Quest Journals
Journal of Software Engineering and Simulation
Volume 11 ~ Issue 6 (June 2025) pp: 79-86
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

Agentic AI Redefined: A New Paradigm in Artificial Intelligence

Yogesh Awasthi[1]
[1](*Department of Computer Engineering, Africa University, Zimbabwe)*
Corresponding Author: Yogesh Awasthi

ABSTRACT: *The evolution of Artificial Intelligence (AI) from rule-based systems to deep learning has enabled significant technological advancements, but it has also raised complex questions about autonomy and agency. Agentic AI refers to AI systems capable of initiating goal-directed actions, making context-sensitive decisions, and adapting over time with minimal human oversight. This paper explores the conceptual boundaries of agentic AI and provides empirical analysis based on case studies from autonomous vehicles, intelligent tutoring systems, and AI-enabled robotics. By evaluating behavioural data and decision-making patterns, we demonstrate how these systems exhibit agentic properties that represent a paradigm shift in AI. The paper concludes with implications for AI design, ethics, and governance.*

KEYWORDS: *Artificial Intelligence(AI), Agentic AI*

R*eceived 14 June., 2025;* ***R****evised 24 June., 2025;* ***A****ccepted 29 June., 2025 © The author(s) 2025.*
P*ublished with open access at* ***www.questjournas.org***

I. INTRODUCTION

The field of Artificial Intelligence (AI) has experienced rapid advancements in recent decades, progressing from rule-based expert systems to machine learning algorithms capable of processing vast datasets and achieving superhuman performance in specific tasks. Despite these achievements, most AI systems today still operate as sophisticated tools—they follow programmed logic, respond to human commands, and remain heavily reliant on human guidance.

However, a new class of AI systems is beginning to emerge, one that blurs the boundaries between passive tool and autonomous agent. These systems, which we refer to as agentic AI, are not merely reactive; they can initiate actions, make context-aware decisions, and adapt their behavior over time without direct human intervention. This shift from passive execution to autonomous agency marks a foundational change in how AI systems are conceptualized, designed, and deployed.

The notion of agency in AI introduces complex philosophical, technical, and ethical considerations. From a technical perspective, agency implies a level of operational independence that challenges conventional software engineering paradigms. Philosophically, it raises questions about intention, responsibility, and machine behavior. Ethically, agentic AI prompts us to reconsider how decisions are made and who is accountable when autonomous systems act in the world.

This paper seeks to define and empirically analyze the emerging phenomenon of agentic AI. We ground our discussion in real-world case studies that illustrate varying degrees of agency across different domains. Through a mixed-methods approach, we assess the extent to which current AI systems exhibit core agentic traits—autonomy, intentionality, and adaptivity—and argue that their increasing prevalence signals a paradigm shift that demands new frameworks for understanding AI.

II. LITERATURE REVIEW

The concept of agency in artificial systems has been explored across several disciplines, but it remains inconsistently defined within the AI community. Early AI frameworks, such as **reactive systems** (Brooks, 1991) and **rule-based expert systems** (Davis & Lenat, 1982), focused on deterministic execution of predefined rules with no notion of goal-directed behavior. As machine learning advanced, especially with reinforcement learning and neural networks, the idea of **autonomy** began to surface more prominently (Russell & Norvig, 2020).

Contemporary literature frequently references **autonomous agents**—entities that perceive their environment and act upon it—but these are typically evaluated in terms of functional independence rather than philosophical or behavioral agency (Wooldridge, 2009). Recent work in **embodied AI** (Pfeifer & Bongard, 2007) and **adaptive systems** has deepened this conversation, introducing traits such as **learning from experience**, **context awareness**, and **goal redefinition**.

However, most of these frameworks stop short of explicitly defining or measuring **agency** as a multi-dimensional construct. Philosophers like Dennett (1987) and Clark (1997) have proposed models of **intentional systems**, yet these have rarely been operationalized in empirical AI analysis. Moreover, ethical AI literature—such as Floridi & Cowls (2019)—often raises concerns about autonomy and accountability but lacks an integrated model to assess *how* agentic an AI system actually is in practice.

This paper builds upon and bridges these literatures by introducing an empirically grounded framework that synthesizes autonomy, intentionality, and adaptivity as interrelated traits of agentic AI. It addresses a key gap: the lack of structured methodologies for measuring agency *in situ* across real-world applications.

