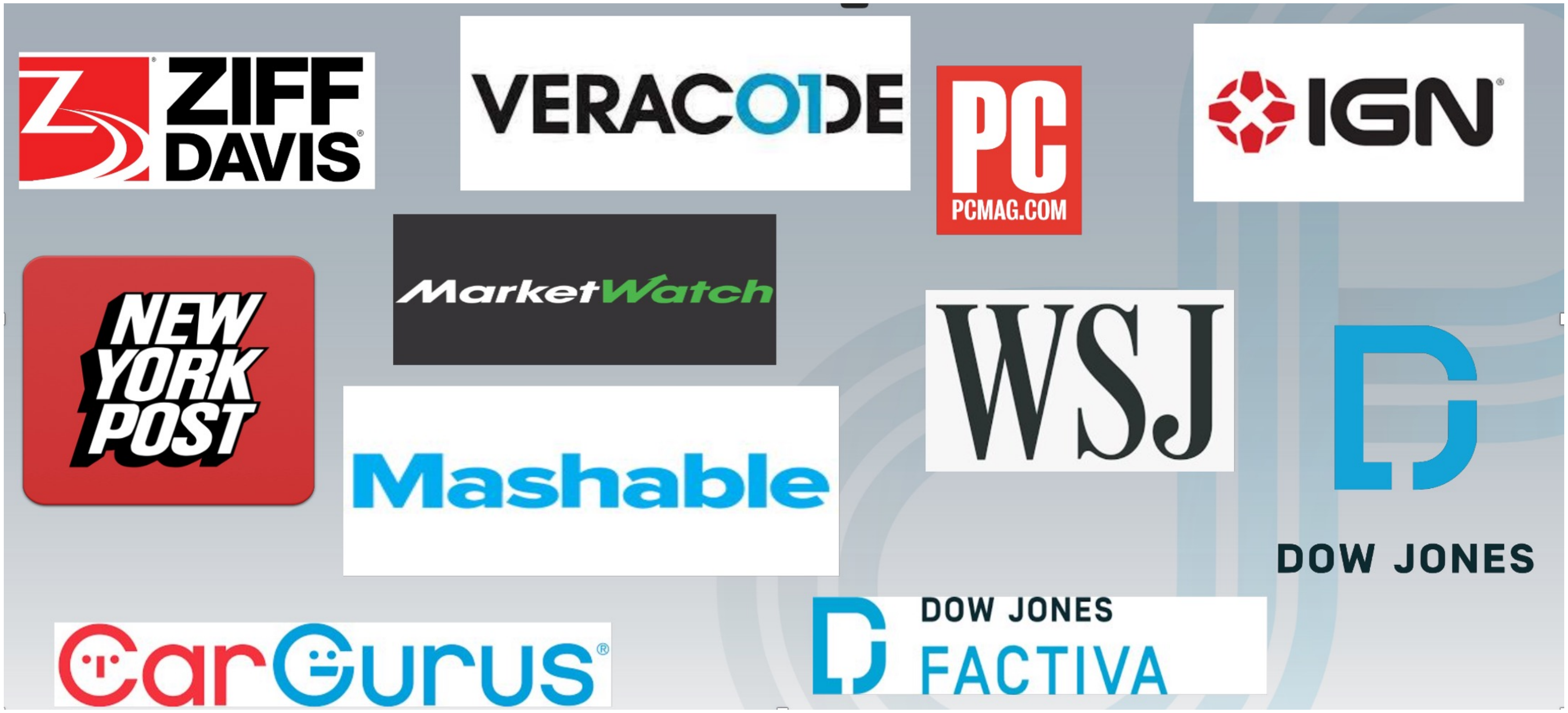
Crossroads of AppSec & Gen-AI

Sohail Iqbal

VP / CISO @ Veracode

VERACODE

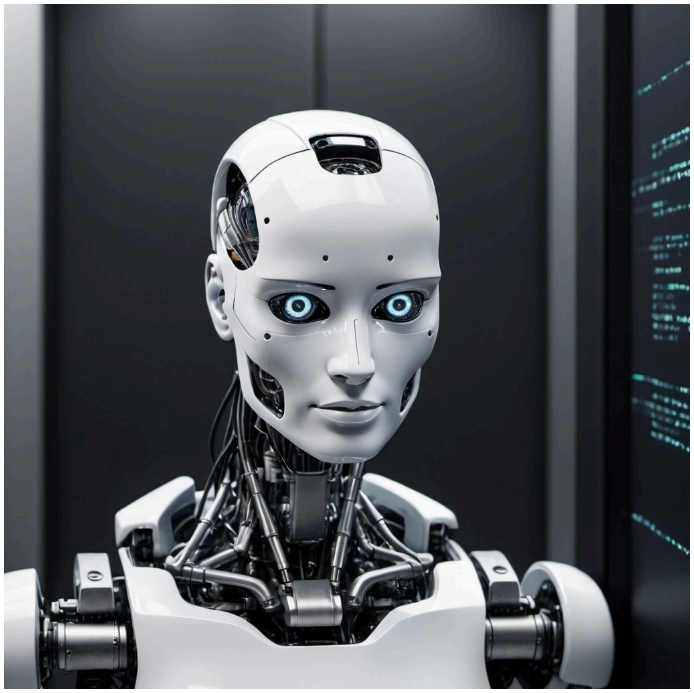
Leadership Roles



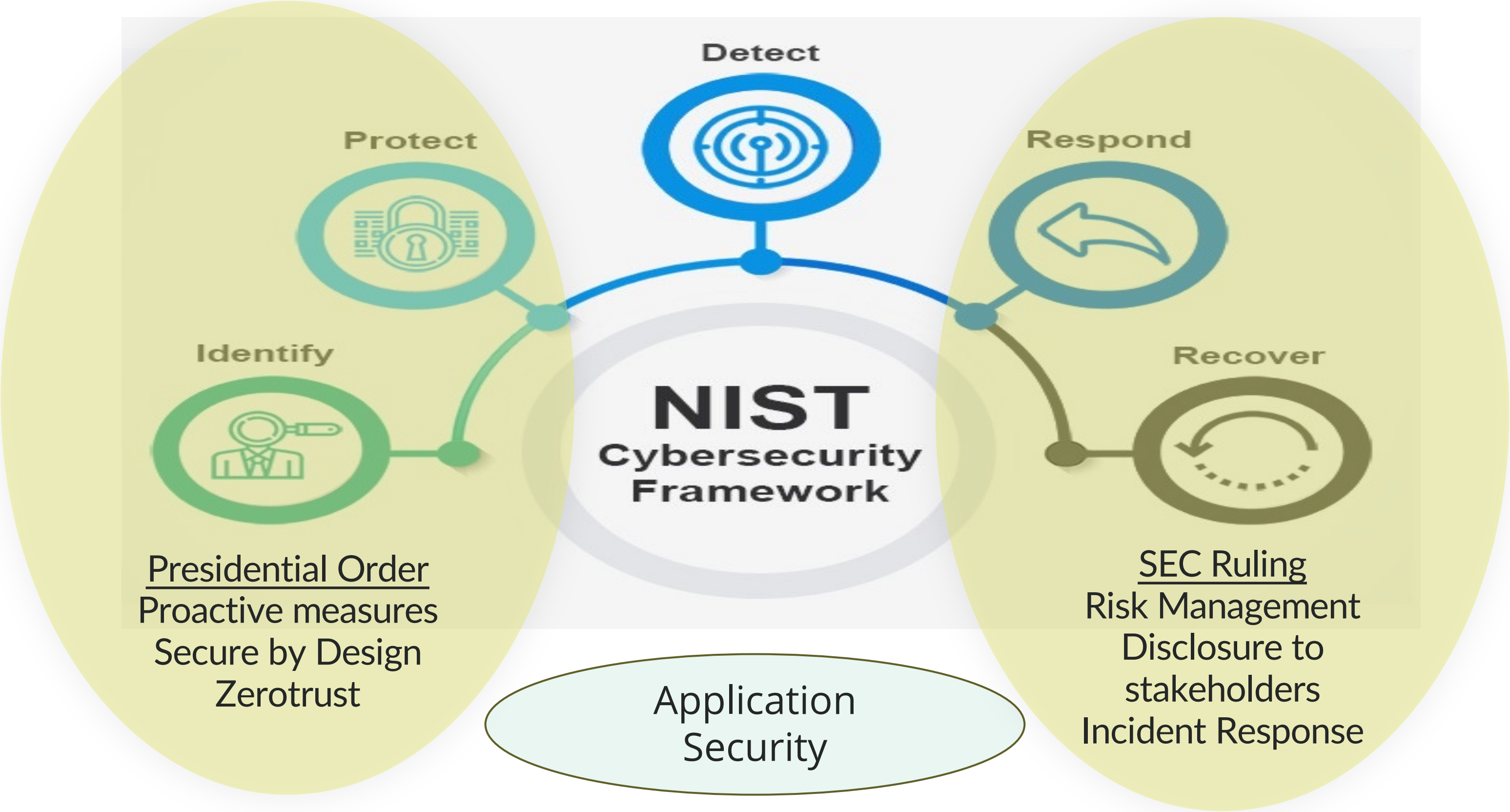
Crossroads of AppSec & Gen-AI

Agenda

- AppSec under spotlight
- Gen-AI good, bad & ugly
- AppSec on Gen-AI steroids



AppSec Under The Spotlight

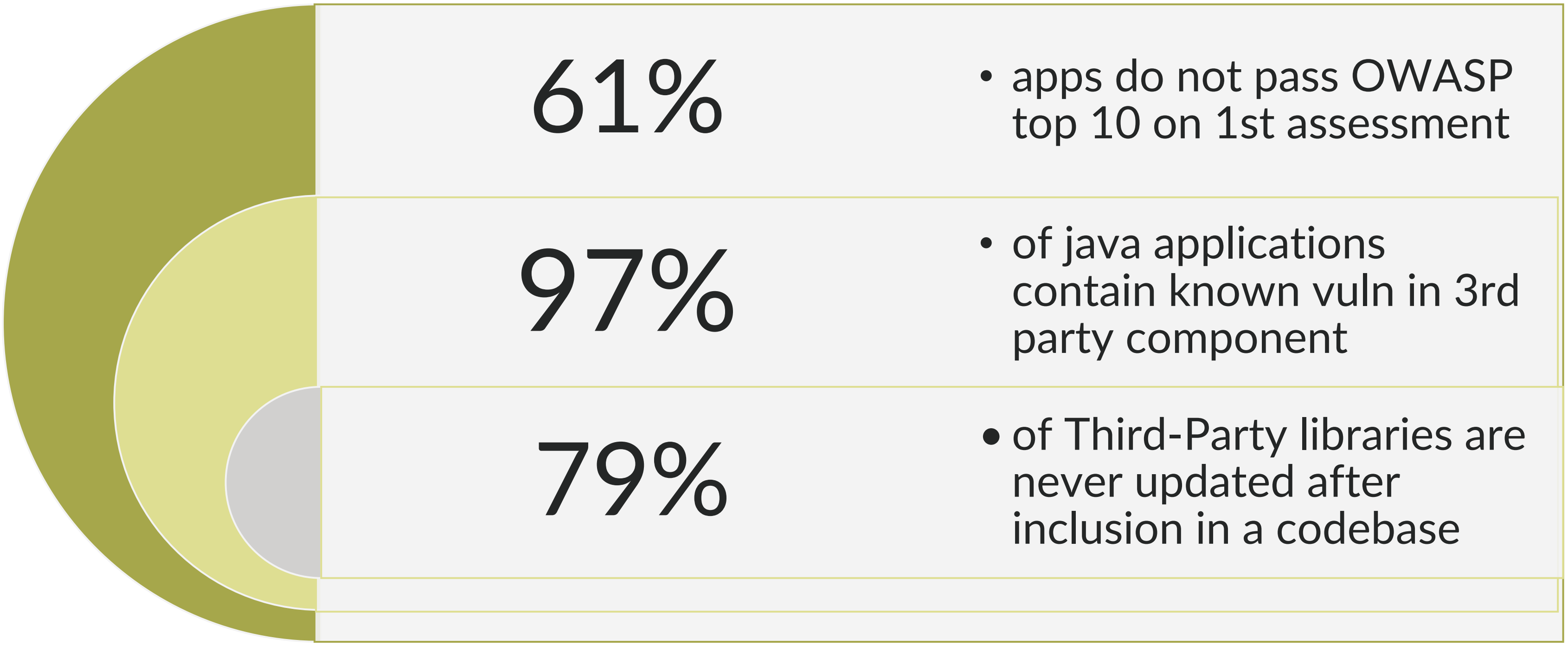


<u>Presidential Order</u>
Proactive measures
Secure by Design
Zerotrust

Application
Security

<u>SEC Ruling</u>
Risk Management
Disclosure to
stakeholders
Incident Response

Current State of Applications Security



61%

- apps do not pass OWASP top 10 on 1st assessment

97%

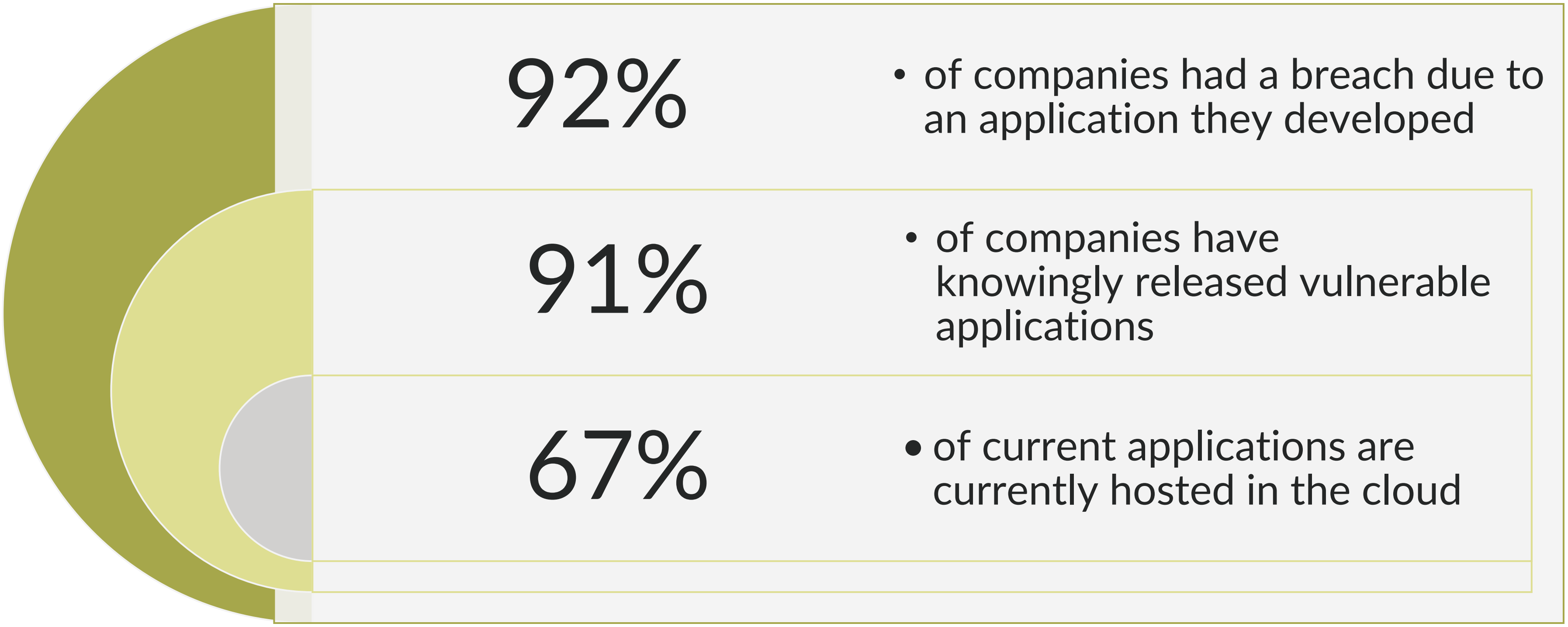
- of java applications contain known vuln in 3rd party component

79%

- of Third-Party libraries are never updated after inclusion in a codebase

Ref: Veracode SOSS

Current State Of Applications Security



92%

- of companies had a breach due to an application they developed

91%

- of companies have knowingly released vulnerable applications

67%

- of current applications are currently hosted in the cloud

Ref: Checkmarx

Current State Of Vuln Exploitation

- Verizon Data Breach Investigations Report (DBIR) 2024, which analyzes 10,626 confirmed data breaches.
- The latest Verizon DBIR shows that vulnerability exploitation as a method for criminals to attack companies with ransomware and extortion grew 180% in 2023.

