Whitepaper

Securing AI Agents: The New Identity Challenge

AI agents are reshaping digital experiences—but securing them requires rethinking identity and access controls built for a human-first world.

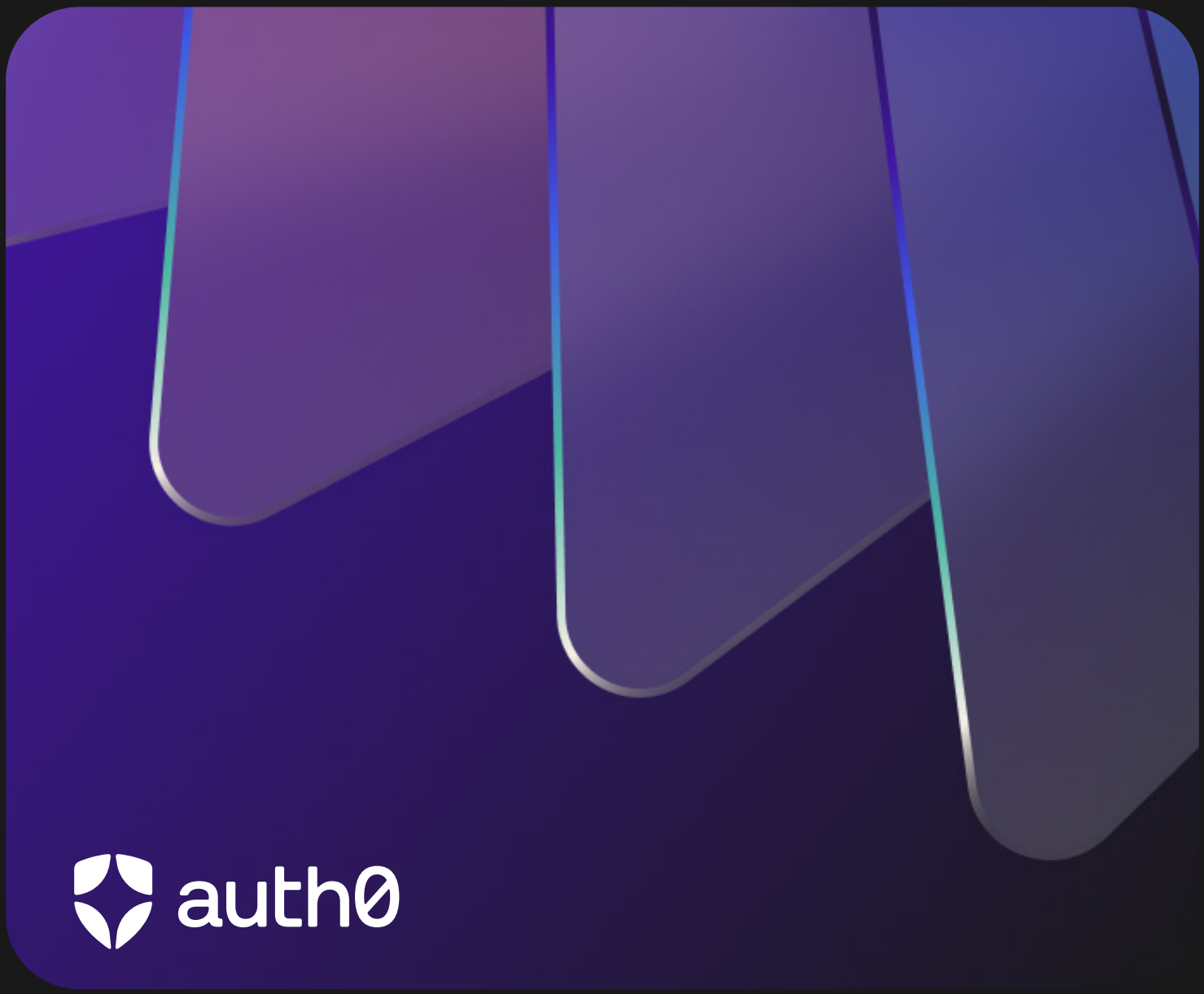


Table of contents

Let the markets speak for themselves

The global generative AI market is projected to grow from USD 67.18 billion in 2024 to USD 967.65 billion by 2032, exhibiting a compound annual growth rate (CAGR) of 39.6% during the forecast period.

Source: Fortune Business Insights, 2025

The AI agents market is expected to expand from USD 5.1 billion in 2024 to USD 47.1 billion by 2030, reflecting a CAGR of 44.8%.

Source: Fortune Business Insights, 2025

A survey by IBM reveals that 59% of enterprises already working with AI intend to accelerate and increase investment in the technology, reflecting a strong commitment to expanding AI capabilities.

Source: IBM, 2024

The Identity Security Gap in AI

The rapid adoption of AI agents has introduced new identity-related challenges. By 2027, 82% of organizations are expected to deploy AI agents, yet most security strategies remain focused on human authentication (source: Capgemini, 2024). This gap creates vulnerabilities that attackers can take advantage of. AI agents often rely on stale credentials, making them prime targets for credential theft, spoofing, and unauthorized access. Additionally, AI agents handle a lot of sensitive data, increasing the risk of exposure if access controls are not specific enough. AI agents could become a major attack space without robust security measures, costing customer trust and exposing businesses to legal risks.

