Forbes INDIA

How AI hallucinations endanger software supply chain

AI coding tools boost productivity but also introduce hallucinations—false APIs, insecure settings, and fake dependencies—that can trigger compliance failures, cyberattacks, and reputational damage. B ... **Read More**

By IIM Calcutta

Last Updated: Aug 28, 2025, 12:51 IST 9 min

AI tools are becoming more common in software development at an amazing rate. They allow code to be written in seconds and systems to be built on demand. However, there is a growing risk associated with this progress: AI hallucinations. These aren't minor bugs; they're sure lies about APIs that don't exist, insecure configurations, and libraries that don't exist. In today's interconnected software supply chains, these kinds of hallucinations can lead to significant problems, such as security breaches, compliance failures, and damage to your reputation. As AI becomes more prevalent in development processes, companies must strike a balance between being cautious and innovative. The real value of AI isn't in unquestioningly trusting it; it's in keeping a close eye on it with clear rules, strict human review, and a mindset that views AI output as a suggestion, not a solution. In this changing world, those who are cautious will be the ones who succeed, not those who fail to question.

RELATED STORIES

1 **How AI is changing India's IT services industry**
Harichandan Arakali

2 **AI, humans and India's role in this tech revolution**
Shobhit Jain

3 **All things AI: How your life will change in 5 years**
Neha Bothra

The Mechanism of AI Hallucinations in Software Supply Chains

AI tools like GitHub Copilot, Amazon CodeWhisperer, and ChatGPT are now integral to the software development process. They help developers write code, suggest libraries, plan system architectures, and even write cloud infrastructure and security policies. These tools can greatly increase productivity, but they also pose a subtle but serious threat: AI hallucinations, which are false or misleading suggestions given with that aren't safe.

Dependency management is a hazardous area. An AI might, for example, think of a package name like requests-proxy. Then, a malicious person could upload that exact package to PyPI, including malware that infiltrates business environments through automated builds. This is a classic example of typo squatting or dependency confusion. These aren't just examples; packages like ua-parser-js and event-stream have already been attacked in the same way.
AI-generated hallucinations can also include infrastructure-as-code and security policies. For example, an AI might suggest AWS IAM permissions that are too broad (such as *:* access) or Kubernetes RBAC roles that grant admins excessive rights. Because these hallucinations are presented with such confidence, developers may use them without careful consideration, leaving systems vulnerable to attack. The line between convenience and being open to attack is now thinner than ever.

Business Implications of AI Hallucinations in the Supply Chain

AI hallucinations pose not only technical risks but also threats to business continuity, compliance with regulations, brand reputation, and financial stability. This new threat has several significant effects on businesses, which are outlined below.

1. **Increased Exposure to Cybersecurity Breaches -** AI hallucinations can put insecure code, misconfigured settings, or harmful dependencies directly into software systems, which makes an organisation's attack surface much bigger. If an attacker releases a fake version of a package that an AI tool recommends, it could compromise the supply chain. Similarly, security settings that don't make sense, such as IAM roles that are too permissive, can grant attackers access they don't need. These kinds of breaches have effects that go beyond fixing the problem. If an exploit is successful, it can result in data theft, intellectual property theft, service outages, and ransomware attacks. These results damage customer trust, complicate operations, and incur significant costs to rectify. It's also important to remember that being publicly linked

2. **Erosion of Trust in Digital Products -** In any business relationship, trust is crucial, but it's particularly critical for organisations that produce software or provide digital services. If AI hallucinations are used to exploit vulnerabilities, customers may lose trust in the company's ability to safeguard their data or ensure that their services are always available. This loss of trust directly leads to reduced sales, fewer clients, and difficulties in acquiring new business. Customers and investors may lose faith in a production system simply because it uses shoddy or insecure components, even if a vulnerability is discovered before it can be exploited. Businesses may also be examined more closely during vendor evaluations and security audits.
3. **Regulatory and Compliance Risks -** Strict rules govern industries like finance, healthcare, and critical infrastructure. Some examples are GDPR, HIPAA, PCI DSS, SOC 2, and ISO 27001. These frameworks require safe methods for software development, such as safe supply chain management. AI hallucinations that create weaknesses can cause non-compliance, which can lead to fines, penalties, or legal action. Regulators are also paying more attention to AI governance itself. Companies that don't manage the risks of AI hallucinations could run afoul of new AI-specific rules that require AI tools to be open, accountable, and able to handle risks.
4. **Financial Consequences -** Hallucination-induced security incidents can have a huge effect on a company's bottom line. Costs include fixing the breach immediately (forensics, legal fees, and communications), paying fines to the government, settling lawsuits, and running campaigns to repair your reputation. Indirect costs like lost revenue, lower market value, and higher insurance premiums also cause long-term financial harm. Think about how much it would cost to redesign systems that have AI-generated parts that aren't safe. Companies may also have to wait longer to release products while they eliminate risky dependencies or review decisions made with AI.
5. **Intellectual Property Risks -** AI hallucinations can make code snippets from training data that don't have clear licensing. Developers might add code that breaks open-source licenses, which don't align with the company's usage policies or business

very important. If you break a license, you may have to open-source your own code, go to court, or re-engineer your products, which is a time-consuming and costly process.