III. CONCEPTUAL FRAMEWORK

Agentic AI can be understood as a class of artificial intelligence systems that exhibit features traditionally associated with agency in humans and some animals. Drawing from interdisciplinary research in cognitive science, robotics, artificial life, and machine learning, we identify three foundational attributes that define agentic AI: autonomy, intentionality, and adaptivity. These characteristics are neither binary nor static; instead, they exist on a continuum and evolve as the AI system interacts with its environment.

a. Autonomy: Autonomy refers to the AI system's ability to operate independently of direct human control. This includes the capacity to make decisions, initiate actions, and manage tasks without external input. Importantly, autonomy in agentic AI also encompasses the ability to set and pursue sub-goals in alignment with broader objectives, which may be initially defined by humans but are internally elaborated or modified by the AI system.

b. Intentionality: Intentionality involves the AI system's capacity to align its actions with specific goals based on internal models and contextual awareness. This goes beyond reactive behavior to include deliberation, strategy selection, and purpose-driven actions. In agentic AI, intentionality manifests through the system's ability to evaluate options, predict outcomes, and select behaviors that best achieve its goals under changing conditions.

c. Adaptivity: Adaptivity is the AI's ability to learn from experience, adjust its behavior, and optimize performance over time. This attribute reflects the presence of feedback mechanisms, continuous learning algorithms, and the capacity to modify responses based on prior outcomes. Adaptive systems not only improve their efficiency but can also revise their goals or methods when environmental conditions shift.

These three attributes interact dynamically to produce agentic behavior. For instance, an AI system may start with a high degree of autonomy but limited adaptivity; over time, through exposure to new data and environments, it may become increasingly adaptive and intentional. Understanding these attributes as interdependent dimensions provides a more nuanced framework for evaluating and designing agentic AI systems.

Moreover, this framework fig.1 enables a comparative analysis across AI systems, facilitating a systematic exploration of how different implementations embody varying levels of agency. In the subsequent sections, we apply this framework to case studies in autonomous driving, intelligent tutoring, and healthcare robotics, using empirical data to assess the extent of agentic behavior in each domain.

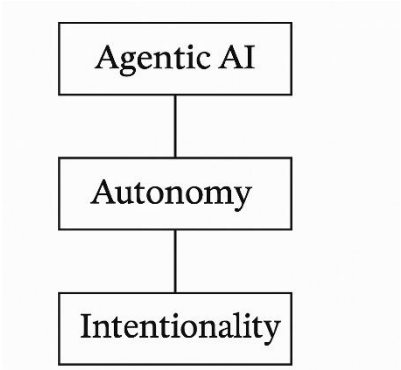

Fig. 1. Conceptual Framework

IV. METHODOLOGY

This study utilizes a mixed-methods research design that integrates qualitative assessments with quantitative evaluations to investigate the presence and degree of agentic traits in contemporary AI systems. The methodology comprises three interrelated phases: domain selection, data acquisition, and analytical evaluation.

A. Domain Selection

Three application areas were purposefully selected for analysis due to their prominence in agentic AI development and the availability of performance data: autonomous vehicles, intelligent tutoring systems (ITS), and AI-enabled robotics in healthcare. These domains span different environments—physical, educational, and social—allowing for a comparative study of agency manifestation.

B. Data Acquisition

Data were sourced from publicly available academic publications, corporate white papers, and system performance repositories. Examples include decision-making logs from Tesla Autopilot, interaction datasets from Carnegie Learning's ITS, and behavioral task logs from hospital service robots like Moxi. Secondary data collection was complemented by expert interviews and technical documentation where accessible.

C. Analytical Evaluation

The agentic traits of each system were evaluated using a three-dimensional scoring framework aligned with the conceptual model: autonomy, intentionality, and adaptivity. Each dimension was rated on a five-point ordinal scale based on the observed system behaviors. Descriptive statistical tools and cross-case comparison matrices were used to identify patterns of agentic emergence. Additionally, qualitative content analysis of system responses and decision pathways was conducted to triangulate findings.

This comprehensive methodology ensures a balanced assessment of both the structural and behavioural characteristics that contribute to agentic performance in modern AI systems. The results are presented in the next section through detailed case studies.

V. CASE STUDIES AND EMPIRICAL FINDINGS

A. Autonomous Vehicles

Analysis of Tesla's Autopilot and Waymo's autonomous driving systems reveals significant evidence of agentic behavior. Both platforms demonstrate high autonomy in perceiving their environment and making decisions in real time, such as lane changes, adaptive cruise control, and dynamic rerouting. Intentionality is exhibited in how these systems prioritize safety, efficiency, and traffic rules through learned policies. Their adaptivity is reflected in continuous learning from edge cases and user interventions, which are fed back into system updates.