Understanding AI Agents and Their Security Implications

What Are AI Agents?

AI agents are autonomous software systems that leverage large language models (LLMs), machine learning (ML), and APIs to perform tasks without direct human intervention. Unlike traditional software, AI agents can interpret and respond to natural language inputs, analyze real-time data, and take actions on behalf of users. These capabilities make them powerful business tools, but they also require a fundamental shift in identity security approaches. Traditional identity security frameworks were designed for human users, not autonomous AI software making independent decisions. As enterprises adopt AI agents at scale, they must rethink identity and access controls to optimize security without compromising user experience.

Why Enterprises Are Adopting AI Agents

Organizations across industries are integrating AI agents to improve efficiency, reduce costs, and enhance customer experiences. From AI chatbots handling customer service inquiries to decision-making agents analyzing business data, these agents streamline operations and unlock new capabilities.

There's already many areas where AI Agents are plugged into to help workflows and those use cases are only going to continue to expand as more enterprises become comfortable with using AI. Some of the common uses cases today include:

- **Customer Support Automation:** AI agents act as frontline support, answering common questions, resolving issues, and escalating complex cases to human agents. This helps with reducing both costs and response times, while keeping customers happy.
- **Sales and Marketing Acceleration:** AI Agents can qualify leads, draft personalized outreach, and assist with campaign execution based on real-time CRM data. They help teams move faster with existing headcount.
- **Internal Productivity and IT Support:** Within the enterprise, GenAI agents help employees reset passwords, request access, answer HR or policy questions, and even troubleshoot code—offloading routine tasks from IT and operations teams.

AI Agents' increasing autonomy presents a security challenge – non-human agents now require authentication, authorization, and governance. As AI adoption expands, different stakeholders within organizations have distinct concerns regarding security risks:

- **CTOs (Chief Technology Officers)** balance AI innovation with security, ensuring AI agents integrate safely into existing infrastructures without creating vulnerabilities.
- **CPOs (Chief Product Officers)** focus on maintaining frictionless customer experiences while ensuring that AI-driven interactions are secure and compliant.
- **CIOs (Chief Information Officers)** oversee regulatory compliance, risk management, and data governance to ensure AI agents adhere to industry standards and security best practices.

The Security Challenges of AI Agents

What is RAG?

RAG (Retrieval-Augmented Generation) is an AI method that improves answers by first searching a set of documents for relevant information, then using that data to generate a more accurate response.

Why is token vaulting important?

Auth for GenAI keeps user tokens safe in a secure vault and handles storage, refresh, and access so developers don't have to build their own token system.

AI systems and AI-powered applications are complex and exposed to different risks. Typically, a vulnerability in an AI system also affects the AI-powered applications that depend on it. There are a number of security challenges that everyone must be focused on when building AI agents:

1. First, **user authentication**. The agent or app needs to know who the user is. For example, a chatbot might need to display my chat history or know my age and country of residence to customize replies. This requires some form of identification, which can be done with authentication.
2. Second, **calling APIs on behalf of users**. AI agents connect to far more apps than a typical web application. As GenAI apps integrate with more products, calling APIs and storing them securely will be critical.
3. Third, **asynchronous workflows**. AI agents may need to take more time to complete tasks or wait for complex conditions to be met. It might be minutes or hours, but it could also be days. Users won't wait that long. These cases will become mainstream and will be implemented as asynchronous workflows, with agents running in the background. For these scenarios, humans will act as supervisors, approving or rejecting actions when away from a chatbot.
4. Fourth, Authorization for **Retrieval Augmented Generation (RAG).** Almost all GenAI apps can feed information from multiple systems to AI models in order to implement RAG. To avoid sensitive information disclosure, all data fed to AI models to respond or act on behalf of a user must be data the user has permission to access.

To realize GenAI's full potential, we must solve all four requirements. Whether you are building your own custom GenAI framework on top of a language like Python or using one of the many fast-growing frameworks that have emerged in the past two years, these requirements need to be addressed.

There has been no blueprint for building AI securely into applications. Developers at organizations have been figuring out DIY solutions to building the AI Agent itself.

Defend Your GenAI Apps from identity threats

That’s why we built Auth for GenAI. Auth for GenAI packages what we've learned working with GenAI frameworks and product builders and builds upon Auth0's decade of experience in identity.

With Auth for GenAI, you get:

Authentication for GenAI

Implement a tailor-made login experience for AI agents. This includes account linking of all accounts for the user profile and step-up authentication.

Token Vault

Use secure standards to connect AI agents to tools like Gmail and Slack using OAuth 2.0 for token management while also automatically handling token refreshes and exchanges.

Asynchronous Authorization

Enable AI agents to perform tasks with human-in-the-loop approvals.

Fine-Grained Authorization for RAGs

Protect sensitive data by only allowing AI Agents to retrieve documents that the user has access to.