6. **Supply Chain Contagion -** Hallucinations often create vulnerabilities that don't stay isolated. When insecure or fake parts are added to open-source projects, they can spread quickly through dependency trees. Companies that use these projects without being aware of them take on these risks. This "contagion effect" increases the possible blast radius. An organisation could be responsible not only for its own safety but also for the safety of its partners and customers. In the worst cases, this could lead to legal claims from third parties who were hurt or the end of strategic partnerships.

Mitigation Strategies for Businesses

Businesses need to take a proactive and organised approach to protect their software supply chains from the effects of AI hallucinations, given these risks. The following strategies employ a combination of technical controls, process enhancements, and behavioural changes to ensure that software supply chains are secure from vulnerabilities.

1. **Reinforce Human-in-the-Loop Oversight -** Artificial intelligence tools should help people make decisions, not take their place. Establish processes that require all code and configurations generated by AI to be reviewed by a person. This is especially important for things that have to do with security. Developers and reviewers should verify that suggested APIs exist, that package recommendations are accurate, and that security settings align with trusted documentation. Companies can create review checklists that focus solely on identifying AI-related risks, ensuring that hallucinations are not overlooked.

2. **Adopt and Enforce Software Bill of Materials (SBOM) Practices -** SBOMs let businesses see the parts of their software, such as libraries, dependencies, and licenses. Companies can find out where each part came from, figure out where hallucinated suggestions may have added risks, and respond quickly to

can make sure that no unknown or unauthorised components get into production. This openness is also beneficial for audits, customer trust, and compliance with regulations.

3. **Strengthen Dependency Management Controls -** Organisations should use automated tools that make sure dependencies are kept clean. Some of these tools include OWASP Dependency-Check, Snyk, and Dependabot, which can identify security holes and unauthorised changes. These groups should also use cryptographic signing and verification of packages and keep lists of approved repositories and packages. Regular audits of dependencies ensure that the organisation doesn't unknowingly acquire components that are fake or harmful. To avoid confusion about dependencies, it's also essential to keep an eye out for typo-squatting attacks on package registries.

4. **Enhance AI Usage Governance -** To effectively manage the use of AI in software development, organisations should set up formal governance structures. This involves clearly defining what AI can and can't do, documenting the steps for reviewing and utilising AI outputs, and verifying changes made with the assistance of AI. To ensure that these rules are followed and that people are held accountable, there should be specific individuals in charge, such as AI risk officers or security champions. Governance rules must make it clear how AI tools are used and hold people accountable for any choices made based on AI-generated content. To keep development teams on the same page, aware, and ready to deal with AI-related risks, they need to get regular training and talk about these policies.

5. **Promote Developer Education and Awareness -** Security awareness programs need to change to deal with risks that are unique to AI. This will ensure that developers know how to identify and address potential problems. Training should include real-life examples of AI hallucinations, hands-on ways to check AI-generated outputs, and instructions on how to spot and report suspicious package recommendations or misconfigurations. Additionally, developers must be aware of the licensing issues that arise when using AI-generated code, particularly when working with open-source content. Organisations can prevent people from unquestioningly trusting AI outputs

6. **Integrate Security Automation and Monitoring -** To find security holes caused by AI hallucinations before they get to production, it's important to include security automation in CI/CD pipelines. This involves utilising static application security testing (SAST) and dynamic application security testing (DAST) to examine how code functions and its performance. Infrastructure-as-code (IaC) scanning can identify errors in deployment templates, and secret detection tools can uncover hardcoded credentials, a mistake that AI suggestions often make. Continuous monitoring of production environments for anomalies provides an additional layer of defence, catching security issues that may have slipped through earlier checks and making the software development lifecycle more resilient.

7. **Monitor and Respond to Ecosystem Manipulation -** Organisations should keep an eye on communities like GitHub, Stack Overflow, and package registries for signs of fake content that is meant to mess up AI training datasets. If possible, participate in industry efforts to protect open-source ecosystems and report any suspicious behaviour. Taking steps to maintain a healthy ecosystem reduces the likelihood of receiving fake outputs that attackers have planted.

8. **Legal and Compliance Readiness -** Legal teams should work with engineering and security to make sure there are clear rules about how to use AI, intellectual property, and licensing. Contracts with vendors and partners should spell out how AI can be used, SBOM requirements, and security practices. Keeping detailed records of AI-assisted decisions helps with legal defences and ensures compliance with regulations. Getting ready for possible audits now lowers your risk later.

Authors: *Ranjan Pal (MIT Sloan School of Management, USA), Douglas Granillo (MIT EECS), Bodhibrata Nag (Indian Institute of Management Calcutta)*

This article has been published with permission from IIM Calcutta. ***www.iimcal.ac.in*** *Views expressed are personal.*

First Published: Aug 28, 2025, 12:47

Artificial Intelligence (AI)

Home / **Thought-leadership** / **Iim-calcutta** / How-ai-hallucinations-endanger-software-supply-chain

FORBES LIST

AI LIST 2025

TogetherAI's Vipul Ved Prakash: Democratising AI access with open-source solutions

Naini Thaker

Creating an army of AI employees: Surojit Chatterjee

Naini Thaker

Prof Sunita Sarawagi: Helping machines make sense

Samidha Jain

Explore All List

LATEST NEWS