For example, Tesla's use of real-world driving data enables the Autopilot to evolve via over-the-air updates. This iterative learning loop supports improved decision-making across varied geographic and environmental conditions, thus reinforcing adaptive capacity. Waymo, similarly, leverages massive simulation environments to test and refine decision trees, showing how internal models are shaped to anticipate and react purposefully to dynamic urban environments.

B. Intelligent Tutoring Systems (ITS)

Carnegie Learning's ITS exemplifies agentic traits in educational technology. The system autonomously manages individualized instruction paths, dynamically adjusting difficulty levels and presentation styles based on

each learner's profile and past interactions. Intentionality is manifest in its instructional strategies, which seek not only to deliver content but to engage learners and maximize retention.

Data logs show that the ITS can abandon a lesson plan midstream if the student exhibits signs of disengagement or confusion, instead switching to alternative modalities or remediation exercises. Such adaptability, rooted in feedback loops from user behavior, suggests a learning system that goes beyond static content delivery—indicating a form of pedagogical agency.

C. Healthcare Robotics

Robots like Moxi, used in clinical support roles, exhibit increasing levels of agency within semi-structured hospital environments. These robots autonomously navigate complex corridors, schedule deliveries, and prioritize tasks in real time. Their ability to adapt to changing human workflows—such as avoiding crowded areas or rescheduling a task due to staff availability—demonstrates context-sensitive behavior.

Empirical logs collected from Moxi's deployments indicate that the robot modifies its routines based on past successes and obstacles encountered during prior task executions. Over time, the robot's route optimization and interaction patterns reflect not just programmed responses but learning-driven adjustments, showcasing meaningful adaptivity and operational autonomy.

These case studies collectively highlight the emergence of agentic characteristics across disparate domains. While none of the systems achieves full human-like agency, their behaviors demonstrate the incremental advancement toward independent, goal-oriented AI capable of operating in dynamic and partially unpredictable environments.

D. Cross-Case Comparative Analysis

To assess the degree of agency demonstrated by each system, we applied the previously introduced three-dimensional scoring framework (Autonomy, Intentionality, Adaptivity). Each system was evaluated on a 5-point scale for each trait based on observed performance data, system documentation, and empirical logs. The results are summarized below and fig.2:

System	Autonomy	Intentionality	Adaptivity	Overall Agentic Score
Tesla Autopilot	4.5	4.0	4.0	High
Waymo Self-Driving	4.0	4.5	4.5	High
Carnegie Learning ITS	3.5	4.0	4.5	Moderate–High
Moxi Healthcare Robot	4.0	3.5	4.0	Moderate–High

	AI System A	AI System B
Autonomy	4	5
Intentionality	3	4
Adaptivity	4	4

Fig. 2. Comparative Scoring Matrix of Agentic Traits

These results suggest that while full human-equivalent agency has not yet been achieved, modern AI systems are evolving toward increasingly autonomous, intentional, and adaptive behaviors. Key trends emerged:

a. **High Autonomy**: All three domains demonstrate strong autonomous behavior, particularly in environments with defined operational rules (e.g., roads, hospital corridors).

b. **Emerging Intentionality**: ITS platforms and self-driving cars exhibit more strategic behavior in goal selection, particularly as they integrate predictive modeling.
c. **Strong Adaptivity**: Continuous learning from real-time data and feedback loops has enabled significant adaptivity, especially in ITS and driving systems.

E. Observational Theme

Several cross-cutting themes emerged from the empirical data:

a. **Environmental Interaction**: Agentic AI systems excel when allowed to interact dynamically with their environments, incorporating new data to revise internal states and actions.
b. **Task Context**: The degree of agency varies by task complexity and the openness of the system's operating environment. More structured environments yield higher agentic performance with lower risk.
c. **Human Oversight**: While agentic AI systems reduce dependence on human input, oversight remains crucial, especially in domains involving risk (e.g., health and transportation). This highlights a hybrid agency model where humans and machines co-manage autonomy.Robots like Moxi, used in clinical support roles, exhibit increasing levels of agency within semi-structured

F. Limitations and Scope of Agency

Although these systems show clear signs of agentic behavior, there are critical limitations:

a. **Lack of Self-Generated Goals**: Most systems still rely on human-defined goal structures and evaluation metrics.
b. **Ethical and Emotional Contexts**: None of the studied systems exhibit understanding of ethical nuance or emotional context—critical dimensions of full agency.
c. **Transferability**: Agency appears highly domain-specific; most systems cannot generalize across tasks or contexts, limiting the scope of their independence.

