SPOTLIGHT
Sponsored by: OpenText

Generative AI is revolutionizing application security by enhancing cybersecurity tools. However, this technology is also introducing new security risks, necessitating that organizations adopt innovative solutions that leverage its strengths while mitigating its vulnerabilities.

The Peril and Promise of Generative AI in Application Security

July 2024

Written by: Katie Norton, Research Manager, DevSecOps and Software Supply Chain Security

Introduction

In the rapidly evolving landscape of cybersecurity, generative AI (GenAI) has emerged as a transformative force, bringing with it new challenges as well as opportunities for application security. Since it burst into public consciousness in late 2022, GenAI has captivated organizations worldwide, promising to revolutionize efficiency and productivity across various sectors. IDC research has found that as of April 2024, 17% of organizations have already introduced GenAI applications or services into production. Another 38% are investing significantly with an established spending plan for GenAI-enhanced software (source: IDC's *Future Enterprise Resiliency and Spending Survey, Wave 4,* April 2024).

The rapid adoption of this technology means that attacks on GenAI systems are becoming more sophisticated and frequent, presenting a formidable threat vector. With their heavy reliance on large, often sensitive data sets for training, GenAI applications will become prime targets for data breaches. The significance of this threat vector is magnified as GenAI applications become integral to critical systems and applications in domains such as healthcare and automotives.

Application security in the GenAI era is marked by both promise and peril. On one hand, GenAI offers unprecedented opportunities to enhance cybersecurity tools and processes, making them more robust and intelligent. On the other hand, it introduces novel security risks that demand innovative solutions. To navigate this complex landscape, a new paradigm is required — one that ensures AI applications' security while simultaneously leveraging GenAI tools to bolster application security efforts.

AT A GLANCE

KEY TAKEAWAYS

- GenAI has emerged as a transformative force in application security. However, it has a dual nature in that it needs to be not only secured but also leveraged to enhance security tools and practices.
- GenAI applications have both a supply chain to be secured and distinct vulnerabilities, requiring the adoption of application security tools specifically designed to address these challenges.
- Incorporating GenAI into application security processes offers the potential to significantly enhance cybersecurity defenses, making applications more secure and resilient. GenAI can augment traditional practices such as threat modeling, vulnerability explanation, code review, and security testing, providing a more adaptive and comprehensive approach to application security.

Securing GenAI Applications

IDC predicts that by 2026, 40% of net-new applications will be intelligent, incorporating AI to enhance user experiences and create novel use cases. The cornerstone of these intelligent applications lies in machine learning (ML) models. These models, capable of processing vast volumes of data and learning from it, are indispensable in today's digital economy, driving performance, functionality, and innovation. As businesses and technologies evolve, the integration of ML models as foundational components deepens, marking a significant shift in how applications must be developed and secured.

