SALT

Securing the Future of Agentic AI: Building Consumer Trust through Robust API Security

Table of Contents

01 Executive Summary

The integration of AI agents into websites and applications is accelerating, reshaping how businesses operate and how users interact with digital services. From chatbots and recommendation engines to automation bots - agentic AI has gone mainstream. More than half of organizations already using agentic AI say they're deploying it, or planning to, for customer interactions. On the consumer side, 64% report encountering AI chatbots more frequently than a year ago, and four out of five say they've shared personal details during these interactions.

But here's the problem: trust hasn't kept pace with adoption.

Half of consumers say they feel uncomfortable inputting personal information into a chatbot, and 44% have felt pressured to do so just to complete a task. This gap between usage and trust raises a key question: How can organizations harness the benefits of agentic AI while respecting - and earning - the trust of the consumers they serve?

What's often overlooked is that this AI revolution is powered by Application Programming Interfaces (APIs) that form the connective fabric of the AI ecosystem, enabling agents to retrieve information, trigger actions, and communicate seamlessly between systems. As agentic AI spreads, so too do the risks. While consumers notice the growing presence of AI and increasingly engage with it, concerns about security loom large. In fact, 62% believe chatbots are more vulnerable to manipulation by hackers than their human counterparts.

This report* explores the evolving architecture of AI agents, how both consumers and organizations are engaging with them, and the critical role APIs play in enabling their capabilities. It also examines the security challenges posed by this rapid evolution and offers best practices to help organizations protect their APIs, maintain consumer trust, and ensure the safe future of agentic AI.

*Methodology: Censuswide carried out a survey on behalf of Salt Security of 1000 US-based consumers and 250 organizations with 250+ employees who are already using agentic AI.

64%

of consumers report encountering AI chatbots more frequently than a year ago

02 Introduction

AI agents are software-driven entities designed to simulate intelligent behavior in tasks like customer service, content generation, and business automation.

Common examples include:

- Chatbots and virtual assistants
- Recommendation engines
- Task automation bots

These agents rely heavily on APIs to access internal databases, third-party services, and company systems, which enable them to function efficiently and dynamically. For example, a chatbot may use APIs to access a customer's order history or to initiate a return request.

Almost two-thirds (64%) of consumers noted an uptick in how much they encountered AI agents compared with the previous year and 81% of those who had encountered an AI chatbot said they entered personal details into it.

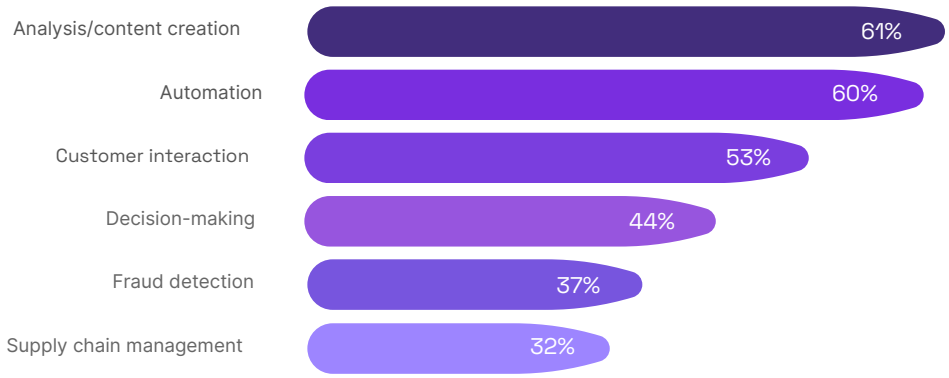
From a business perspective, the scope of AI operations is broad with nearly half (48%) of organizations that use agentic AI currently employing between 6 and 20 types of AI agents and 19% deploying between 21 and 50. A significant 37% of organizations report that 1–100 AI agents are currently active within their systems, while almost a fifth (18%) host between 501–1000 agents.

Businesses are using or planning to use agentic AI in a wide range of applications, from content creation and data analysis to customer interactions, fraud detection, supply chain optimization, and even internal process automation. As adoption deepens, agentic AI is increasingly embedded into core business functions, driving both efficiency and innovation.