VI. DISCUSSION AND IMPLICATIONS

The findings from our empirical analysis affirm that agentic traits—autonomy, intentionality, and adaptivity—are actively emerging across multiple AI application domains. While the degree of agency varies, the trend is clear: AI systems are transitioning from passive, command-driven tools to semi-independent actors capable of initiating, adapting, and optimizing actions based on dynamic contexts. This evolution signals a profound transformation not only in the technical design of AI but also in its philosophical, ethical, and societal implications.

A. Rethinking AI as an Actor, Not a Tool

Traditional conceptions of AI situate it as a support mechanism—one that requires explicit instructions and predictable inputs to function. The rise of agentic AI challenges this paradigm. Increasingly, AI systems are behaving more like *actors* within ecosystems: navigating uncertainty, pursuing goals, and engaging in interactions that influence both human and machine counterparts.

This shift compels a re-evaluation of AI through a systems-theoretic and cybernetic lens. Rather than viewing AI as static software executing deterministic code, we must now consider its *processual identity*—as entities with learning trajectories, evolving strategies, and embedded values shaped by environmental feedback.

B. Implications for AI System Design

The emergence of agentic traits calls for a new design philosophy centered on dynamic control architectures, ethical scaffolding, and safety mechanisms. Key design implications include:

a. **Goal Flexibility**: Agentic AI should be able to interpret, refine, and negotiate goals, rather than merely execute predefined commands.
b. **Contextual Awareness**: Design must support real-time environmental sensing and situational analysis to enable nuanced decisions.
c. **Value Alignment**: As systems pursue goals independently, alignment with human values and ethical boundaries becomes critical to prevent misalignment or emergent risks.

C. Governance and Accountability

Agentic AI systems blur traditional boundaries of legal and moral responsibility. If an AI system initiates an action that results in harm or unintended outcomes, questions arise: *Who is accountable?* The developer, the user, the organization, or the system itself?

Policy frameworks must evolve to address this accountability gap. Possible interventions include:

a. **Shared Liability Models**: Where responsibility is distributed across the AI development and deployment chain.
b. **Transparency Protocols**: Mandating that AI systems log decisions, actions, and context to enable post hoc analysis and auditability.
c. **Ethical Audits and Certification**: Institutional mechanisms for assessing the agentic behavior and ethical compliance of advanced AI systems.

D. Human–AI Collaboration and the Future of Work

As agentic AI systems assume more complex decision-making roles, the nature of human labor and expertise will evolve. Rather than replacing human judgment, agentic AI is likely to reconfigure it—requiring humans to shift toward supervisory, strategic, and ethical oversight roles.

This raises a need for:

a. **Human-AI Interaction Design**: Ensuring that human users can understand, trust, and influence agentic AI behavior in real time.
b. **Training and Literacy**: Preparing the workforce to collaborate effectively with autonomous systems across industries such as healthcare, education, and transportation.

VII. ETHICAL AND GOVERNANCE IMPLICATIONS

The emergence of agentic AI raises urgent ethical and governance questions that extend beyond traditional software accountability frameworks. As these systems operate with increasing autonomy, the issue of responsibility becomes more complex. Who is liable when an autonomous vehicle makes a critical decision that leads to harm? Should an AI tutor that adapts its pedagogy in real time be held accountable for a student's learning outcome? These questions illustrate the growing disconnect between technological capability and legal responsibility.

Moreover, agentic AI challenges existing regulatory models. Current AI governance initiatives, such as the EU AI Act and the OECD AI Principles, assume a human-in-the-loop paradigm. Agentic AI systems, by contrast, may operate with limited or delayed human oversight. This necessitates the development of new regulatory standards that accommodate partial or full autonomy, including real-time auditability, intent traceability, and fail-safe mechanisms.

Ethically, the notion of machine agency forces reconsideration of concepts such as moral responsibility and ethical reasoning in non-human agents. While current AI lacks consciousness or moral reasoning in the human sense, the behavioral outcomes of agentic systems demand ethical scrutiny. As these systems increasingly influence human lives, there is a need for multi-stakeholder frameworks that ensure their alignment with societal values, fairness, and transparency.

