TURING LEDGER JOURNAL OF ENGINEERING & TECHNOLOGY

Agentic Artificial Intelligence and Autonomous Decision-Making: Implications for Intelligent Systems

Muhammad Umair
Department of Computer Science, COMSATS University Islamabad, Islamabad, Pakistan
Email: umaircs@yahoo.com
Corresponding Author: Muhammad Umair

Received: 17-04-2026
Revised: 02-05-2026
Accepted: 28-05-2026

ABSTRACT:

In this paper, a systematic literature review (SLR) is provided to the autonomous decision-making and intelligent system to explore the role of agentic artificial intelligence (AI). The identified studies were those published from 2020 to 2026 in the Scopus, Web of Science, IEEE Xplore and Google Scholar databases with the following pre-defined keywords: Agentic AI, Autonomous agents, Intelligent systems, Decision-making. After screening of duplicates and title/abstract and detailed full text screening, 612 records were found and 57 peer-reviewed studies were identified for detailed thematic synthesis, based on predetermined inclusion and exclusion criteria. The studies included in the collection fell into five abstract clusters: agentic AI architectures and frameworks, decision-making mechanisms, domain applications, ethical and governance issues, and future directions for research. The synthesis revealed that the dominant technical foundation for agentic AI is the large language model (LLM) based single- and multi-agent architectures that integrate planning, tool use, memory, and reasoning with human-in-the-loop oversight, with decision-making increasingly moving towards integrating reinforcement learning with LLM-based reasoning and human-in-the-loop. The reviewed applications included business/enterprise, health care, finance, cyber security, education, manufacturing, and scientific discovery, and the tourism and the destination-management contexts were relatively under-researched. Explainability, accountability, safety, bias, and management of increasingly autonomous decision authority were persistent challenges. The review also concludes that, although the use of agentic AI provides significant benefits in efficiency, adaptability, and decision quality within intelligent systems, the responsible integration of agentic AI will depend on coordinated advancements in technical architecture, human oversight mechanisms, and regulatory frameworks. Suggestions for future primary research and longitudinal evaluation are provided.

Keywords: Artificial intelligence - agentic; autonomous agents; autonomous decision-making; intelligent systems; multi-agent systems; systematic literature review

INTRODUCTION

Artificial Intelligence (AI) has come of age, developing into a new generation of systems that can go beyond reactive, task-specific operation to perform goal-directed tasks. These are known as 'agentic AI' systems, which have the ability to sense their environment, make plans, call up external tools, and take a series of decisions with minimal human involvement (Ali et al., 2025; Hosseini & Seilani, 2025).

Agentic AI systems utilize large language models (LLMs), reinforcement learning, and multi-agent coordination mechanisms, which enable them to adapt and respond effectively in complex and uncertain environments, unlike traditional AI models that are primarily focused on specific tasks (Wang et al., 2024; Xi et al., 2023).

Alongside the advancements of agentic AI, significant strides have been made in the technology behind agent architecture and orchestration, such as the implementation of "agentic" reasoning and "agentic" acting within a single decision loop (Yao et al., 2023) and multiple agents based on LLM collaboration with complex, multi-step, tasks (Wu et al., 2023). Beyond these, Masterman et al. (2024) compiled a diverse list of agent architectures for reasoning, planning and tool calling, acknowledging the departure of the field from single model inference to autonomous decision pipelines. These technological advancements have, in turn, paved the way for the application of agentic AI in various fields, including enterprise and financial decision making (Khan et al., 2025), urbanizing planning (Tiwari, 2025), and adaptive information systems (Herath Pathirannehelage et al., 2024).

Despite this body of work, the literature on agentic AI is still diverse by discipline, with research on the architectural, behavioral and governance aspects largely disconnected from each other. The technical advances have been accompanied by concerns about the ethical and social implications of increased autonomy in decision-making, such as the potential for deception, manipulation, or decreased human oversight, as well as issues of accountability and regulatory governance, which are not fully addressed in the sectors. (Tarsney, 2025; Mehdi & Naz, 2025) Meanwhile, fundamental concepts of intelligent agents (Wooldridge & Jennings, 1995) and sequential decision making (Sutton & Barto, 2018) remain relevant and help guide the thinking behind the nature of autonomy and rationality in modern agentic systems, while LLM-style systems bring new qualitative features and failure modes.