81%

say they have entered personal details into an AI chatbot

In what ways does your business use (or plan to use) AI agents most?



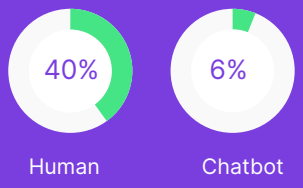
02 Introduction

What is the Value Proposition of AI agents?

Businesses are turning to AI agents to gain a competitive edge. These agents offer more than just automation, they enable round-the-clock service availability, hyper-personalized user experiences, and scalable operations that can adapt in real time to customer needs. By handling routine tasks and streamlining complex workflows, AI agents reduce operational costs while freeing up human teams to focus on higher-value activities.

But this value comes with a caveat: the growing reliance on APIs which power these AI capabilities also introduces new security challenges. As AI agents interact with sensitive data, initiate transactions, and make autonomous decisions, the infrastructure that supports them must be not only high-performing, but also resilient, secure, and trustworthy. Ensuring API integrity is no longer just a technical concern, it's a foundational requirement for protecting business value and maintaining consumer trust.

Who do you trust more with your personal data?



This disparity emphasizes a persistent trust gap between humans and machines, particularly when sensitive data is involved.

What Public Skepticism and Trust Issues are There?

Consumers still prefer human interaction for personal data sharing. Despite the growing presence and capabilities of AI agents, human interaction remains the preferred channel when it comes to sharing personal information. While 54% of consumers are comfortable disclosing personal details in-person and 37% by phone, only 22% feel comfortable doing so through a chatbot. This disparity underscores a persistent trust gap between humans and machines, particularly when sensitive data is involved.

The data also reveals a deeper tension: convenience often trumps caution. Nearly half (44%) of users admit to feeling pressured into providing personal information to chatbots, often just to complete a transaction or access a service. This suggests that while consumers may be wary of AI, they frequently go along with the interaction out of necessity, not confidence.

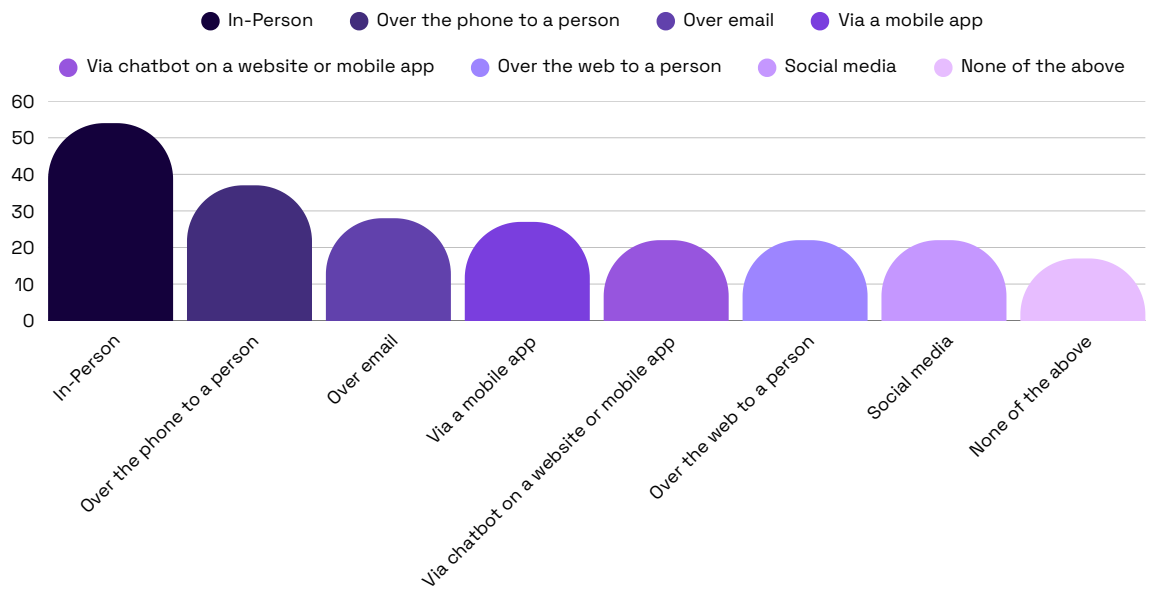
62%

Agree that chatbots can more easily be tricked by hackers to give up personal data than a human

