

# The Future of Artificial Intelligence: Toward AI Systems that can Develop and Improve Themselves

Ahmed Alsiwar<sup>1</sup>

<sup>1</sup>Saudi Aramco

Publication Date: 2026/09/26

**Abstract:** Artificial Intelligence (AI) is progressing from systems designed primarily to assist human users toward increasingly autonomous systems capable of reasoning, planning, coding, experimentation, and scientific discovery. One of the most significant potential developments in this field is the emergence of AI systems that can contribute to the development and improvement of future AI systems. Rather than representing an abrupt transition to completely autonomous intelligence, this development is more realistically understood as a gradual expansion of AI-assisted research and engineering. Recent advances in AI agents, automated software engineering, machine-learning experimentation, and multi-agent scientific research demonstrate that AI can already perform important components of the development cycle. Research systems are increasingly capable of generating hypotheses, writing and testing code, conducting experiments, evaluating alternative solutions, and iteratively improving their outputs. However, current systems remain limited in reliability, generalization, independent scientific reasoning, and their ability to consistently reproduce or surpass expert human research.

**How to Cite:** Ahmed Alsiwar (2026) The Future of Artificial Intelligence: Toward AI Systems that can Develop and Improve Themselves. *International Journal of Innovative Science and Research Technology*, 11(9), 1581-1583.  
<https://doi.org/10.38124/ijisrt/26sep680>

## I. INTRODUCTION

Artificial Intelligence has evolved from specialized computational systems into increasingly general-purpose technologies capable of performing tasks that previously required significant human cognitive effort. This evolution has created a new stage in the development of computing: AI is no longer limited to executing predefined instructions, but can increasingly participate in processes that involve reasoning, experimentation, and decision-making. A particularly important question has emerged from this progress: What happens when AI becomes capable of contributing to the development of AI itself? The question should not be interpreted simply as the idea of a machine independently rewriting its own intelligence. A more realistic and technically meaningful concept is the gradual automation of parts of the AI development lifecycle. An AI system may assist researchers in designing architectures, generating and testing code, optimizing training procedures, analyzing experimental results, identifying weaknesses, and proposing new research directions. As these capabilities improve, AI could become an increasingly important participant in the process through which future AI systems are designed. This transition could have a major impact. If AI can help speed up software development, research, algorithms, and AI models, technological progress could happen much faster. However, more AI independence also makes safety, verification, accountability, and human control more important. The goal

of this article is not to suggest that AI will suddenly become fully autonomous.

Instead, it examines the more plausible progression toward increasingly capable AI systems that participate in their own development and explores the opportunities and challenges that this expansion may create.

### ➤ Current Level of Developing the System

The development of AI systems capable of contributing to their own improvement is still an emerging field. Current technology should not be described as fully self-developing AI. Instead, the industry is moving toward systems in which AI agents can independently perform increasingly long and complex sequences of research and engineering tasks while remaining within human-designed environments and constraints. One of the clearest areas of progress is AI-assisted software engineering. Modern coding agents can inspect large codebases, identify software problems, generate solutions, execute tests, revise unsuccessful approaches, and implement complex changes. These capabilities are particularly relevant because modern AI systems themselves depend heavily on software engineering and experimentation. A second important area is machine-learning engineering. AI agents are being tested on tasks involving dataset preparation, model training, experimentation, and optimization. The MLE-bench evaluation, for example, assessed AI agents across 75

machine-learning engineering competitions and demonstrated that current systems can perform meaningful parts of the ML development process, while remaining far from consistently matching expert performance. A third area is AI-assisted scientific research. Systems such as Co-Scientist demonstrate a more advanced form of AI participation. Rather than simply answering questions, the system can generate hypotheses, evaluate competing ideas, and iteratively refine promising research directions. Google DeepMind reports applications in areas including antimicrobial resistance and liver fibrosis. These developments can be viewed as components of a potential future self-improvement pipeline: Research → hypothesis generation → implementation → experimentation → evaluation → refinement → improved system. At present, humans remain essential throughout this process. AI systems can generate ideas and execute substantial technical work, but they still require reliable evaluation, human-defined objectives, appropriate infrastructure, and oversight. Consequently, the current state should be characterized as AI-assisted AI development rather than fully autonomous AI self-development. The distinction is important. The technology is advancing toward greater autonomy, but the complete automation of AI research and development has not yet been achieved.