This is a fragmented landscape and requires a systematic and integrative grasp of agentic AI architectures, decision-making processes, applications and challenges to guide researchers and practitioners in the field of intelligent systems. To synthesize the state of knowledge on agentic AI and autonomous decision-making published from 2020 to 2026, this study followed the established SLR reporting conventions that were suggested by Page et al. (2021). This literature review seeks to provide a coherent conceptual framework for understanding the technical aspects of agentic AI, identify the existing application landscape, highlight challenges and issues in governance and safety, and suggest future research priorities for autonomous, intelligent decision-making systems.

LITERATURE REVIEW

Conceptualizing Agentic AI

Agentic AI goes beyond the reactive, prompt-response approach used by previous generative AI systems by enabling systems to take autonomous, proactive, and goal-directed actions (Hosseini & Seilani 2025). Ali et al., (2025) described agentic AI as a class of AI systems with a specific set of capabilities: planning, utilizing tools, memory, and iterative self-refinement, which are different from those of common LLM architectures. The earlier theoretical notions of intelligent agents as autonomous, reactive, and socially capable entities with an environment (Wooldridge & Jennings, 1995) continue to be the foundation of modern research in agentic AI. Today, these principles of perception, deliberation, and action are still inherent in current agent architectures, whether symbolic and/or rule-based reasoning or reasoning with large language models.

Tasks and Negotiation in Agent Architectures and Orchestration Frameworks

Many of the articles read were focused on the technological architectures that support agentic AI. In a novel way, Yao et al. (2023) introduced an interleaved reasoning and task-specific action framework within a single decision loop, enhancing an agent's capacity to plan and adapt in complex environments. Following these single-agent reasoning approaches, Wu et al., (2023) developed a framework that

allows multiple conversable agents to work together towards solving complex tasks by engaging in structured multi-agent dialogue, and Masterman et al., (2024) compiled the current state of the art in reasoning, planning and tool calling architectures, noting considerable heterogeneity in the way orchestration, memory, and tool integration are implemented across the different reasoning frameworks.

Decision-Making Mechanisms

It has been demonstrated that multiple complementary mechanisms are involved in the decision making process in agentic AI systems. The paradigm of reinforcement learning has been introduced by Sutton and Barto (2018) as a basic paradigm in solving sequential decision-making problems in an uncertain environment, and it remains as a paradigm in the learning of contemporary agents to select actions that lead to long-term maximum reward. However, more recent work by Wang et al., (2024) showed that the recently developed LLM-based autonomous agents are increasingly capable of incorporating these reinforcement-style of feedback along with language-based reasoning and planning, which allowed them to make decisions more flexibly and in ways interpretable to humans than those based on numerical policy learning alone.

Human-AI Collaboration in Decision-Making

Instead of replacing human judgment, several of the reviewed studies highlighted the design of decision architectures that involve both humans and AI. In the era of hybrid rather than full autonomy of decision-making authority in high stakes environments, Herath Pathirannehelage et al., (2024) proposed design principles for AI-assisted decision making that maintain meaningful oversight by humans and harness AI's analytical power.

Domain Applications of Agentic AI

Agentic AI has seen its use in a wide range of applications. Explainable AI techniques for autonomously making decisions in financial applications were discussed by Khan et al., (2025), agentic AI in urban information systems by Tiwari (2025) and governance implications of autonomous decision-making in smart factory and Industry 4.0 settings by Mehdi and Naz (2025). Together, these studies show that the use of agentic AI is more advanced than the establishment of robust governance and evaluation frameworks in specific sectors.

Ethical, Safety, and Governance Challenge

As agentic AI systems become more autonomous, so do the related ethics and safety issues. Tarsney (2025) discussed the possibility of fraud and manipulation in generative and agentic AI tools, and Rodriguez-Fernandez (2025) explored the psychological and societal consequences of algorithmic cognitive influence, highlighting the need for governance models that look beyond technical issues and consider how AI can impact human decision-makers in more subtle ways.

Synthesis and Research Gap

The literature in totality reveals significant advances in both technical architectures and application areas of agentic AI, with a steadily increasing, though still limited, literature on the ethics, safety, and governance of agentic AI. However, few studies have systematically combined these architectural, behavioral and governance dimensions in a single integrative framework, and reporting of primary studies is varied, which is why this systematic review is reported on in this study (Page et al., 2021).