02 Introduction

This dynamic has serious implications. It points to a fragile trust ecosystem in which users may comply with AI systems under duress or frustration, rather than feeling secure and in control. For organizations, this highlights the importance of designing transparent, ethical, and secure AI experiences, not only to meet user expectations but to avoid a long-term erosion of trust.

In which scenarios, if any are you/would you be comfortable inputting or communicating you personal data?



03 The Role of APIs in AI Agent Functionality

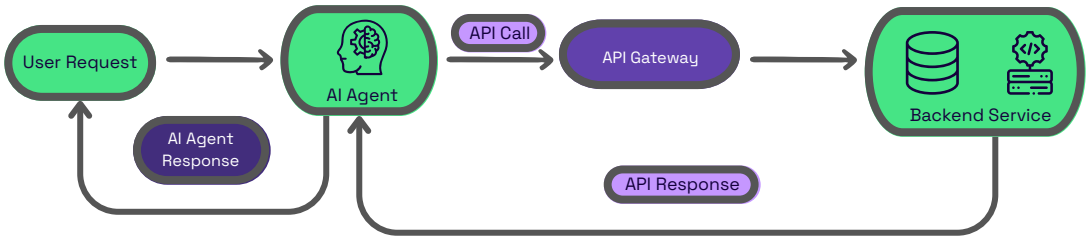
APIs enable AI agents to access structured data from both internal systems and external services. For example, an AI-powered customer support agent might use APIs to retrieve a customer's purchase history from a backend database in real time, allowing it to provide more relevant and informed responses. Similarly, APIs allow AI agents to take action within enterprise platforms such as CRMs, ERPs, or e-commerce systems triggering workflows like updating records, generating invoices, or processing returns.

Beyond simple data retrieval or task execution, APIs also allow AI agents to orchestrate complex, multi-step processes. Consider a logistics bot that coordinates a delivery: it may check warehouse inventory, schedule a pickup with a shipping carrier via their API, and send a confirmation to the customer - all without human intervention. In the financial sector, an AI advisor might pull real-time market data from third-party APIs, analyze the results, and generate investment recommendations tailored to an individual's portfolio. Each API serves as a "bridge," enabling the AI to function intelligently and autonomously.

In essence, APIs transform AI agents from isolated algorithms into powerful digital collaborators capable of navigating and acting within interconnected systems. This deep integration is what gives agentic AI its utility, but it also highlights why securing those APIs is so critical. If the interfaces that fuel AI are vulnerable, the agents themselves become liabilities rather than assets.

"APIs allow AI agents to take action within enterprise platforms such as CRMs, ERPs, or e-commerce systems triggering workflows like updating records, generating invoices, or processing returns."

AI Agent API Interaction Flow



04 API Security Challenges in the Context of AI Agents

As AI agents become more embedded across digital systems, they dramatically increase the number of API interactions occurring within and between platforms. This proliferation creates a vastly expanded attack surface, as each new API connection introduces a potential entry point for threat actors seeking to exploit vulnerabilities. Unlike traditional applications, AI agents often operate autonomously and at scale, amplifying the impact of any single security lapse.

One of the key challenges lies in authentication and privilege management. AI agents frequently require access to multiple systems and datasets, and are often granted elevated permissions to perform their functions efficiently. However, if their credentials are compromised, whether through phishing, poor storage practices, or insecure token handling, the consequences can be severe. A single set of stolen credentials could enable an attacker to access sensitive data, execute fraudulent transactions, or manipulate business operations without immediate detection.

“A single set of stolen credentials could enable an attacker to access sensitive data, execute fraudulent transactions, or manipulate business operations without immediate detection.”