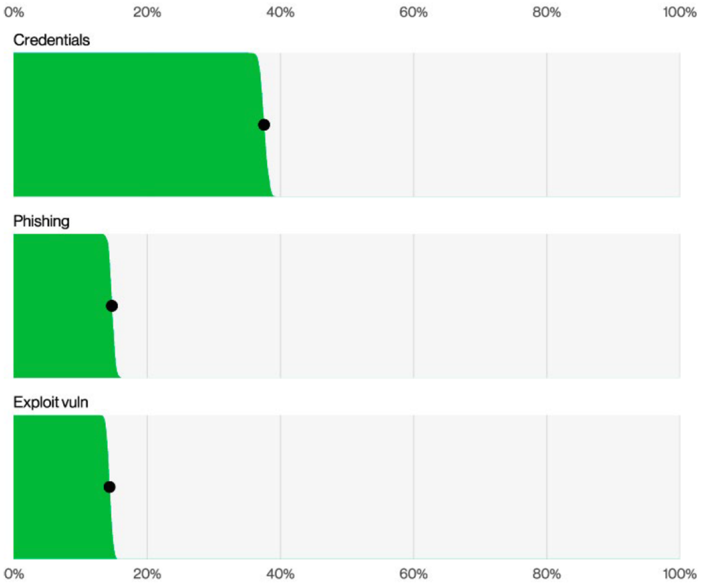
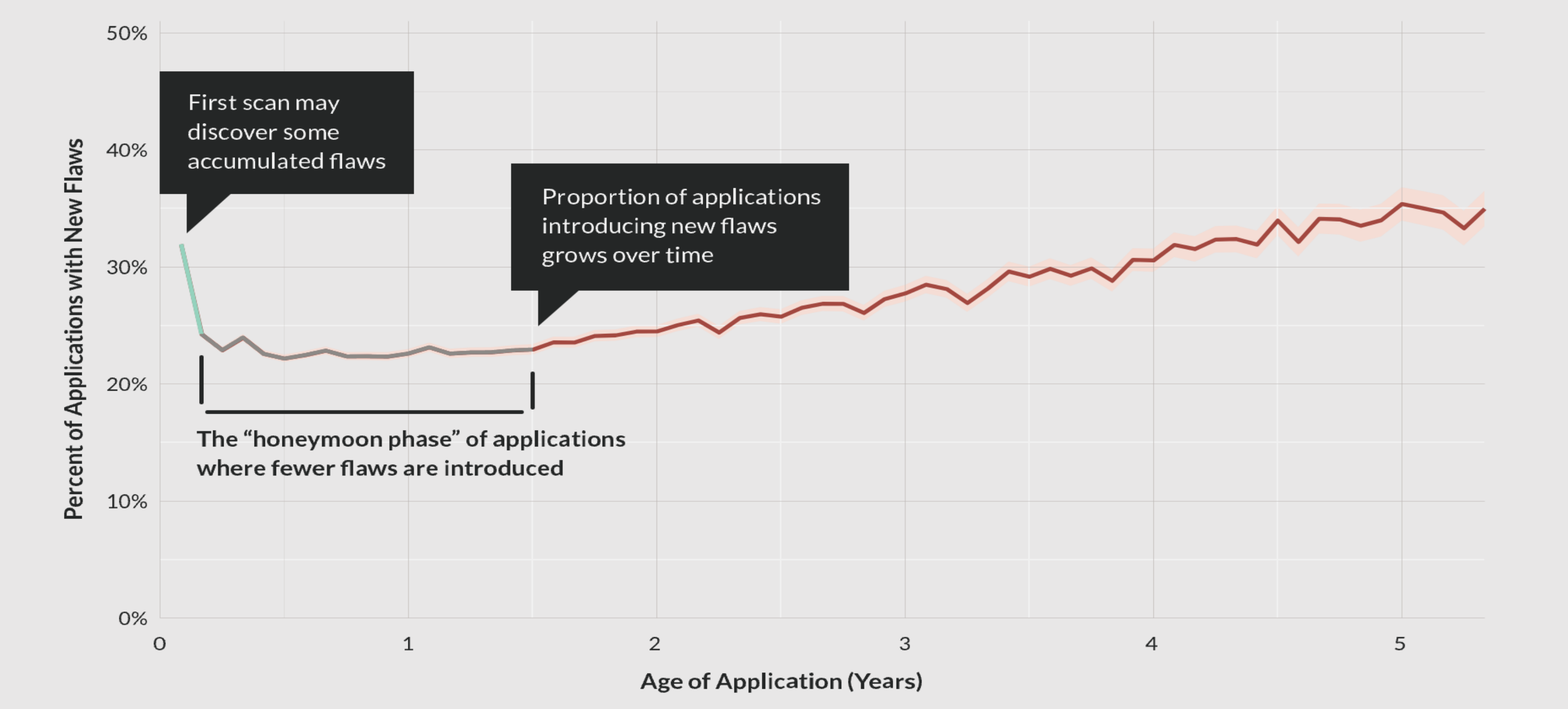


Figure 1. Select ways-in enumerations in non-Error, non-Misuse breaches (n=6,963)