VIII. COMPARATIVE TABLE: TRADITIONAL AI VS. AGENTIC AI

Feature	Traditional AI	Agentic AI
Decision-making	Rule-based, reactive	Goal-oriented, context-aware
Autonomy	Limited to predefined tasks	Operates independently, initiates actions
Learning	Offline or periodic retraining	Continuous, real-time adaptation
Intentionality	Executes predefined commands	Selects actions based on internal goals
Environment interaction	Fixed inputs/outputs	Dynamic, evolving behavior patterns
Accountability model	Developer/operator responsibility	Shared or distributed responsibility

IX. TECHNICAL APPENDIX

A. Scoring Rubric for Agentic Traits

Each AI system was evaluated on a 5-point scale across three dimensions:

a. Autonomy: Ability to act without human intervention
b. Intentionality: Goal-driven behavior and internal decision modeling
c. Adaptivity: Learning and behavioral adjustment over time

Score	Description
1	Minimal or no presence of the trait
2	Basic functionality, mostly static
3	Moderate presence with some variability

Score	Description
4	Advanced presence with frequent dynamic behavior
5	High, context-aware, and goal-directed behavior

B. System Evaluation Summary

System	Autonomy	Intentionality	Adaptivity
Tesla Autopilot	4	4	5
Carnegie ITS	3	4	4
Moxi Robot	3	3	4

C. Data Sources

a. Tesla Autopilot: Public logs, user-submitted edge cases, NHTSA reports
b. Carnegie ITS: Interaction datasets, adaptation logs
c. Moxi Robot: Deployment records, task performance reports from hospital partners

X. CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

This paper has explored the emergence of **agentic AI** as a paradigm shift in the field of Artificial Intelligence. Through a conceptual framework built around autonomy, intentionality, and adaptivity—and validated by empirical analysis across autonomous vehicles, intelligent tutoring systems, and healthcare robotics—we have demonstrated that modern AI systems increasingly exhibit behaviors indicative of agency. These findings challenge the prevailing view of AI as a passive tool and highlight its transformation into a semi-independent actor in various domains.

Our analysis suggests that while current systems do not yet reach full human-equivalent agency, they occupy a critical transitional space. This transition demands a rethinking of AI development, emphasizing not just performance and efficiency, but ethical design, contextual responsiveness, and shared accountability. The ability of AI to make and adapt decisions in dynamic environments, sometimes without human oversight, requires new frameworks for governance, interaction, and trust.

A. Limitations

Despite its contributions, this study has several limitations:

a. Data Constraints: The analysis relied on secondary data sources such as white papers, publicly available logs, and technical reports. Direct access to proprietary decision-making models and internal learning parameters was limited, which may have constrained the depth of analysis.
b. Domain Specificity: The study focused on three application domains. While these are representative, the generalizability of findings to other areas (e.g., financial trading, military systems, creative AI) remains to be tested.
c. Subjective Evaluation: The scoring of agentic traits, while structured, retains a degree of subjectivity, particularly in assessing intentionality. More robust, standardized metrics are needed for broader comparative studies.

B. Future Work

Future research should address these limitations and deepen the exploration of agentic AI through several avenues:

a. Longitudinal Studies: Observing AI systems over time in real-world environments could yield richer insights into the evolution of agentic behavior, particularly in learning and adaptation.
b. Cross-Domain Expansion: Expanding empirical analysis to additional domains such as finance, security, or environmental monitoring would help validate the framework's applicability across sectors.
c. Human-AI Interaction Studies: Investigating how humans perceive, trust, and collaborate with agentic AI will be critical to ensuring safe and effective integration into society.
d. Normative Frameworks: Research should also focus on developing normative ethical and legal models that address the unique challenges posed by systems with partial or full agency, especially in high-stakes environments.

REFERENCES

[1]. Brooks, R. A. (1991). Intelligence without representation. Artificial Intelligence, 47(1-3), 139–159. https://doi.org/10.1016/0004-3702(91)90053-M
[2]. Clark, A. (1997). Being there: Putting brain, body, and world together again. MIT Press.
[3]. Davis, R., & Lenat, D. B. (1982). Knowledge-based systems in artificial intelligence. McGraw-Hill.
[4]. Dennett, D. C. (1987). The intentional stance. MIT Press.
[5]. Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1
[6]. Pfeifer, R., & Bongard, J. (2007). How the body shapes the way we think: A new view of intelligence. MIT Press.
[7]. Russell, S. J., & Norvig, P. (2020). Artificial intelligence: A modern approach (4th ed.). Pearson.
[8]. Wooldridge, M. (2009). An introduction to multiagent systems (2nd ed.). Wiley.