Another emerging risk is the potential for prompt injection attacks, where malicious inputs are crafted to manipulate the AI’s behavior or bypass safety controls. When these inputs are processed through APIs that feed data directly into agent workflows, the danger increases, especially if input validation and sanitization are inadequate.

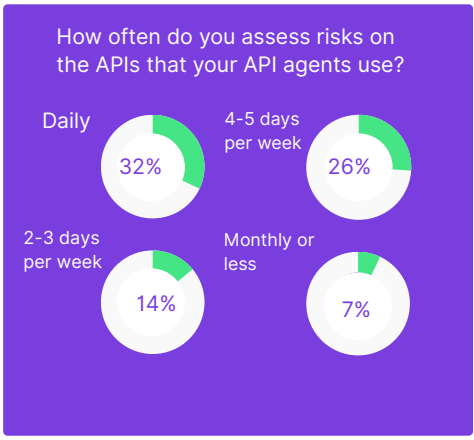
Monitoring and detection also become significantly more complex in environments where AI agents generate a high volume of API traffic. Distinguishing “normal” behavior from anomalies is far more challenging when millions of requests are processed dynamically, especially when those patterns evolve in real time. Traditional security tools may struggle to keep pace without adaptive, AI-driven monitoring solutions.

Additionally, APIs are frequently the weakest link in otherwise secure environments. Poorly configured endpoints, insufficient access controls, lack of input validation, and vulnerabilities in third-party APIs all present opportunities for exploitation. Even with strong perimeter defenses, an overlooked API can provide a direct pathway into critical systems.

04 API Security Challenges in the Context of AI Agents

Survey data underscores the complexity of these challenges. While a notable percentage of organizations (32%) conduct API risk assessments daily and another 26% do so several times a week, a concerning minority (7%) report performing them monthly or even less frequently. This inconsistency presents discrepancies between awareness and action. In a world where AI agents are operating continuously and autonomously, frequent, systematic API security assessments must become the norm, not the exception. Without them, organizations risk leaving their most advanced digital capabilities exposed.

Organizations must recognize that frequent, consistent API security assessments are
essential in mitigating AI-related risks.



How Are Privacy Concerns Addressed by Businesses?

As AI agents become more deeply integrated into business operations, addressing privacy concerns is critical, given the clear hesitancy from consumers to interact with it. Encouragingly, businesses appear to be taking a multifaceted approach to managing these concerns, though practices vary in maturity and range.

Survey results show that organizations are adopting several core strategies to mitigate privacy risks associated with AI. The most common approach, cited by 44% of respondents, involves actively monitoring AI decision-making to detect and prevent unintended privacy violations. This reflects a growing awareness that AI systems can generate outcomes that may conflict with privacy expectations, even when those outcomes are not explicitly programmed. Real-time oversight is seen as essential for spotting these issues early and ensuring that AI systems behave within acceptable boundaries.

04 API Security Challenges in the Context of AI Agents

How, if in any way, do you address data privacy concerns related to AI usage within the business?

We monitor AI decision-making to prevent unintended privacy violations	44.40%
We conduct regular audits to ensure compliance with data privacy laws	43.20%
We provide clear privacy policies that inform users how their data is used	42.40%
We use AI governance frameworks to align with ethical and legal requirements	42.00%
We encrypt all data both in transit and at rest to prevent unauthorized access	39.60%
We implement strict access controls to limit who can interact with sensitive data	37.60%
We conduct AI explainability testing to ensure transparency in automated decisions	37.60%
We use a dedicated API security solution	37.20%
We have a dedicated data privacy team overseeing AI projects	37.20%
We conduct regular penetration testing to identify and address AI security vulnerabilities	36.80%
We offer user control options, such as opt-outs or data deletion requests	36.40%
We follow global privacy regulations (e.g., GDPR, CCPA, HIPAA, etc.)	35.20%
We anonymize or pseudonymize data before using it for AI training	29.60%
We don't address data privacy concerns related to AI usage within the business in any way in particular	1.20%

04 API Security Challenges in the Context of AI Agents

How Are Privacy Concerns Addressed by Businesses?