New Flaws Introduced, By Application Age

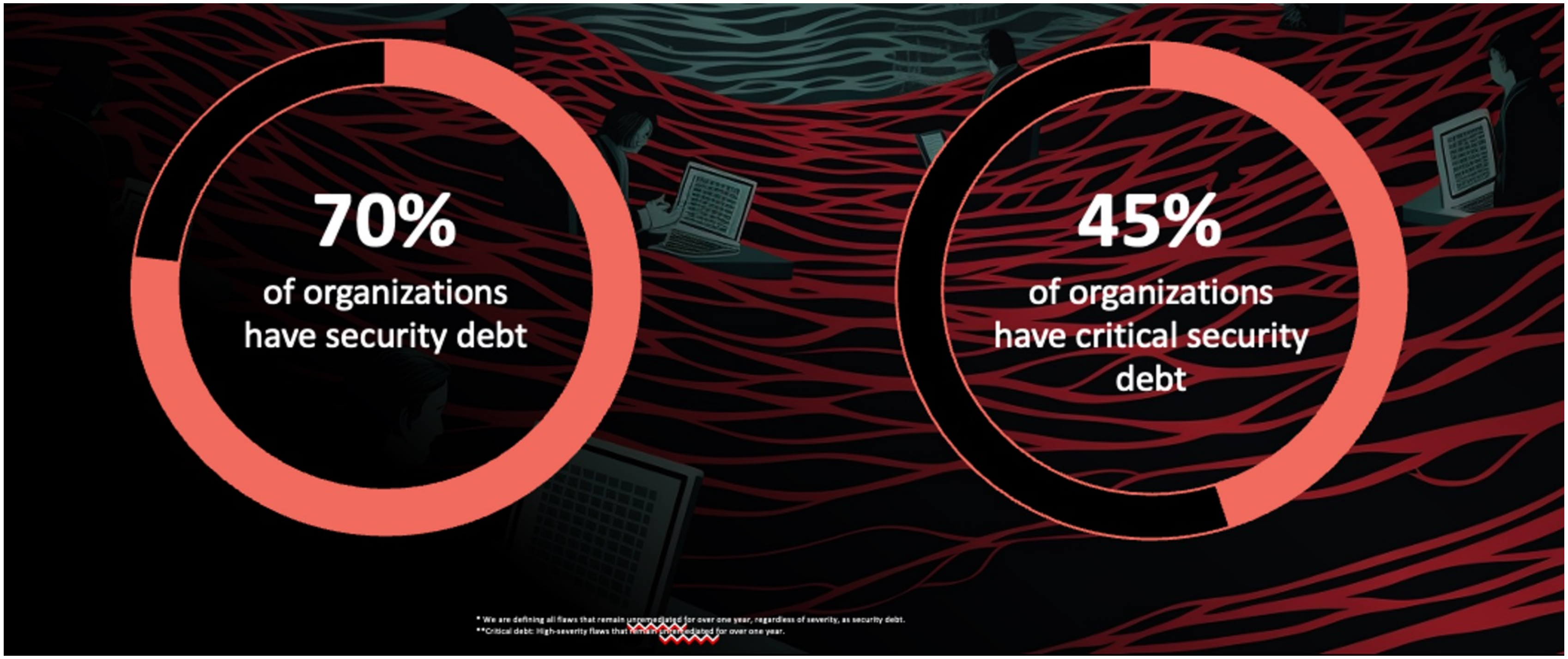


First scan may discover some accumulated flaws

Proportion of applications introducing new flaws grows over time

The “honeymoon phase” of applications where fewer flaws are introduced

Application Security Debt



70%

of organizations have security debt

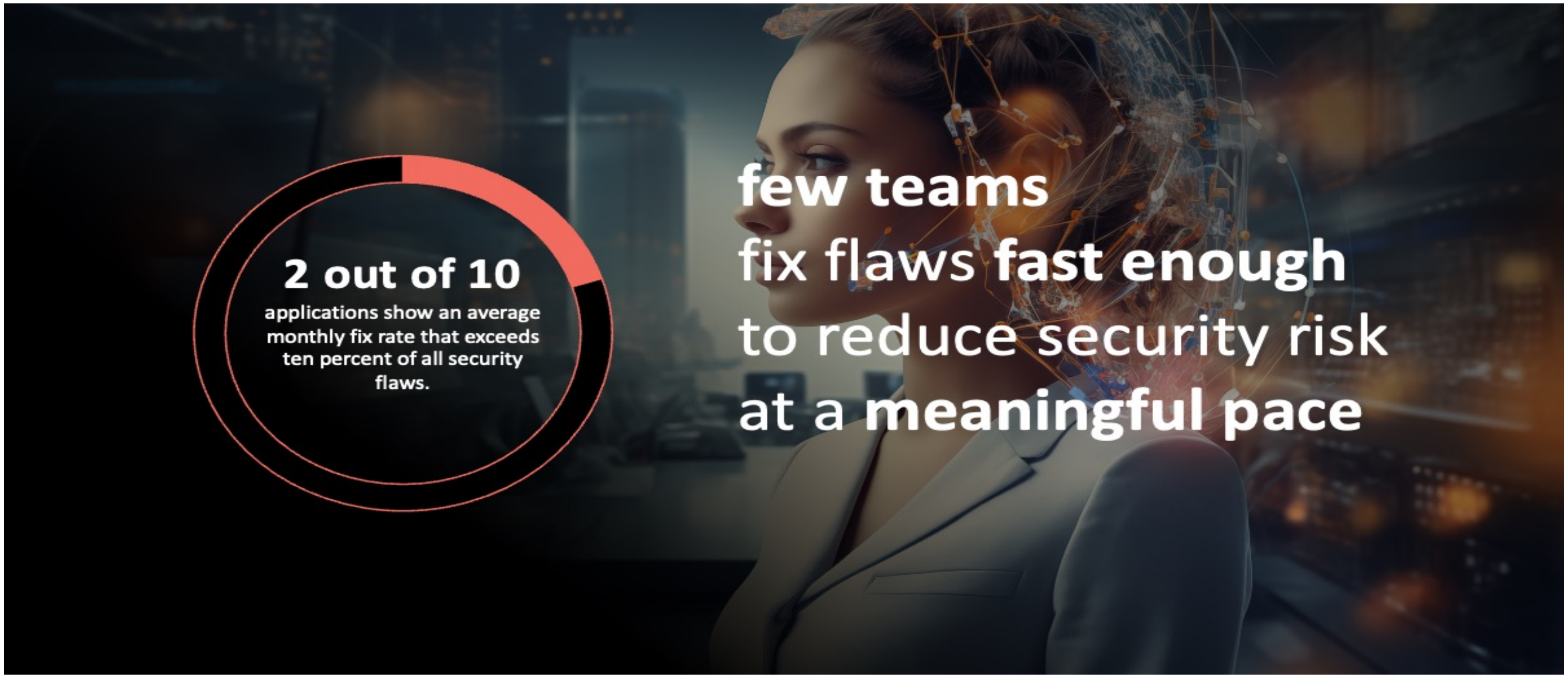
45%

of organizations have critical security debt

* We are defining all flaws that remain unremediated for over one year, regardless of severity, as security debt.

**Critical debt: High-severity flaws that remain unremediated for over one year.

Remediation Challenges


2 out of 10
applications show an average monthly fix rate that exceeds ten percent of all security flaws.
few teams
fix flaws fast enough
to reduce security risk
at a meaningful pace

Velocity Issue

Today we are finding software security flaws faster than we can fix them

Now let's add the exciting new potential of generative AI that can write code!

I TRAINED MY MODEL ON ALL THE WORLDS CODE!
YOU REMOVED THE EXPLOITABLE STUFF RIGHT?
imgflip.com

Generative-AI (gen-AI)

Increased productivity

Enhanced creativity

Improved decision-making

good genie (gen-AI)

evil genie (gen-AI)

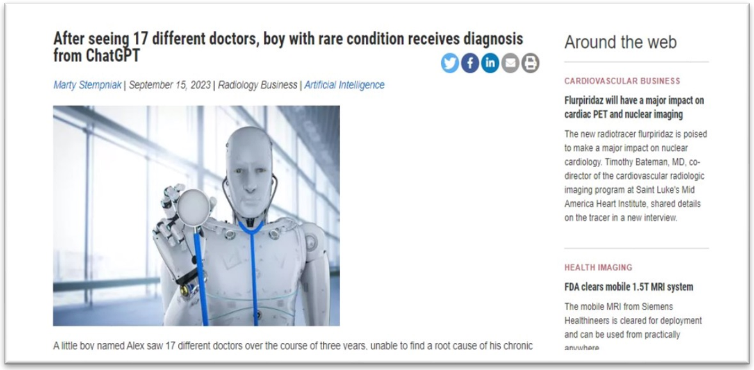
Bias

Security risks

Privacy concerns

Gen-AI Impacts

✓



After seeing 17 different doctors, boy with rare condition receives diagnosis from ChatGPT

Marty Stempniak | September 15, 2023 | Radiology Business | Artificial Intelligence

A little boy named Alex saw 17 different doctors over the course of three years, unable to find a root cause of his chronic

Around the web

CARDIOVASCULAR BUSINESS

Flurpiridaz will have a major impact on cardiac PET and nuclear imaging

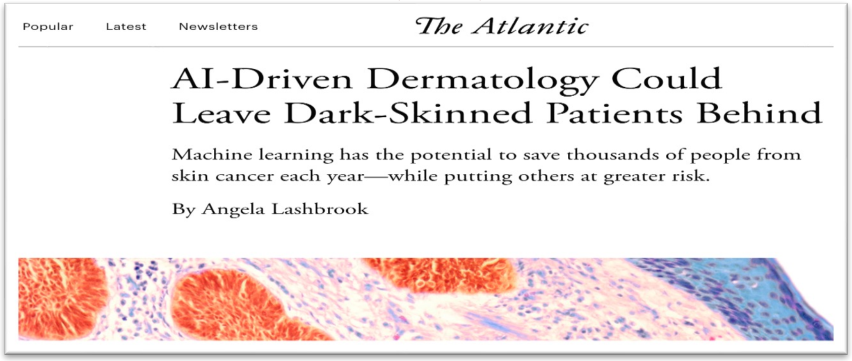
The new radiotracer flurpiridaz is poised to make a major impact on nuclear cardiology. Timothy Bateman, MD, co-director of the cardiovascular radiologic imaging program at Saint Luke's Mid America Heart Institute, shared details on the tracer in a new interview.

HEALTH IMAGING

FDA clears mobile 1.5T MRI system

The mobile MRI from Siemens Healthineers is cleared for deployment and can be used from practically anywhere

✗

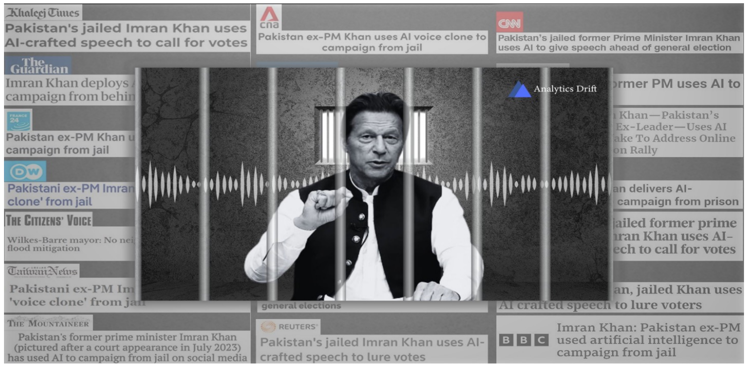


Popular Latest Newsletters *The Atlantic*

AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind

Machine learning has the potential to save thousands of people from skin cancer each year—while putting others at greater risk.