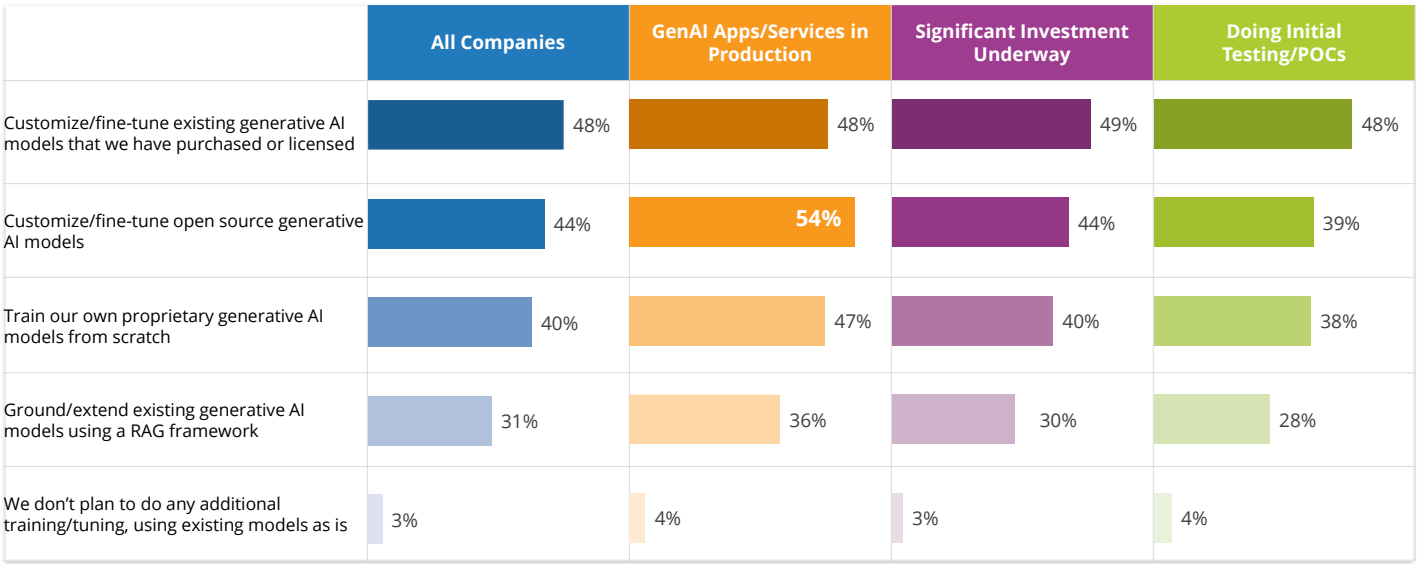
Open Source and Third-Party Models

GenAI applications rely on a complex network of data sources, pretrained models, libraries, and other components, some of which are often untraceable. Due to its complexity and lack of transparency, the supply chain for GenAI applications presents malicious actors with a vast attack surface.

In addition to open source software components, GenAI applications also often leverage open source models and training data downloaded from public repositories. As shown in Figure 1, IDC research finds that 44% of organizations are customizing or fine-tuning open source GenAI models, with an even higher rate of adoption at 54% for the subset of organizations that have GenAI applications or services in production.

FIGURE 1: ***Organizations Are Using Open Source Models***

Q Which of the following approaches are you using/evaluating for training or tuning of GenAI models?



n = 889

Source: IDC's Future Enterprise Resiliency and Spending Survey, Wave 4, April 2024

While fostering innovation and enabling more rapid deployment of GenAI applications, open source models also expose organizations to risks. Attackers have the same access to these online repositories and can deploy a backdoor or malware

into them. Once uploaded back into the repository, they can become an entry point to anyone that downloads the infected model.

A hallmark of a supply chain attack is its blast radius. A successful attack against one open source model or training data set can enable the attacker to compromise multiple GenAI applications. For instance, in May 2024, the llama-cpp-python package was found vulnerable to server-side template injection. This susceptibility could lead to remote code execution and the prospect of full system compromise, data breaches, and other malicious activities. The flaw, tracked as CVE-2024-34359, is a popular Python library for LLMs and was in use by more than 6,000 models on Hugging Face. This incident highlights the critical need for securing the AI supply chain, including practices for ensuring the integrity of open source models.

Several important practices exist for safeguarding against AI software supply chain vulnerabilities. These include threat modeling any new open source model incorporated into an application, creating an AI Bill of Materials (AIBOM) to manage all the components and dependencies of a model, managing and monitoring those dependencies to ensure all components are up to date and free of known vulnerabilities, tracking the provenance of all model artifacts, and using cryptographic signatures to verify the model's authenticity.

Distinct Security Vulnerabilities

GenAI applications have ushered in a new frontier of technological complexity, revealing novel risks that challenge traditional security paradigms. GenAI models operate as black boxes and exhibit highly dynamic behavior. Traditional security tools often rely on understanding the application's logic to detect anomalies or vulnerabilities, which is challenging with opaque AI models.