Close behind, 43% of businesses report conducting regular audits to ensure compliance with data protection regulations such as GDPR, CCPA, and other global privacy laws. These audits are critical for verifying that data is collected, processed, and stored appropriately, especially in dynamic environments where AI agents are continuously learning and adapting.

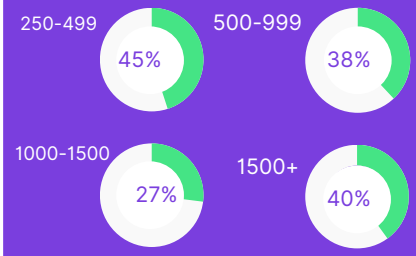
Transparency also plays a central role in privacy efforts. About 42% of organizations say they provide clear privacy policies to inform users about how their data is being used. While this is a baseline requirement in many jurisdictions, making these policies understandable and accessible is key to building trust, particularly when AI is involved in decision-making or personalization.

n parallel, 42% of respondents noted the use of AI governance frameworks to ensure alignment with both legal and ethical standards. These frameworks typically include oversight mechanisms, internal review boards, and guidelines for responsible AI use, factors that are becoming increasingly important as regulators and the public demand greater accountability from AI systems.

Interestingly, only 37% of organizations report using a dedicated API security solution, demonstrating the recognition that data privacy is closely linked to how information flows through systems. However, there is clear room for improvement, since APIs serve as the primary channels through which AI agents access and exchange data, meaning that securing them is a fundamental privacy safeguard. Interestingly, organizations with 250-499 employees were more likely to have a dedicated API security solution (45%) than the largest organizations.

An equal percentage (37%) also state they have a dedicated data privacy team overseeing AI initiatives. Perhaps unsurprisingly this number rises to 46% in the largest organizations of more than 1,500 employees. This signals a more mature governance approach, where privacy isn't treated as an afterthought or delegated to general IT functions, but is instead actively managed by experts who understand both legal requirements and the unique risks posed by AI.

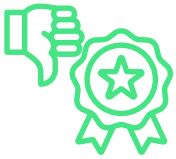
Organizations that use a dedicated API security solution by number of employees



Taken together, these responses suggest that while organizations are broadly aware of privacy risks, there's no single silver bullet. Instead, the most effective strategies appear to be those that combine technical controls, organizational structures, and policy transparency, a layered defense model that adapts to the changing AI landscape.

04 API Security Challenges in the Context of AI Agents

Data Exposure Unauthorized access to personal or financial records



Reputational Damage Loss of consumer trust and customer churn

Financial Loss Regulatory fines and incident response costs

Operational Disruption Compromised AI agents can disrupt services or output faulty data

While not always explicitly linked to APIs, these exploits often involve poorly secured backend connections or prompt manipulation, both of which can be API-driven.

05 Best Practices for Securing APIs Used by AI Agents

As AI agents become more integrated into business operations and consumer experiences, the APIs that power them become critical points of vulnerability. Without strong security measures, these APIs can be exploited which can potentially undermine user trust and expose sensitive data. The following table outlines key best practices organizations should adopt to secure APIs used by AI agents, ensuring safe, reliable, and trustworthy deployments across digital environments.

Category	Best Practices
Monitoring and Detection	• Monitor all API traffic for anomalies. • Log every interaction for forensic analysis. • Use AI-based threat detection tools to flag suspicious activity.
Authentication and Access Control	• Use OAuth 2.0 or multi-factor authentication. • Enforce least privilege access for AI agents. • Secure credential management (e.g., rotate and protect API keys).
Input Sanitization	• Validate all input data formats. • Sanitize user inputs to prevent prompt injection and data poisoning.
Governance and Developer Training	• Use an API gateway to enforce policies. • Train developers on secure coding practices. • Establish an internal security review for every new AI integration.
Performance Throttling and DoS Protection	• Apply rate limits and quotas per endpoint or user. • Detect and block traffic spikes indicative of a DoS attack.
Encryption	• Use TLS/HTTPS for data in transit. • Apply encryption at rest for sensitive data. • Implement masking for confidential information shared across APIs.
Security Testing	• Conduct regular penetration testing. • Automate security scanning within CI/CD pipelines. • Run fuzz tests to detect edge-case vulnerabilities.