By Angela Lashbrook



Khaleej Times: Pakistan's jailed Imran Khan uses AI-crafted speech to call for votes

cna: Pakistan ex-PM Khan uses AI voice clone to campaign from jail

CNN: Pakistan's jailed former Prime Minister Imran Khan uses AI to give speech ahead of general election

The Guardian: Imran Khan deploys A campaign from behin

France 24: Pakistan ex-PM Khan u campaign from jail

DW: Pakistani ex-PM Imran clone' from jail

The Citizens' Voice: Wilkes-Barre mayor: No nei flood mitigation

Taiwan News: Pakistani ex-PM Im 'voice clone' from jail

The Mountaineer: Pakistan's former prime minister Imran Khan (pictured after a court appearance in July 2023) has used AI to campaign from jail on social media

rmer PM uses AI to

Khan—Pakistan's Ex-Leader—Uses AI ake To Address Online on Rally

an delivers AI- campaign from prison

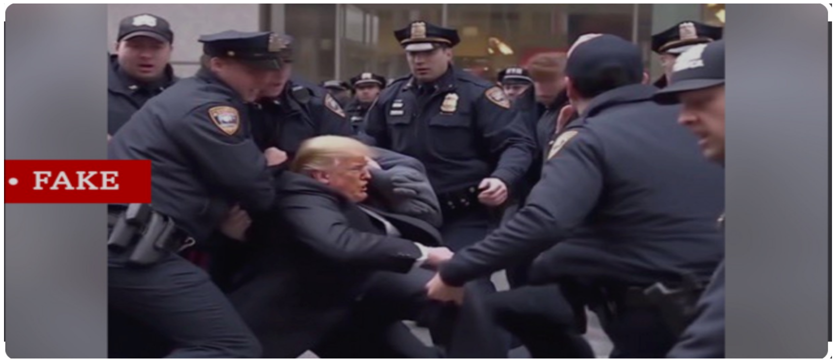
jailed former prime ran Khan uses AI- ech to call for votes

an, jailed Khan uses AI crafted speech to lure voters

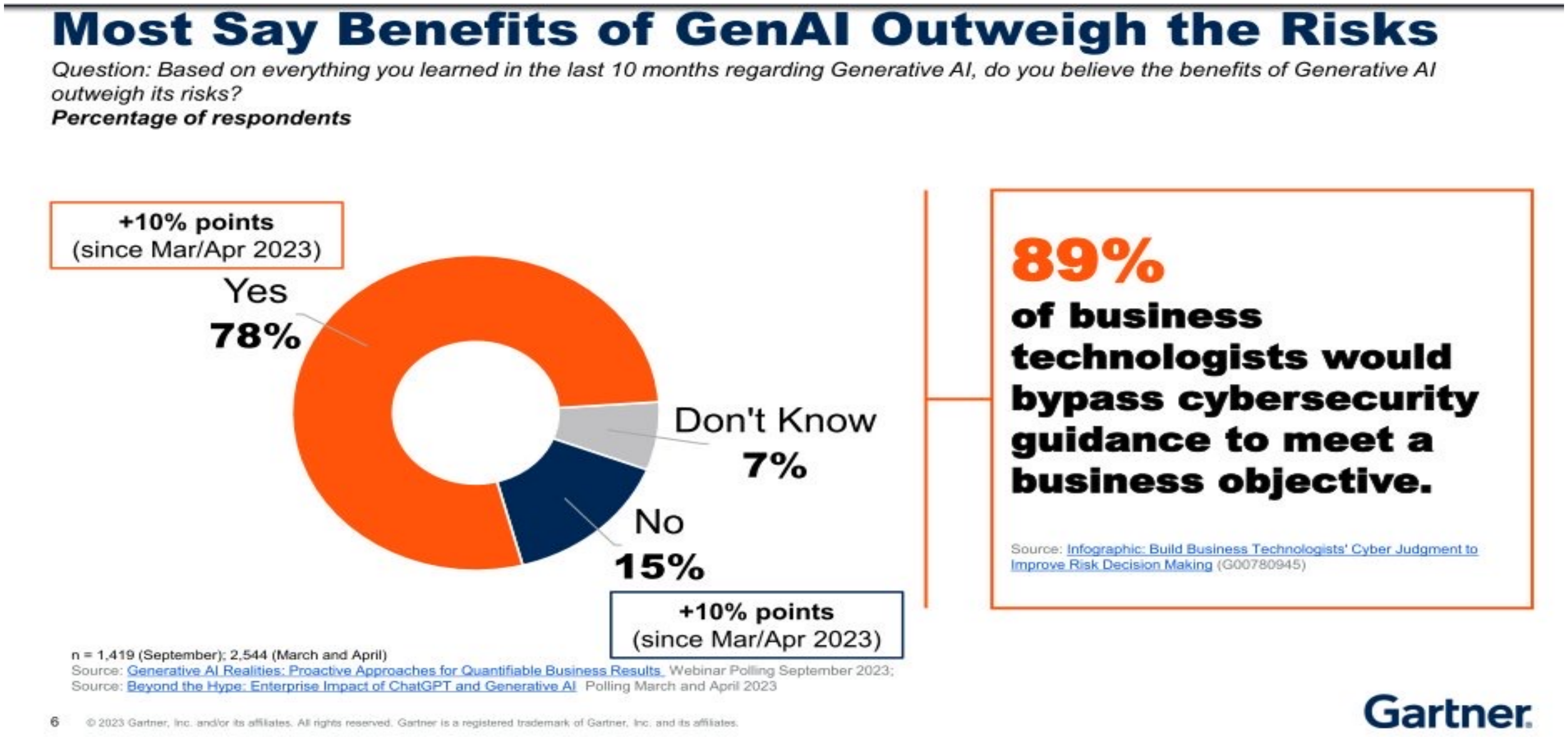
general elections

Reuters: Pakistan's jailed Imran Khan uses AI-crafted speech to lure votes

BBC: Imran Khan: Pakistan ex-PM used artificial intelligence to campaign from jail



Gen-AI Future



Most Say Benefits of GenAI Outweigh the Risks

Question: Based on everything you learned in the last 10 months regarding Generative AI, do you believe the benefits of Generative AI outweigh its risks?
Percentage of respondents

89%
of business technologists would bypass cybersecurity guidance to meet a business objective.

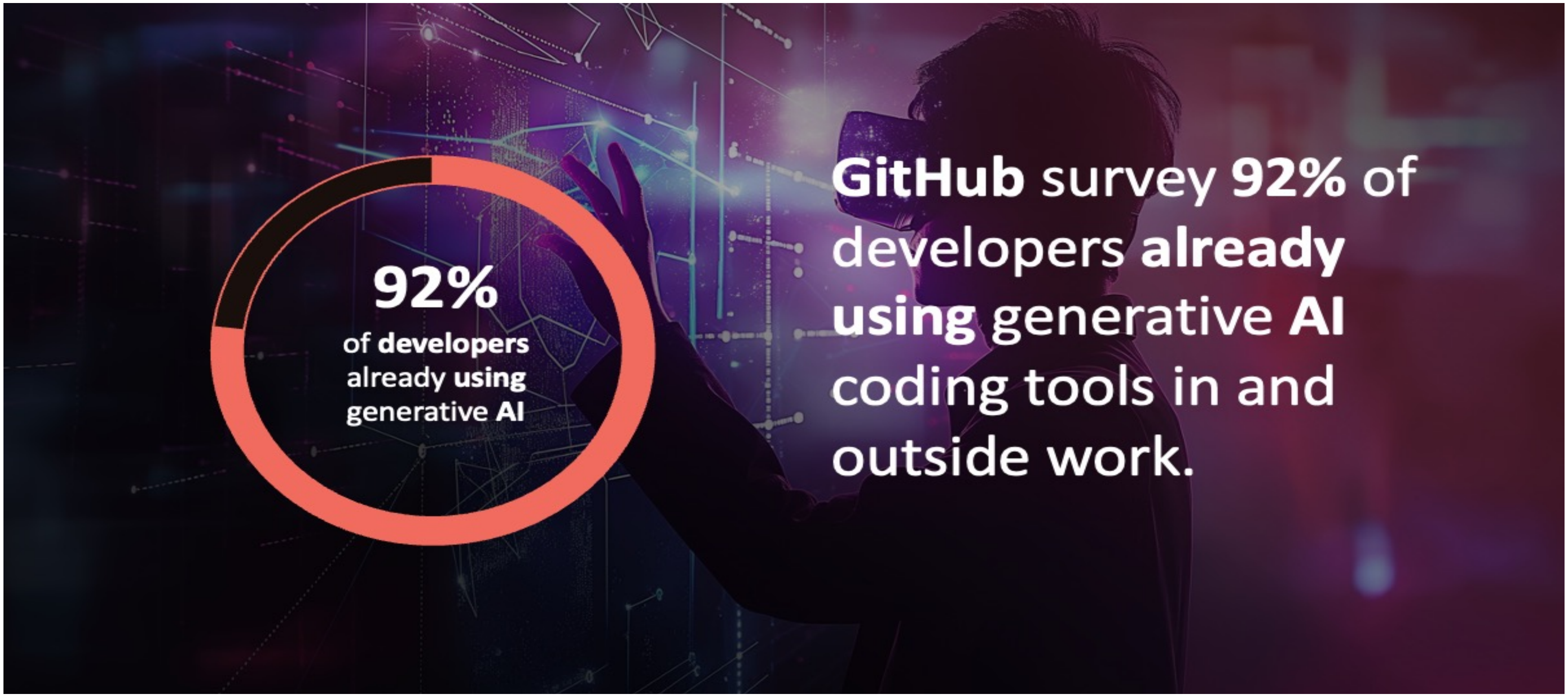
Source: Infographic: Build Business Technologists' Cyber Judgment to Improve Risk Decision Making (G00780945)

