers in Artificial Intelligence*, **9**, Article 1800407. <https://doi.org/10.3389/frai.2026.1800407>