06 Emerging Trends and Future Challenges

The future of AI agent security is closely tied to innovation in API security. As AI agents become more intelligent, autonomous, and deeply embedded in digital ecosystems, their security cannot be treated in isolation. The APIs that underpin their functionality are evolving into critical control points, not just for enabling capabilities, but also for defending against increasingly sophisticated threats. The future of AI agent security will be defined by how well organizations can anticipate and respond to the shifting dynamics of API risk.

"The future of AI agent security will be defined by how well organizations can anticipate and respond to the shifting dynamics of API risk."

Emerging threats are already outpacing traditional defenses. AI-powered attacks which are driven by adversarial machine learning, can adapt and evolve to exploit weaknesses that signature-based or static rule systems simply cannot detect.

These attacks may use automated reconnaissance, mimic legitimate user behavior, or even manipulate prompts and data inputs to distort AI outputs. As these techniques grow more refined, they blur the line between valid and malicious interactions, making it harder for security systems to intervene.

At the same time, architectural shifts are adding layers of complexity.

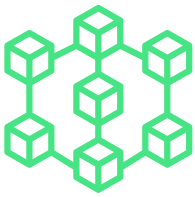
Decentralized infrastructures, including microservices and serverless environments, have introduced agility and scalability, but at the cost of visibility. In these architectures, an AI agent's logic may be distributed across multiple ephemeral services, complicating traditional approaches to logging, auditing, and policy enforcement. Traceability becomes a challenge, especially when APIs span hybrid or multi-cloud environments.

Moreover, vulnerabilities unique to the AI context, such as prompt injection, model inversion, and data poisoning, are likely to increase as generative AI becomes more prevalent. These are not hypothetical risks; they exploit the very interfaces AI agents rely on to learn, generate, and act. Addressing them requires both new thinking and new tools.

Fortunately, innovation in API security is rising to meet these challenges.

07 Innovation in API Security

AI-driven security tools apply machine learning to monitor API behavior continuously. These tools can detect anomalies that deviate from historical baselines, flag suspicious patterns in real time, and even take automated action to limit exposure.



Blockchain is also emerging as a powerful solution for enhancing integrity and accountability in API ecosystems. By creating tamper-proof audit logs of every API call, blockchain-based systems can ensure that transactions are both traceable and verifiable, critical features in a landscape where trust is easily undermined.

Homomorphic encryption enables data to be processed without ever being decrypted. This technique holds enormous potential for AI applications that need to analyze sensitive information, such as financial or healthcare data, without risking exposure. If integrated effectively into API workflows, it could allow for privacy-preserving AI operations at scale.

Ultimately, the security of tomorrow's AI agents will depend on how proactively organizations can reimagine their API strategies. This means moving beyond patching vulnerabilities as they arise and toward building strong governance for resilient, intelligent, and adaptive API environments where security is not an obstacle to innovation, but an enabler.

08 Conclusion

AI agents are rapidly reshaping how businesses interact with customers, automate operations, and deliver services. Yet, their success hinges on one foundational element: secure, well-governed APIs. These interfaces are not just technical connectors, they provide the lifelines through which AI agents access data, execute tasks, and integrate across platforms. Without robust API security, even the most advanced AI becomes a vulnerability rather than an asset.

As this report has shown, consumers are engaging with AI more frequently but remain wary, especially when personal data is involved. Organizations, meanwhile, are eager to expand their use of agentic AI but must navigate a landscape filled with evolving threats and compliance pressures. The only sustainable path forward is a proactive approach to API security anchored by governance, encryption, access control, and intelligent monitoring.

Effective API governance is as much about building trust as it is about maintaining security. By embedding security into every layer of the AI ecosystem, businesses can protect their data, maintain user confidence, and build a resilient foundation for innovation. Without secure APIs, AI systems cannot function safely. Strong API security builds the trust and confidence needed for widespread AI adoption.

To learn more, visit: https://salt.security