The attack surface for GenAI applications includes not just the traditional components (e.g., web servers, databases) but also the ML models themselves, which can be manipulated through adversarial attacks. Some of the most critical vulnerabilities unique to GenAI applications lie in their data sets and learning algorithms. Data poisoning, for instance, is a tactic where attackers deliberately "pollute" the data on which the model is being trained, tampering with its decision-making capabilities. As a result of these modifications to the training data, the AI system becomes biased toward the attacker's objective, enabling them to harvest sensitive data or corrupt the model's output.

Another potential vulnerability, prompt injection, manipulates the ML model via inputs, forcing it to execute instructions beyond its intended purpose. In 2022, a remote work company created an X bot, running on the GPT-3 language model by OpenAI, that would respond positively to tweets about remote work. By redirecting the bot with phrases like "ignore the above," users could make it repeat embarrassing or ridiculous text rather than the commentary on remote work for which it was designed.

Prompt injection is analogous to SQL injection or cross-site scripting in web applications. However, it is more challenging to defend against because it takes advantage of how GenAI applications use natural language to consider both instructions and data as the user inputs them. Limiting user inputs or outputs can impede the functionality that makes them useful.

Both prompt injection and data poisoning are included in the Open Web Application Security Project's (OWASP) Top 10 for LLMs and Generative AI Apps. This list aims to bridge the gap between general application security principles and the specific challenges posed by LLMs. It provides practical and actionable guidance to developers, data scientists, and security professionals regarding GenAI application security.

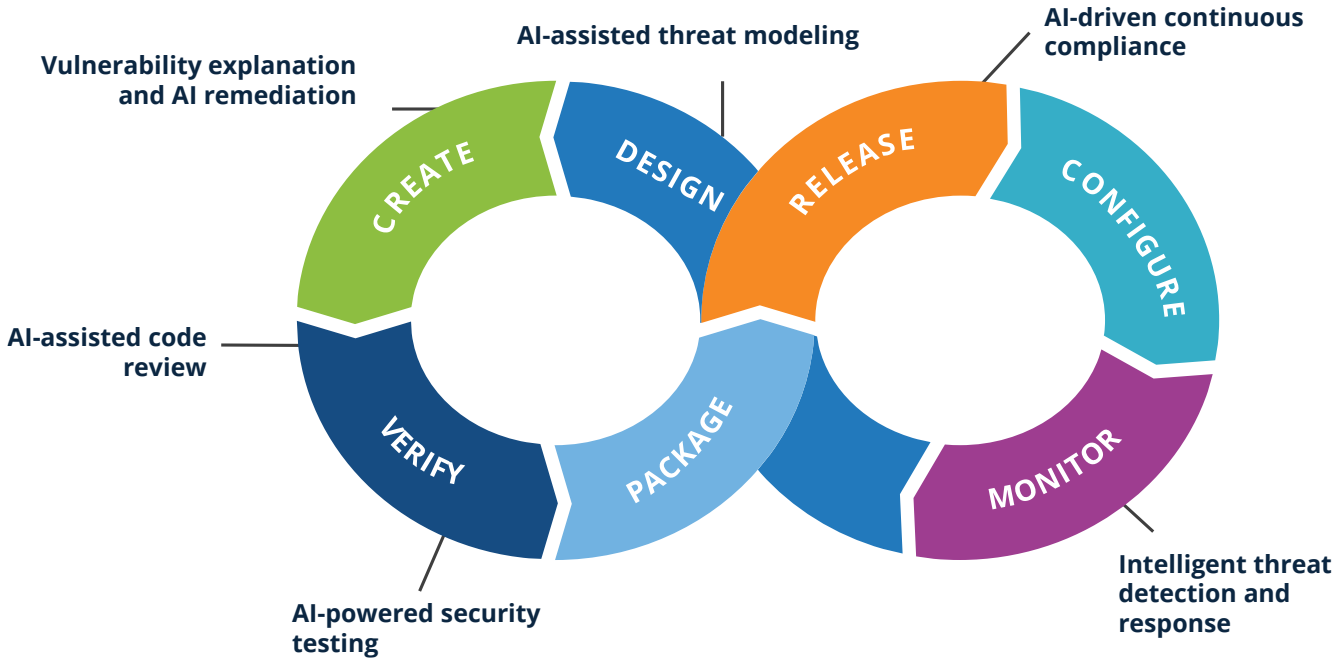
Organizations need to adopt application security tools designed to address the distinct vulnerabilities associated with GenAI applications to ensure application integrity and reliability. These tools should identify code patterns that make the application susceptible to malicious inputs crafted to manipulate the AI model's behavior. Also important is the ability to recognize and understand AI and ML libraries and frameworks (e.g., TensorFlow, PyTorch). Finally, the tools should support compliance with AI-related industry standards and regulations, providing audit trails and documentation to demonstrate adherence to security and privacy laws.