n = 1,419 (September); 2,544 (March and April)
Source: Generative AI Realities: Proactive Approaches for Quantifiable Business Results Webinar Polling September 2023;
Source: Beyond the Hype: Enterprise Impact of ChatGPT and Generative AI Polling March and April 2023

6

Gartner

Gen-AI Adoption



92% of **developers** already **using** generative **AI**

GitHub survey **92%** of developers **already using** generative **AI** coding tools in and outside work.

To Go Faster Developers Leverage

open source

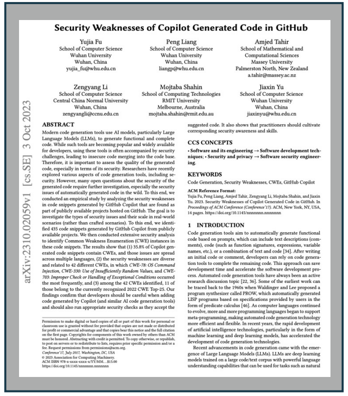
code repositories

code generators

large language models

Case Studies

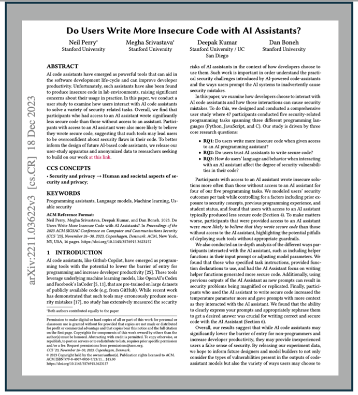
Cornell University Study on AI Code Generators	**Stanford University Study** on AI Code Generators	**New York University Study** on GitHub Copilot	**Purdue University** on ChatGPT accuracy
35%	**52%**	**39%**	**75%**
Out of the **435 Copilot** generated code snippets **35%** contain security weaknesses, across **6** programming languages.	**52%** of **ChatGPT's** answers contain inaccuracies and **77%** are verbose when answering software engineering questions.	**39.34%** of the time users still prefer **ChatGPT's** responses due to their comprehensiveness and articulate language style.	**75%** of developers have said that GitHub **CoPilot** makes them code faster.



arXiv:2310.02059v1 [cs.SE] 3 Oct 2023

Security Weaknesses of Copilot Generated Code in GitHub

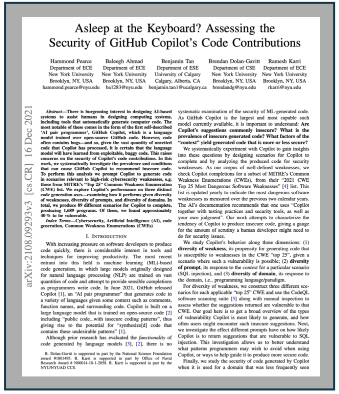
Yujia Fu, School of Computer Science, Wuhan University, Wuhan, China, yujia_fu@whu.edu.cn
Peng Liang, School of Computer Science, Wuhan University, Wuhan, China, liangp@whu.edu.cn
Amjed Tahir, School of Mathematical and Computational Sciences, Massey University, Palmerston North, New Zealand, a.tahir@massey.ac.nz
Zengyang Li, School of Computer Science, Central China Normal University, Wuhan, China, zengyangli@ccnu.edu.cn
Mojtaba Shahin, School of Computing Technologies, RMIT University, Melbourne, Australia, mojtaba.shahin@rmit.edu.au
Jiaxin Yu, School of Computer Science, Wuhan University, Wuhan, China, jiaxinyu@whu.edu.cn



arXiv:2211.03622v3 [cs.CR] 18 Dec 2023

Do Users Write More Insecure Code with AI Assistants?

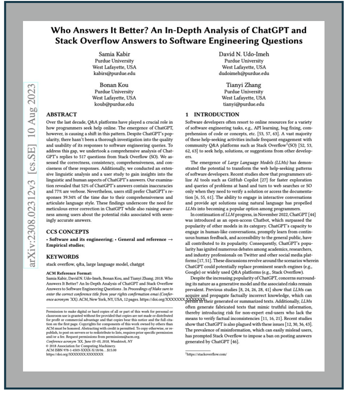
Neil Perry*, Stanford University
Megha Srivastava*, Stanford University
Deepak Kumar, Stanford University / UC San Diego
Dan Boneh, Stanford University



arXiv:2108.09293v3 [cs.CR] 16 Dec 2021

Asleep at the Keyboard? Assessing the Security of GitHub Copilot's Code Contributions

Hammond Pearce, Department of ECE, New York University, Brooklyn, NY, USA, hammond.pearce@nyu.edu
Baleegh Ahmad, Department of ECE, New York University, Brooklyn, NY, USA, ba1283@nyu.edu
Benjamin Tan, Department of ESE, University of Calgary, Calgary, Alberta, CA, benjamin.tan1@ucalgary.ca
Brendan Dolan-Gavitt, Department of CSE, New York University, Brooklyn, NY, USA, brendandg@nyu.edu
Ramesh Karri, Department of ECE, New York University, Brooklyn, NY, USA, rkarri@nyu.edu