#### ➤ *Expected Impact of the New Expansion*

The shift from AI-assisted tasks to AI-assisted AI development could have a major impact beyond the technology sector.

More capable AI could speed up software development, scientific research, and the creation of new technologies.

## II. ACCELERATION OF AI RESEARCH

AI systems could help researchers run experiments faster, test different ideas, find errors, and analyze results. This could make AI research more efficient and create a cycle where better AI helps develop even better AI.

#### ➤ *Transformation of Software Engineering*

Instead of manually implementing every component, engineers may increasingly define objectives, review AI-generated solutions, design system architectures, and validate results. The direction of current development already points toward longer-running coding agents capable of performing complex tasks with limited step-by-step human intervention.

#### ➤ *Acceleration of Scientific Discovery*

The potential impact extends beyond AI itself. AI systems capable of generating and testing scientific hypotheses could help researchers explore enormous search spaces in biology, chemistry, physics, materials science, and engineering.

#### ➤ *Economic and Workforce Effects*

As AI takes on more technical work, productivity may rise while the meaning of human work changes. Tasks once measured in hours could increasingly be measured in decisions, ideas, and oversight. The shift may not simply be

AI replacing people, but people working through increasingly capable AI systems. Jobs built around repetitive analysis, coding, and documentation could shrink, while roles requiring judgment, creativity, responsibility, and direction become more valuable. The likely result is not simply the disappearance of human work, but a redistribution of responsibilities. Human workers may increasingly focus on strategic decisions, creativity, communication, validation, management, and accountability.

## III. NEW SAFETY AND GOVERNANCE CHALLENGES

The benefits of increasingly autonomous AI systems must be considered alongside their risks. As AI takes a larger role in designing and improving AI, errors may no longer remain isolated—they could propagate through the development process at machine speed. The challenge is not only how fast AI can improve, but how safely that improvement can be contained and verified. This creates a need for stronger evaluation frameworks, controlled experimentation, monitoring, security measures, and human oversight. The central challenge will therefore be to increase AI's ability to improve systems without allowing the development process to become impossible to understand, evaluate, or control. The most important expansion may consequently be not simply greater autonomy, but the development of reliable mechanisms that allow humans to supervise increasingly autonomous AI systems.

## IV. SUMMARY

Artificial Intelligence is entering a new phase in which systems are increasingly capable of performing tasks that contribute to research, software engineering, experimentation, and scientific discovery. AI has begun to move beyond using intelligence toward helping engineer intelligence itself. Today's systems can write code, run experiments, analyze results, and support scientific research. The emerging possibility is a continuous loop in which AI identifies a problem, proposes a solution, tests it, learns from the outcome, and helps build the next version. If this loop becomes reliable, AI development could become faster, broader, and increasingly iterative. But greater capability also means greater responsibility: every step toward autonomous development must be matched by stronger verification, security, governance, and human judgment. The future challenge is therefore not simply to build smarter AI, but to ensure that the intelligence we build remains worthy of the responsibility we give it. The success of this transition will depend not only on computational power and algorithmic innovation, but also on our ability to establish appropriate scientific, ethical, and governance frameworks for increasingly autonomous systems.

## REFERENCES

- [1]. Starace, G., Jaffe, O., Aung, J., et al. (2025). *PaperBench: Evaluating AI's Ability to Replicate AI Research*. OpenAI. PaperBench — OpenAI

- [2]. Gottweis, J., Weng, W.-H., Daryin, A., et al. (2025). *Towards an AI co-scientist*. arXiv. Towards an AI co-scientist
- [3]. Google DeepMind. (2026). *Co-Scientist: A multi-agent AI partner to accelerate research*. Google DeepMind — CoScientist
- [4]. OpenAI. (2025). *MLE-bench: Evaluating machine learning agents on machine learning engineering*. MLE-bench — OpenAI
- [5]. OpenAI. (2025). *Introducing SWE-Lancer*. SWE-Lancer — OpenAI
- [6]. OpenAI. (2025). *Introducing GPT-5 for developers*. GPT-5 for Developers — OpenAI
- [7]. Anthropic. (2025). *Anthropic Economic Index: AI's impact on software development*. Anthropic Economic Index
- [8]. OpenAI. (2026). *Research acceleration: The view inside OpenAI*. Research Acceleration — OpenAI