CSET CENTER for SECURITY and EMERGING TECHNOLOGY

# Key Takeaways from "Cybersecurity Risks of AI-Generated Code"

Artificial intelligence (AI) models have become increasingly adept at generating computer code. This makes them powerful and promising tools for software development across many industries, but they can also pose direct and indirect cybersecurity risks.

This report identifies three broad categories of risk associated with AI code generation: **1) models generating insecure code**, **2) models themselves being vulnerable to attack and manipulation**, and **3) downstream cybersecurity impacts** such as feedback loops in training future AI systems.

To further explore the risk of insecure AI-generated code, we also evaluated code generated by five large language models which were given the same set of prompts. Our evaluation results show that **almost half of the code snippets produced by these five different models contain bugs** that are often impactful and could potentially lead to malicious exploitation.

## Key takeaways:

- Industry adoption of AI code generation models may pose risks to software supply chain security, but these risks are unlikely to be evenly distributed across organizations and industries.
- Multiple stakeholders have roles to play in ensuring AI-generated code is secure. This includes users, developers, organizations producing code at scale, and policymakers and industry leaders. Existing cybersecurity guidance should be applied in cases where AI code generation models are used to write code, and other guidance can be expanded to address AI systems that impact software supply chain security.
- It is currently difficult to evaluate code generation models for security. Evaluation benchmarks often do not focus on models' ability to generate secure code, and there is also inadequate transparency to address questions such as the relationship between model performance and code security.

## For more information:

- Download the report: https://cset.georgetown.edu/publication/cybersecurity-risks-of-ai-generated-code/
- Contact us: cset@georgetown.edu