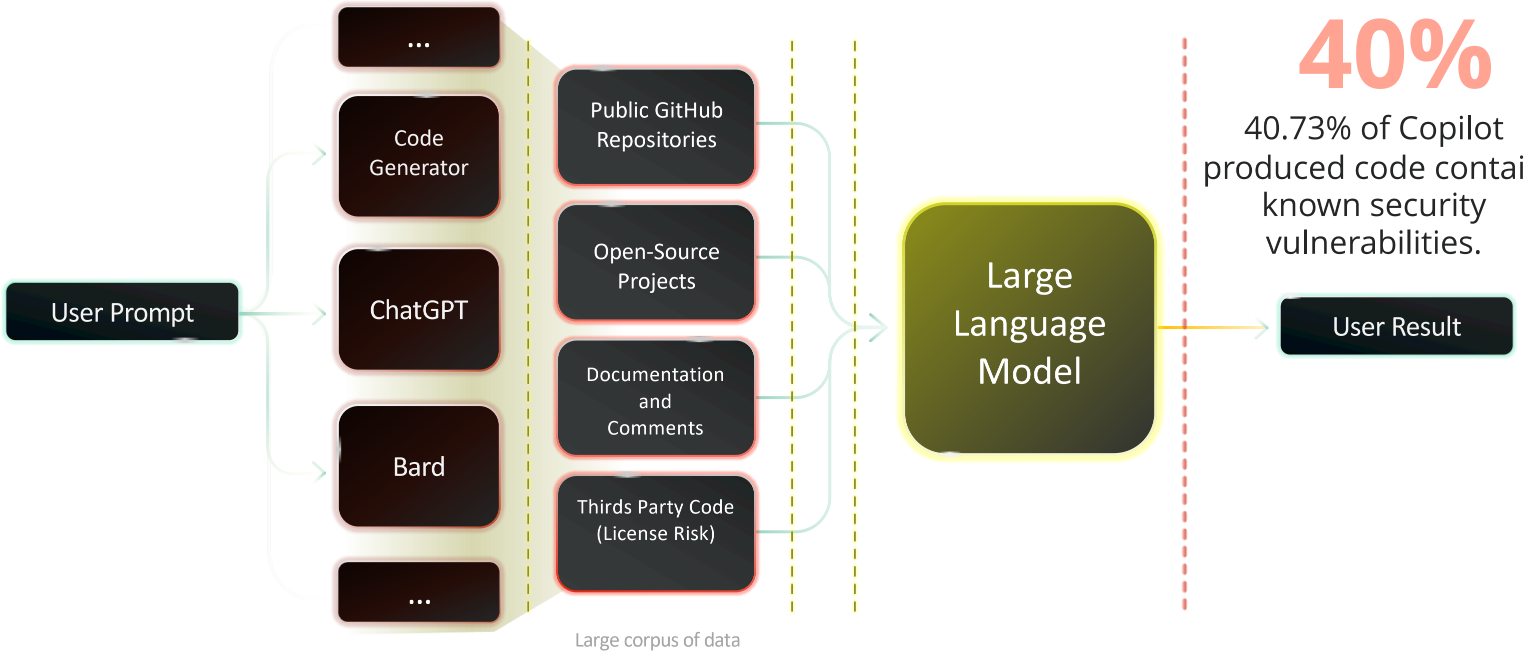


arXiv:2308.02312v3 [cs.SE] 10 Aug 2023

Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions

Samia Kabir, Purdue University, West Lafayette, USA, kabirs@purdue.edu
David N. Udo-Imeh, Purdue University, West Lafayette, USA, dudoimeh@purdue.edu
Bonan Kou, Purdue University, West Lafayette, USA, koub@purdue.edu
Tianyi Zhang, Purdue University, West Lafayette, USA, tianyi@purdue.edu

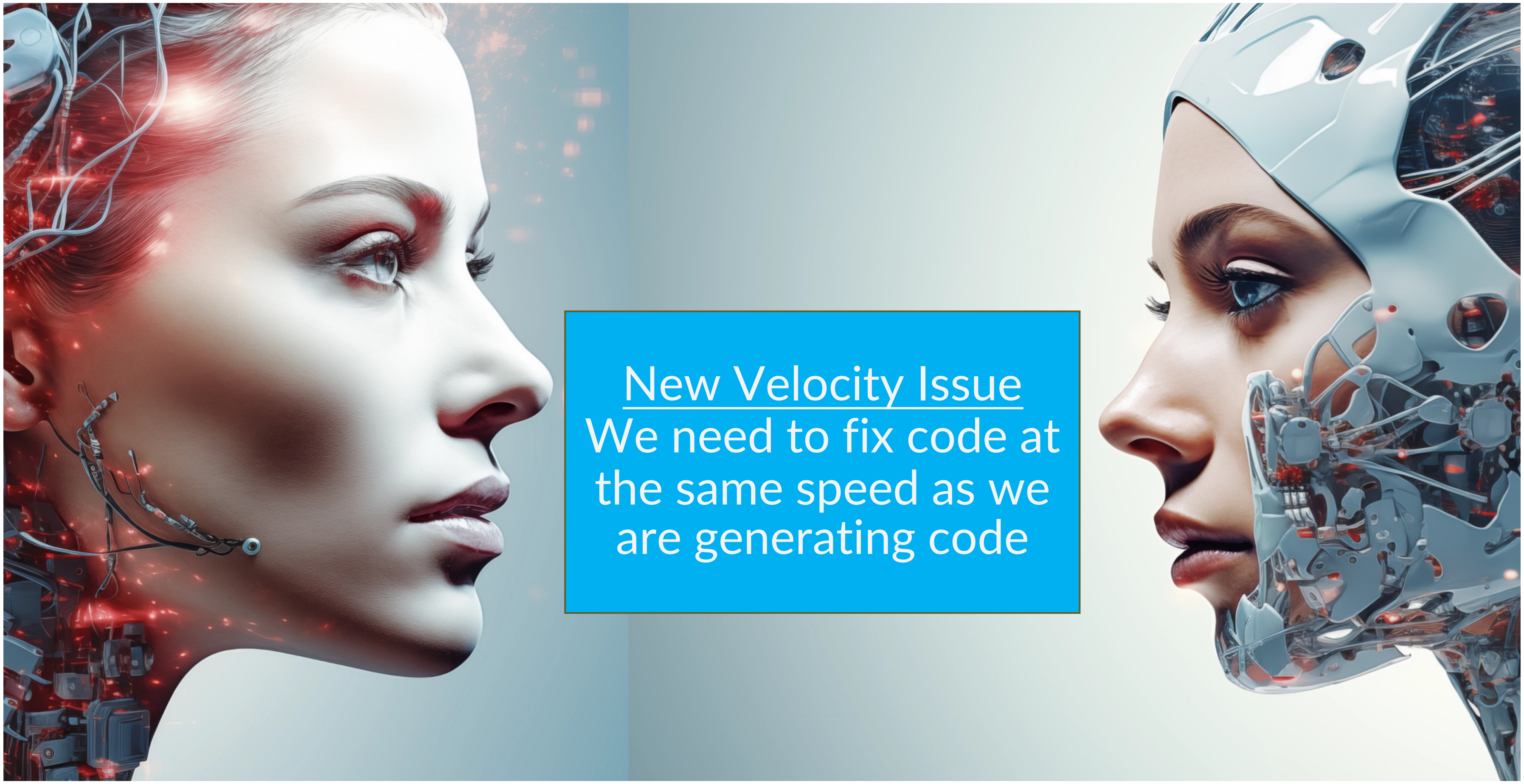
Large Language Models



Large corpus of data that includes open web content.

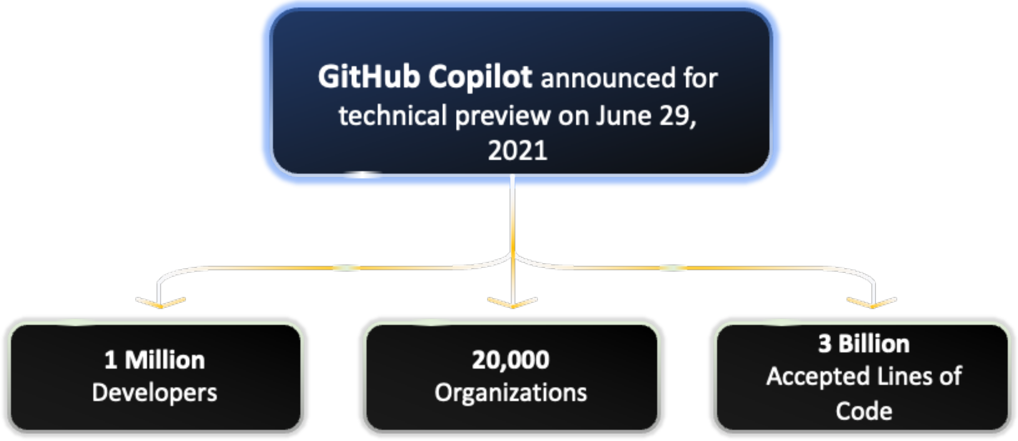
40%

40.73% of Copilot produced code contain known security vulnerabilities.



New Velocity Issue
We need to fix code at the same speed as we are generating code

Gen-AI Driven Remediation – Veracode Fix



GitHub Copilot announced for technical preview on June 29, 2021

1 Million Developers

20,000 Organizations

3 Billion Accepted Lines of Code

The world's most widely adopted AI development tool

▶

40%

of Copilot produced code contains known security vulnerabilities.