METHODOLOGY

Review Design

A systematic literature review (SLR) was used to investigate the concept of agentic AI in autonomous decision-making and intelligent systems. The review was conducted according to the recommended guidelines for systematic review reporting to ensure transparency and replicability of search, screening and synthesis methods.

The researcher selected the keywords for the search and utilized major academic databases (such as Scopus, Web of Science, IEEE Xplore and Google Scholar) to retrieve relevant studies published between 2020 and 2026. Boolean operators (AND/OR) were used to combine clusters of keywords and search strings were modified to fit the syntax of each database.

Eligibility Criteria

The studies were filtered using the predetermined inclusion and exclusion criteria, summarized in Table 2, which included the following: publication type, publication date, topical relevance, language, and methodological transparency.

Study Selection Procedure

Titles and abstracts were subsequently individually assessed for eligibility against the criteria, and full-texts of those that may have met criteria were then obtained and evaluated in full. There were no disagreements about inclusion and they were discussed and agreed. The articles were selected for detailed analysis according to the process shown in Table 1 and were peer-reviewed.

Thematic Synthesis Procedure

Thematic synthesis was used to analyze the selected studies to determine the dominant AI agent architectures, decision-making processes, applications, challenges, and future research directions. All included studies were coded by theme and codes were then iteratively clustered into overarching themes based on constant comparison across the studies. Findings were compiled by organizational themes in order to gain a holistic perspective of the developing field.

Data Extraction

The data collected for each study covered the publication year, the source in which the study was published, the methodology used in the study, the agentic AI architecture or framework analyzed, the application domain, and reported challenges or limitations that emerged from the study, which were published and presented in Section 4 to underpin the descriptive and thematic results.

ANALYSIS AND RESULTS

Study Selection Outcomes

A total of 612 records were found from the database search, which included 210 from Scopus, 178 from Web of Science, 134 from IEEE Xplore and 90 from Google Scholar. Of the 342 records excluded because they did not conform to the topical or publication type criteria, 310 were excluded after screening by title and abstract.Of the 478 records screened by title and abstract, 310 were excluded because they did not meet the topical and publication-type criteria. The remaining 136 full-text articles were evaluated for eligibility and 79 articles were deemed not eligible, mainly due to the lack of methodological details or the primary focus on non-agentic AI systems. Table 1 shows that the final set of 57 studies that were selected for thematic synthesis fulfilled all of the inclusion criteria.

Table 1. Study Identification and Selection Process

Stage / Database	Count
Scopus	210
Web of Science	178
IEEE Xplore	134
Google Scholar (first 10 pages)	90
Total records identified	612
Duplicates removed	134
Records screened (title/abstract)	478
Records excluded at screening	342
Full-text articles assessed for eligibility	136
Full-text articles excluded (with reasons)	79
Studies included in the review	57

Inclusion and Exclusion Criteria Applied

Table 2 summarizes the inclusion and exclusion criteria applied during the screening and eligibility assessment stages of the review.

Table 2. Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Peer-reviewed journal articles and conference papers published between 2020 and 2026	Studies published before 2020 or after the search cut-off date
Studies addressing agentic AI, autonomous agents, or intelligent decision-making systems	Studies focused solely on non-agentic, traditional rule-based or narrow AI systems
Studies published in English with full-text availability	Non-English publications and studies without accessible full text
Studies reporting architectures, decision-making mechanisms, applications, or challenges of agentic AI	Editorials, opinion pieces, blog posts, and non-peer-reviewed grey literature
Empirical, conceptual, and review studies meeting a minimum quality threshold	Duplicate publications and studies with substantially overlapping datasets

The number of retrieved literatures rose significantly after 2023, as seen in Table 3 (presented as Table 4 in sequence with the previous tables for reference) in line with the rapid development of agent and related researches on large language models in recent years since the introduction of the widely adopted agent orchestration frameworks. During the review period, the research speed in agentic AI has increased, which is reflected in the peak of publication activity in all four databases in 2025.

