UDC 004.8:004.775

THE ERA OF «AGENTIC AI»: WHEN CHATBOTS STARTED DOING THINGS

Yefimau M. I., Baulina I. P.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Perevyshko A. I. – Senior Lecturer at the Department of Foreign Languages

Annotation. The article is devoted to the evolution of artificial intelligence (AI) systems from reactive models to autonomous agentic structures. Their main advantages are emphasized. Special attention is given to main characteristics and working principles of Agentic AI. It is pointed out the significance of further advances in the developing AI systems.

Keywords: Agentic AI, autonomous systems, iterative cycle, planning, multi-agent systems, digital transformation.

***Introduction*.** In the modern field of computer science, there is a fundamental transformation of artificial intelligence (AI) systems, characterized by a shift from reactive linguistic models to autonomous agent structures (Agentic AI). While previous iterations of large language models primarily operated within the «query-response» paradigm, generating text content based on statistical patterns, modern agent systems exhibit a distinct sense of agency [1]. The primary objective of this article is to analyze the mechanisms that transform AI from a passive information intermediary into an active digital actor capable of independently planning and executing complex tasks in dynamic environments.

Main part. A distinctive feature of agent-based AI is the presence of an architecture that ensures a closed loop of autonomous functioning [1]. Unlike standard chatbots, the agent's operational cycle includes the stages of environment perception, decomposition of a high-level goal into a sequence of subtasks, and selection of relevant tools. This iterative approach allows the system to interpret execution errors as feedback for optimizing its future strategy. The technological complexity of Agentic AI stems from the integration of several critical cognitive modules, including strategic planning based on chain-of-thought (CoT) reasoning and advanced tree-search algorithms. These algorithms allow the model to simulate multiple potential outcomes before committing to a specific digital action, effectively creating a «mental sandbox» for problem-solving. This simulation-based approach minimizes the probability of stochastic errors by evaluating the syntactic and semantic validity of a planned action before its actual deployment in the production environment.

The use of external vector data stores, often implemented through Retrieval-Augmented Generation (RAG), combined with a short-term context window, enables the accumulation of experience and the preservation of long-term system states [4]. Furthermore, the synchronization between episodic and semantic memory modules allows the agent to generalize from previous experiences, thereby reducing the computational cost of repetitive task planning. This dual-layer memory architecture is crucial for maintaining consistency across long-duration tasks that may span several days or involve thousands of individual sub-steps [6]. Furthermore, the implementation of «self-reflection» loops allows the agent to critique its own intermediate outputs. If a generated solution fails to meet predefined logical constraints, the agent can autonomously trigger a correction cycle, bypassing the need for human intervention.

The most important element is the action space, which is a set of software interfaces for interacting with the external world. This includes code execution in isolated «containerized» environments, direct API interaction, and autonomous navigation in browser-based interfaces. The robustness of the action space is further enhanced by the implementation of feedback-driven wrappers that translate raw system outputs into structured data, suitable for the model’s internal perception module. These capabilities are complemented by a continuous monitoring module that tracks changes in the external environment, ensuring the agent can adapt to real-time data fluctuations. As agents move from sandboxed environments to production-level systems, the necessity for robust «Guardrails» becomes paramount. These are programmatic constraints that prevent the agent from

executing high-risk or irreversible commands, such as unauthorized financial transactions or critical system deletions, without explicit supervisor approval [7].

Special attention should be paid to the development of multi-agent systems (MAS), where problem-solving is achieved through the orchestration of multiple specialized models. In such ecosystems, there is a rigorous division of cognitive labor: separate modules are responsible for generating solutions, performing adversarial critical analysis, and managing resource coordination. This modularity not only increases the reliability of the system but also allows for the integration of heterogeneous models, where each agent is optimized for a specific domain [4]. This shift from individual processing to collaborative cognitive labor necessitates the development of standardized inter-agent communication protocols to prevent logical conflicts and deadlocks during collective problem-solving.

The deployment of Agentic AI also faces the critical challenge of balancing reasoning depth with operational latency. High-level planning requires significant computational overhead, which can lead to delays in execution. To optimize this, modern architectures are adopting a hierarchical approach where a lightweight model manages routine subtasks, while a more sophisticated «orchestrator» handles strategic decision-making [5]. Empirical data suggests that such hierarchical structures significantly reduce the inference latency, making autonomous agents viable for real-time applications in telecommunications and automated financial trading. This division of labor ensures that the system remains responsive while maintaining its ability to solve complex, non-linear problems.

Furthermore, the issue of «explainable agency» is becoming central to the development of trusted autonomous systems [4]. As agents begin to operate in high-stakes environments, it is no longer enough for the system to simply produce a result. It must provide a transparent, human-readable audit trail of its internal reasoning process. This transparency allows human supervisors to intervene if the agent's logic begins to drift toward unintended or risky shortcuts, ensuring that the AI's goals remain strictly aligned with human intent.

In addition to internal logic, the scalability of agentic systems is increasingly dependent on cloud-native integration and distributed computing [7]. By leveraging containerized microservices, agents can dynamically scale their processing power depending on the complexity of the task at hand. This framework is essential for industrial-scale automation where reliability and resource efficiency are the primary benchmarks of success. The integration of these systems into corporate infrastructures also necessitates a rethinking of digital identity. As agents act on behalf of users, the boundaries of legal and technical responsibility become blurred [2]. The problem of goal alignment becomes central, requiring the agent to consider ethical and operational constraints rather than simply finding the shortest path to a result

***Conclusion*.** The analysis confirms that the era of agent-based AI marks a significant leap forward in the development of automation technologies. The transformation of AI from a tool into an autonomous agent changes the nature of human-machine interaction, shifting the focus from detailed instructions to goal-setting management. This transition represents a move toward «Artificial General Intelligence» (AGI) characteristics within specialized domains. The future of this field will depend on the scientific community's ability to ensure the transparency of decision-making algorithms and create architectures that combine high autonomy with guaranteed predictability and safety in dynamic digital environments.

References

1. What is agentic AI? [Electronic resource]. − Mode of access: https://cloud.google.com/discover/what-is-agentic-ai#agentic-ai-versus-generative-ai. − Date of access: 14.02.2026.

2. What is agentic AI? [Electronic resource]. − Mode of access: https://cloud.google.com/discover/what-is-agentic-ai#agentic-ai-versus-generative-ai. − Date of access: 21.02.2026.

3. Agentic AI, explained [Electronic resource]. − Mode of access: https://mitsloan.mit.edu/ideas-made-to-matter/agentic-ai-explained. − Date of access: 01.03.2026.

4. What is agentic AI? [Electronic resource]. − Mode of access: https://aws.amazon.com/what-is/agentic-ai/. − Date of access: 02.03.2026.

5. What is agentic AI? [Electronic resource]. − Mode of access: https://cloud.google.com/discover/what-is-agentic-ai#agentic-ai-versus-generative-ai. − Date of access: 18.02.2026.

6. The Next "Next Big Thing": Agentic AI's Opportunities and Risks [Electronic resource]. − Mode of access: https://scet.berkeley.edu/the-next-next-big-thing-agentic-ais-opportunities-and-risks/. − Date of access: 19.02.2026.

7. Deploying Agentic AI − What Worked, What Broke, and What We Learned [Electronic resource]. − Mode of access: https://realworlddatascience.net/applied-insights/case-studies/posts/2025/08/12/deploying-agentic-ai.html. − Date of access: 20.02.2026.