Enhancing Application Security with GenAI

Incorporating GenAI into application security processes and tools can result in more secure, efficient, and resilient applications while freeing developers and security professionals to focus on more complex and challenging problems.

Figure 2 demonstrates multiple use cases for GenAI application security across the software development life cycle (SDLC). The shift-left use cases that are most impactful for developers are detailed in the sections that follow.

FIGURE 2: ***GenAI-Enhanced Application Security Throughout the SDLC***



Source: IDC, 2024

AI-Assisted Threat Modeling

By analyzing vast data sets of historical security incidents, GenAI can identify potential threats and vulnerabilities, generating sophisticated threat models and assisting in comprehensive risk assessments. This capability enables organizations to proactively mitigate weak points in applications, anticipating attack vectors before they can be exploited.

Vulnerability Explanation and AI Remediation

GenAI can be leveraged to help developers understand a security vulnerability through contextual explanations of the nature of the vulnerability, how it can be exploited, and the potential impact on the application. GenAI can also analyze code in real time as developers write it, identifying security vulnerabilities and suggesting immediate fixes.

AI-Assisted Code Review

By leveraging GenAI, code review can be made more efficient and accurate. Automated code reviews powered by GenAI can identify security vulnerabilities, provide consistent and objective feedback, and generate secure code snippets.

AI-Powered Security Testing

According to IDC's November 2023 *DevOps Practices, Perceptions, and Tooling Survey*, security testing was identified as the top DevOps use case with the most potential to benefit from using GenAI. GenAI-powered application security testing tools can learn from previous scans, identify complex vulnerability patterns, and simulate sophisticated attack scenarios. This adaptability improves detection rates and accuracy, offering a more robust defense against emerging threats.

Considering OpenText Fortify

OpenText has its headquarters in Waterloo, Canada. The company significantly bolstered its application security offerings through the acquisition of Micro Focus in January 2023, which included the Fortify application security portfolio.

The core products of OpenText's application security suite include OpenText Fortify Static Code Analyzer (SCA) for static application security testing (SAST), OpenText Fortify WebInspect (DAST), OpenText Debricked (SCA), and OpenText Fortify On Demand, a SaaS platform offering SAST, SCA, DAST, and MAST capabilities. Recognizing organizations need to both secure GenAI applications and integrate GenAI into application security tools, OpenText recently introduced several new product features.

Fortify SCA detects 815 unique categories of vulnerabilities in 27 programming languages and integrates into all major integrated development environments (IDEs) for real-time analysis, as well as CI/CD tools for automation in the SDLC. It can be flexibly deployed as SaaS, on premises, or hybrid. Fortify SCA can now also help organizations scan GenAI applications, addressing vulnerabilities specific to ML models such as buffer overflows, injection flaws, and insecure coding practices. The tool can also detect weaknesses resulting from implicit trust of responses from AI/ML model APIs. Fortify SCA supports Python projects that consume OpenAI API, Amazon Web Services SageMaker, or LangChain. Users can identify and mitigate noncompliant code practices in LLM applications to ensure compliance with industry standards and regulations (such as GDPR, HIPAA, and PCI DSS). The product road map includes plans to expand coverage across more of the OWASP Top 10 for LLMs.