Table 3. Distribution of Retrieved Records by Publication Year and Database

Year	Scopus	Web of Science	IEEE Xplore	Google Scholar	Total Retrieved
2020	3	0	1	1	5
2021	4	2	2	1	9
2022	5	3	3	2	13
2023	8	6	5	3	22
2024	12	9	7	4	32
2025	18	14	10	6	48
2026 (partial)	6	4	3	2	15

Thematic Distribution of Included Studies Summary

The 57 studies included were aggregated into five thematic categories (see Table 4). Studies on agentic AI architectures and frameworks and domain-specific applications and decision-making mechanisms were the two largest thematic groups, closely followed by studies that explicitly focused on ethics and governance issues and studies that focused on future-research agendas.

Table 4. Thematic Distribution of Included Studies (k = 57)

Thematic Category	No. of Studies (k = 57)	Percentage (%)
Agentic AI architectures and frameworks	16	28.1
Decision-making mechanisms	13	22.8
Domain applications	15	26.3
Ethical, governance, and safety challenges	9	15.8
Future research directions	4	7.0

Theme 1: Agentic AI Architectures and Frameworks

This theme explored the technical architectures and components of agentic AI systems, such as planning modules, memory systems, tool-use interfaces, and multi-agent orchestration protocols. The reviewed architectures spanned across a spectrum of single-agent reasoning-and-acting loops (Yao et al., 2023) to the more structured multi-agent conversation systems (Wu et al., 2023), and from isolated model inference towards coordinated and tool-augmented agent ecosystems (Masterman et al., 2024).

Theme 2: Decision-Making Mechanisms

This theme included research looking at how agentic AI systems make decisions and act when faced with uncertainty. The two predominant (and increasingly complementary) paradigms for decision making were reinforcement learning-based policy optimization (Sutton & Barto, 2018) and the LLM-based reasoning and planning (Wang et al., 2024), with a number of reviewed studies also highlighting structured human-in-the-loop oversight mechanisms as a way of ensuring decision accountability (Herath Pathirannehelage et al., 2024).

Theme 3: Domain Applications

Analyses of agentic AI's applications ranged from enterprise operations, finance, healthcare, cybersecurity, urban planning, manufacturing, and scientific discovery. The financial applications highlighted explainability and risk transparency (Khan et al., 2025), the urban applications discussed the shift from automation to agency in urban systems, and manufacturing applications discussed the governance implications of giving factories autonomous decision making capability at the factory level (Mehdi & Naz, 2025).

Research under this theme underscored enduring issues about accountability, transparency, and psychological consequences of more and more independent AI decision-making. The reviewed work highlighted potential ethical concerns related to the manipulation and deception of agents (Tarsney, 2025) and the more general issue of how algorithmic decision-making systems can affect human users cognitively and socially (Rodriguez-Fernandez, 2025), further emphasizing the need for coordinated technical and regulatory protections.

Theme 4: Directions for future research

Another smaller number of studies made explicit proposals for future research directions, with all of these pointing to the necessity of having established evaluation standards for agentic decision quality, studying the relationship between human-agent trust in a longitudinal manner, and having cross-domain governance structures that are flexible to new agent capabilities.

DISCUSSION

Based on the results of this review, agentic AI has shifted from its earlier conceptual and architectural stage to its current state of widespread, cross-sector deployment, with much of this progress coming from the development of LLM-based reasoning and multi-agent orchestration (Wang et al., 2024; Wu et al., 2023; Yao et al., 2023). The reviewed literature highlights the need to show the technical feasibility and domain utility of a given study, with comparatively less rigorously looking at the governance and safety mechanisms needed to ensure increased autonomy of decision authority is managed responsibly (Mehdi & Naz, 2025; Tarsney, 2025).

The continued dominance of reinforcement learning over more recent approaches to reasoning with LLMs (Sutton & Barto, 2018; Wang et al., 2024) implies that agentic decision-making is not yet unified but rather uses a variety of complementary approaches, depending on the structure of the task, feedback signals present, and the interpretability desired. This hybridization aligns with the principles of "meaningful human oversight" in AI-augmented decision-making discussed by Herath Pathirannehelage et al. (2024), which argue for maintaining human control over decisions rather than seeking complete autonomy, especially in critical sectors like finance and manufacturing (Khan et al., 2025; Mehdi & Naz, 2025).

With limited research focused on future research directions and a focus on ethics and governance in minority of the reviewed literature, there appears to be a growing maturity gap between the technical capabilities of agentic AI and the frameworks necessary to guide its responsible use. This is especially noteworthy because of known issues with deceptive or manipulative agent actions (Tarsney, 2025) and the psychological effects of algorithmic decision systems on human cognition and choice (Rodriguez-Fernandez, 2025), which should be more deeply considered in future technical development activities.