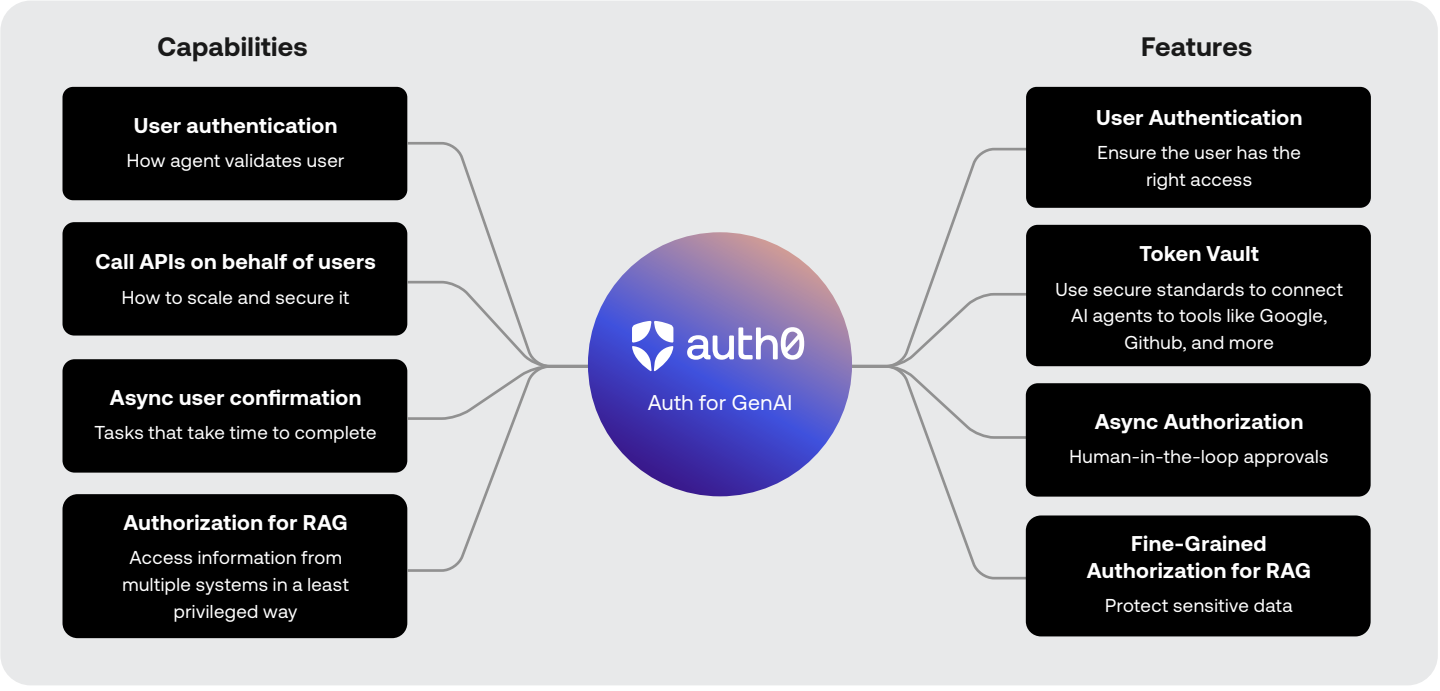
Auth for GenAI allows developers to achieve the same results with only a few lines of code, rather than spending countless hours coding to ensure their AI apps are secured.

A Strategic Path Forward

As AI adoption accelerates, security leaders must proactively address identity risks associated with AI agents. Organizations should:

1. **Evaluate existing AI security gaps** by auditing current access policies and authentication mechanisms.
2. **Implement AI-specific IAM solutions** to strengthen authentication, authorization, and monitoring controls.
3. **Adopt a continuous security posture** that includes real-time anomaly detection and automated response mechanisms.
4. **Support regulatory compliance** by aligning AI agent workflows with industry standards and governance policies.
5. **Invest in future-proof identity security** to support the evolving landscape of AI-driven automation.

AI is changing the way humans interact with technology and with each other. In the next decade, we will see the rise of a huge AI agent ecosystem—networks of interconnected AI programs that integrate into our applications and act autonomously for us. While GenAI has many positives, it also introduces significant security risks that must be considered when building AI applications. Enabling builders to securely integrate GenAI into their apps to make them AI and enterprise-ready is crucial.



In addition, as GenAI Agents begin operating across a growing number of apps and services, structured access to user context becomes essential. New standards like Anthropic's Model Context Protocol (MCP) provide a secure and standardized way for AI agents to retrieve context—such as calendar events, emails, or documents—while respecting privacy and permissions.

But context sharing introduces risk: if access isn't properly gated by identity and authorization controls, agents could expose sensitive data or act inappropriately. Auth0's implementation of an MCP server further facilitates secure management of authentication and authorization processes within this framework. AI builders must integrate fine-grained access checks into these new context flows so only the right information is shared with the right agent for the right task.

AI is a reality, for better or for worse. It brings countless benefits to users and builders, but at the same time, concerns and new challenges on the security side and all up throughout every organization.

With the Auth0 platform, Okta is here to help take the Identity piece off your plate. Learn more about building GenAI applications securely at auth0.com/ai.