OpenText enables AI-powered application security through Fortify Aviator, part of the company's larger OpenText Aviator Platform, offering code fix suggestions in context, reducing the time developers spend remediating code security issues. Fortify Aviator audits and explains security issues in the context of a developer's code. This approach not only boosts developer productivity but also integrates AI fix suggestions seamlessly into developers' workflows for a frictionless experience.

Challenges

Despite GenAI's potential to revolutionize application security, its adoption may be gradual as organizations weigh the risks and benefits, potentially impacting the uptake of a tool like Fortify Aviator. Across IDC's GenAI research, privacy and data security concerns surface repeatedly as a top barrier to the technology's adoption. However, Fortify Aviator builds upon OpenText's expertise in AI, security, information governance, and risk management, offering a trusted solution to these challenges.

Innovative start-ups aiming to secure GenAI applications are emerging in the market, offering best-of-breed tools focused on the nuances of the vulnerabilities and risks associated with GenAI applications and the ML model supply chain. These solutions are often developed by specialists who are focused on the GenAI domain, leading to more rapid innovation and robust features. Such solutions might be attractive to organizations looking to secure their GenAI applications.

At the same time, there is a trend toward consolidation in application security. Some 86% of organizations are either consolidating or planning to consolidate their security tools, with an average of nearly 50 tools per organization, according to IDC's December 2023 *North American Tools and Vendors Consolidation Survey*. This trend may help OpenText defend against emerging players by allowing users to leverage the same tool to secure both GenAI and traditional web applications.

Conclusion

The power of GenAI to transform application security is profound, offering the ability to enhance cybersecurity defenses significantly. However, this transformation is not without its challenges. The dual nature of AI — as both a tool for advancing security measures and a potential source of vulnerabilities — requires a balanced approach. Organizations must embrace the advantages of AI, leveraging its strengths to bolster their cybersecurity efforts, while taking proactive measures to minimize its potential pitfalls.

Organizations must embrace the advantages of AI, leveraging its strengths to bolster their cybersecurity efforts, while taking proactive measures to minimize its potential pitfalls.

This evolution demands a proactive and dynamic approach from businesses to stay ahead of emerging threats and leverage the full potential of AI technologies. Organizations must remain vigilant, continuously assessing and addressing the risks associated with GenAI use. Continuous improvement and staying abreast of the latest AI security developments is crucial for organizations looking to harness GenAI's benefits. By making informed decisions and adopting a comprehensive security strategy, businesses can ensure their applications are secure and maintain their competitive edge.

The journey toward integrating GenAI into application security is marked by both opportunities and challenges. By embracing AI's transformative potential and addressing its inherent risks, organizations can unlock new avenues for innovation and security, paving the way for a more secure and efficient digital future.

About the Analyst

Katie Norton, *Research Manager, DevSecOps and Software Supply Chain Security*

Katie Norton is a Research Manager for IDC's DevSecOps and Software Supply Chain Security research practice. In this role, she is responsible for researching, writing, and advising clients on the fast-evolving DevSecOps and software supply chain security markets. With her background in research administration and data analytics, Katie takes a data-first approach in her market analysis.

MESSAGE FROM THE SPONSOR

As the world becomes increasingly digitized, enterprises are turning to AI-powered solutions to streamline their operations, enhance decision-making, and stay ahead of the competition. However, the integration of AI into critical business processes requires a thoughtful and strategic approach. With over a decade of experience using and securing AI, OpenText Cybersecurity and Fortify's application solutions offer a formative pathway for enterprises to harness the power of AI while prioritizing security, compliance, and responsible implementation.

Through Fortify's suite of tools, enterprises can assess their AI readiness, identify potential risks and vulnerabilities, and develop a comprehensive road map for AI integration. By blending security expertise with cutting-edge analytics, Fortify empowers organizations to make informed decisions, mitigate legal and regulatory challenges, and foster trust in their AI-driven initiatives.

To learn more about OpenText Fortify, visit https://www.opentext.com/products/application-security

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

IDC Research, Inc.
140 Kendrick Street
Building B
Needham, MA 02494, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
blogs.idc.com
www.idc.com

735-000012-001 | O | 08/24 | © 2024 Open Text