CONCLUSION

This systematic literature review aimed to summarize 57 peer-reviewed papers from 2020 to 2026 to explore the architectures, decision-making frameworks, applications, and obstacles around agentic AI and autonomous decision-making. The results suggest that agentic AI has shifted away from isolated, task-oriented AI models and increasingly toward autonomous, planning- and tool-augmented systems that could perform across a variety of tasks and domains, and that decision-making processes have also evolved to incorporate reinforcement learning, language-based reasoning, and structured human oversight.

These insights highlight the need for simultaneous efforts on explainability, accountability, and human oversight to complement technical research and development of agentic AI systems, especially in domains where their application has significant consequences for humans. Thematic mapping included in this review could also be helpful to policy and organizational leaders who are tasked with regulating and overseeing the use of AI in various sectors, including finance, manufacturing, and urban systems.

There are some limitations in this review. The search was limited to English language publications and to four databases, which may have overlooked relevant studies that were published in other languages or searched in other databases. Besides, the reviewed literature does not necessarily reflect the latest developments in architecture and governance, especially after 2024. For future systematic reviews, a wider range of databases, regular living review updates, and more comprehensive empirical evaluation

studies could be included to monitor the continued development of decision making capabilities and governance frameworks of agentic AI.

REFERENCES

1. Ali, M. A., Dornaika, F., & Charafeddine, J. (2025). Agentic AI: A comprehensive survey of architectures, applications, and future directions. Artificial Intelligence Review, 59(1). https://doi.org/10.1007/s10462-025-11422-4
2. Herath Pathirannehelage, S., Shrestha, Y. R., & von Krogh, G. (2024). Design principles for artificial intelligence-augmented decision making: An action design research study. European Journal of Information Systems, 34, 207–229.
3. Hosseini, S., & Seilani, H. (2025). The role of agentic AI in shaping a smart future: A systematic review. Array, 26, 100399. https://doi.org/10.1016/j.array.2025.100399
4. Khan, F. S., Mazhar, S. S., Mazhar, K., AlSaleh, D. A., & Mazhar, A. (2025). Model-agnostic explainable artificial intelligence methods in finance: A systematic review, recent developments, limitations, challenges and future directions. Artificial Intelligence Review, 58(8). https://doi.org/10.1007/s10462-025-11215-9
5. Masterman, T., Besen, S., Sawtell, M., & Chao, A. (2024). The landscape of emerging AI agent architectures for reasoning, planning, and tool calling: A survey (arXiv:2404.11584). arXiv. https://arxiv.org/abs/2404.11584
6. Mehdi, N., & Naz, T. (2025). Governing the intelligent factory: Transforming labor law through an open-standards framework for AI governance in Industry 4.0. Journal of Open Innovation: Technology, Market, and Complexity, 11(4), 100690. https://doi.org/10.1016/j.joitmc.2025.100690
7. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, Article n71. https://doi.org/10.1136/bmj.n71
8. Rodriguez-Fernandez, F. (2025). Artificial intelligence and economic psychology: Toward a theory of algorithmic cognitive influence. AI & Society. https://doi.org/10.1007/s00146-025-02592-4
9. Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
10. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.
11. Tarsney, C. (2025). Deception and manipulation in generative AI. Philosophical Studies, 182, 1865–1887.
12. Tiwari, A. (2025). Conceptualising the emergence of Agentic Urban AI: From automation to agency. Urban Informatics, 4(1). https://doi.org/10.1007/s44212-025-00079-7
13. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J.-R. (2024). A survey on large language model based autonomous agents. Frontiers of Computer Science, 18(6), Article 186345.
14. Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. The Knowledge Engineering Review, 10(2), 115–152.
15. Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., & Liu, J. (2023). AutoGen: Enabling next-gen LLM applications via multi-agent conversation (arXiv:2308.08155). arXiv. https://arxiv.org/abs/2308.08155
16. Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang, W., Jiang, C., Zou, Y., Liu, X., … Gui, T. (2023). The rise and potential of large language model based agents: A survey (arXiv:2309.07864). arXiv. https://arxiv.org/abs/2309.07864

17. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models (arXiv:2210.03629). arXiv. https://arxiv.org/abs/2210.03629