International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (Online): 2347-5552, Volume-13, Issue-5, September 2025
DOI: https:/doi.org/10.55524/ijircst.2025.13.5.3
Article ID IRP-1674, Pages 14-24
www.ijircst.org

AI Governance in the Era of Agentic Generative AI and AGI: Frameworks, Risks, and Policy Directions

Satyadhar Joshi

Independent Researcher, Alumnus, International MBA, Bar-Ilan University, Israel

Correspondence should be addressed to Satyadhar Joshi; satyadhar.joshi@gmail.com

Received 27 July 2025; Revised 11 August 2025; Accepted 26 August 2025

ABSTRACT- The accelerating development of agentic artificial intelligence (AI) and the prospect of artificial general intelligence (AGI) create unprecedented opportunities alongside complex governance challenges. This paper examines the ethical, regulatory, and technical dimensions of governing highly autonomous AI systems, drawing upon more than fifty contemporary academic and policy sources. Three core insights emerge. First, current governance structures provide limited coverage of risks linked to recursive self-improvement and multi-agent coordination, with only an estimated 10–15% of safety research addressing impacts that arise after deployment. Second, economic projections suggest that agentic AI could generate between 2.6 and 4.4 trillion USD in added global output by 2030, yet automation could replace approximately 28–42% of existing job tasks, making proactive workforce transition strategies a policy necessity. Third, fragmented regulatory approaches remain a concern; in the United States, for example, 70–75% of critical infrastructure is considered vulnerable to adversarial autonomous systems. To address these issues, we propose a governance model built on three pillars: modular agent design, adaptive safety mechanisms, and international coordination. Policy measures such as licensing thresholds for high-computer systems exceeding 10^25 FLOPs, structured red-team testing across public and private sectors, and fiscal incentives for governance-by-design practices are advanced as actionable pathways. Overall, the study argues for adaptive, globally coordinated governance frameworks that balance innovation with systemic risk mitigation in the era of agentic AI and AGI. his is a pure review paper and all results, proposals and findings are from the cited literature.

KEYWORDS- AI Governance, Agentic AI, Generative AI, Artificial General Intelligence (AGI), Ethics, Policy, Risk Management, Recursive Self-Improvement, Multi-Agent Systems, Workforce Transition, International Coordination

I. INTRODUCTION

The emergence of *agentic artificial intelligence (AI)*—systems capable of autonomous goal-setting, decision-making, and task execution—marks a fundamental paradigm shift in computational intelligence. Unlike conventional generative AI, which operates within static prompt–response frameworks, agentic AI demonstrates dynamic adaptability, recursive self-improvement (RSI), and multi-agent collaboration [1], [2]. These capabilities enable agentic systems not only to generate content but also to plan, execute, and optimize tasks with minimal human oversight. Parallel advancements in *artificial general intelligence (AGI)* further amplify both the transformative potential and systemic risks of these technologies. Current economic projections suggest that agentic AI could contribute between $2.6 and $4.4 trillion to global GDP by 2030, while simultaneously automating 28–42% of job-related tasks [3], [4]. Such forecasts underscore the dual challenge of harnessing productivity gains while mitigating widespread labor market disruptions.

Despite growing awareness, existing governance frameworks remain ill-equipped to manage the complexities of agentic AI and AGI. This paper identifies three central gaps in current approaches. First, regulatory and ethical models provide insufficient guidance for addressing risks associated with recursive self-improvement and autonomous multi-agent coordination. Second, regulatory fragmentation across jurisdictions hinders coherent oversight; in the United States alone, estimates indicate that 70–75% of critical infrastructure remains exposed to adversarial exploitation by autonomous systems [5]. Third, policy tools remain largely static, lacking the adaptive mechanisms necessary to keep pace with rapidly evolving agentic technologies.

To address these issues, this paper synthesizes insights from over fifty contemporary sources, including academic literature (32%), industry reports (28%), and government publications (20%), with a particular emphasis on policy developments in 2024–2025, such as the European Union's *AI Act* [6] and the United States' *Executive Order 14110* [7]. Our contributions are fourfold:

- Conceptual foundations: We review the terminology and governance challenges of agentic AI and AGI, supported by definitional clarity (Table I) and forward-looking projections (Figure 1).
- Technical governance frameworks: We introduce compliance models such as governance scoring (Eq. 1) and RSI optimization methods (Algorithm 1).
- Comparative policy analysis: We evaluate governance strategies across jurisdictions and sectors, summarized in Table 4.
- Tripartite governance architecture: We propose an integrated model combining modular agent design,