Intelligent Remediation Engine

Veracode Fix

Secure All of Your Code

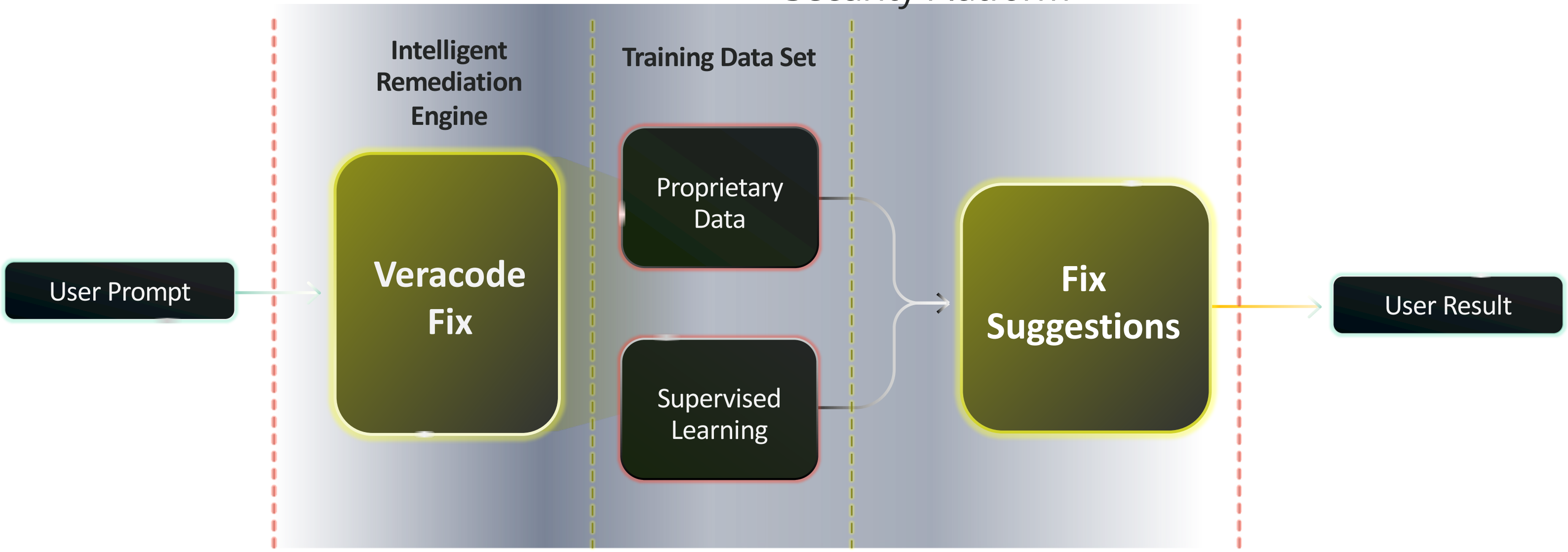
Veracode Fix Approach

Veracode Curated Dataset

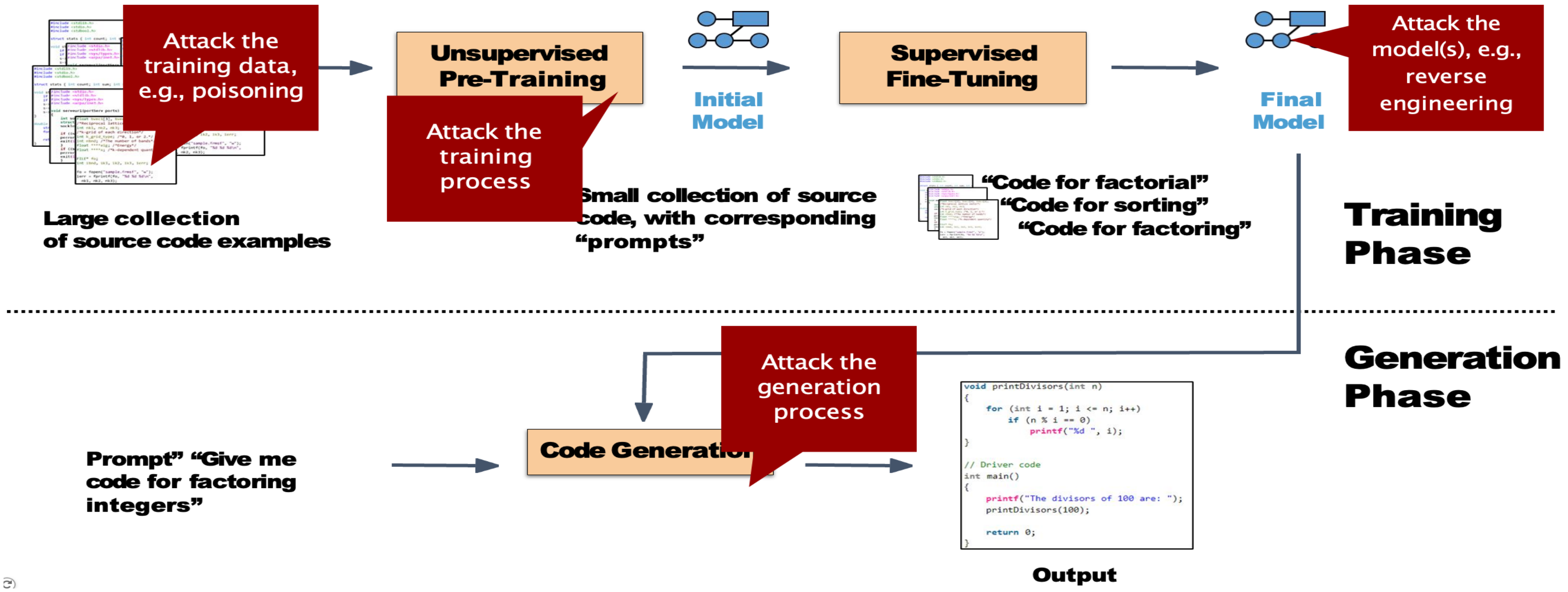
Code Provenance Assurance

Backed by Veracode's Intelligent Software Security Platform

Coverage for all that matters



Attacking the SDLC



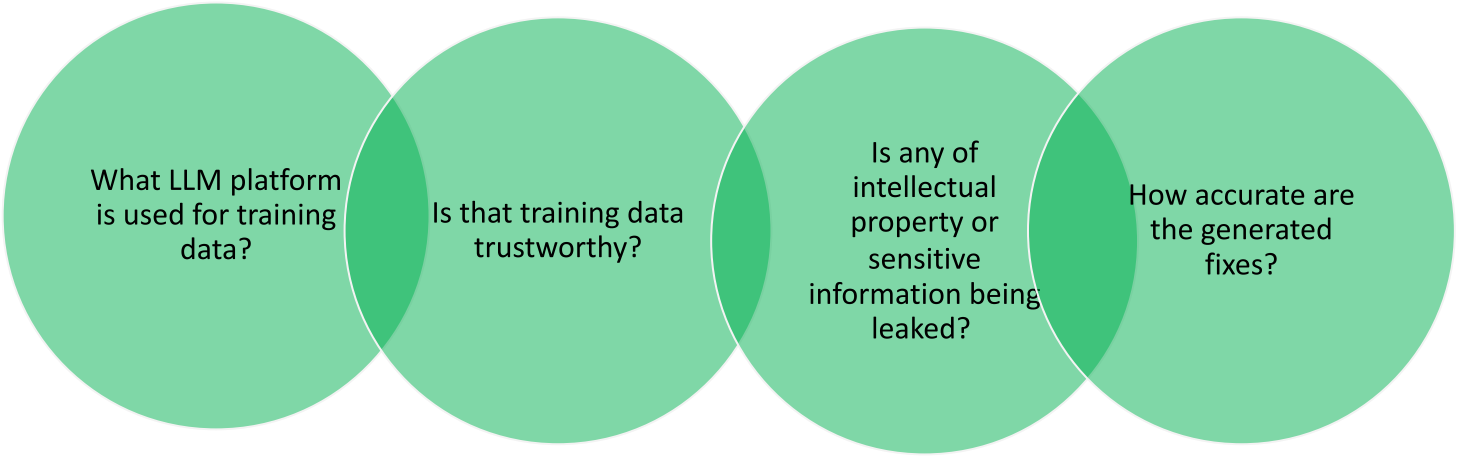
Attack the training data, e.g., poisoning
Unsupervised Pre-Training
Attack the training process
Initial Model
Supervised Fine-Tuning
Final Model
Attack the model(s), e.g., reverse engineering
Large collection of source code examples
Small collection of source code, with corresponding "prompts"
"Code for factorial"
"Code for sorting"
"Code for factoring"
Training Phase
Generation Phase
Attack the generation process
Prompt" "Give me code for factoring integers"
Code Generation
Output

```
void printDivisors(int n)
{
    for (int i = 1; i <= n; i++)
        if (n % i == 0)
            printf("%d ", i);
}

// Driver code
int main()
{
    printf("The divisors of 100 are: ");
    printDivisors(100);

    return 0;
}
```

Ref: MIT Lincoln Lab

Recommendations for Gen-AI and Code Security

Carefully consider the implementation details if you are thinking about leveraging AI for developing and/or securing code



What LLM platform is used for training data?

Is that training data trustworthy?

Is any of intellectual property or sensitive information being leaked?

How accurate are the generated fixes?

Be aware of human biases that trick us into feeling overly confident about the correctness of AI-generated content

Takeaways

For Corporate

- Document Gen-AI policy & standards for clear guidelines of adoption.
- Build a virtual fence for your sensitive data.
- Mature 3rd party risk assessment program to perform due diligence against vendors/partners using Gen-AI.

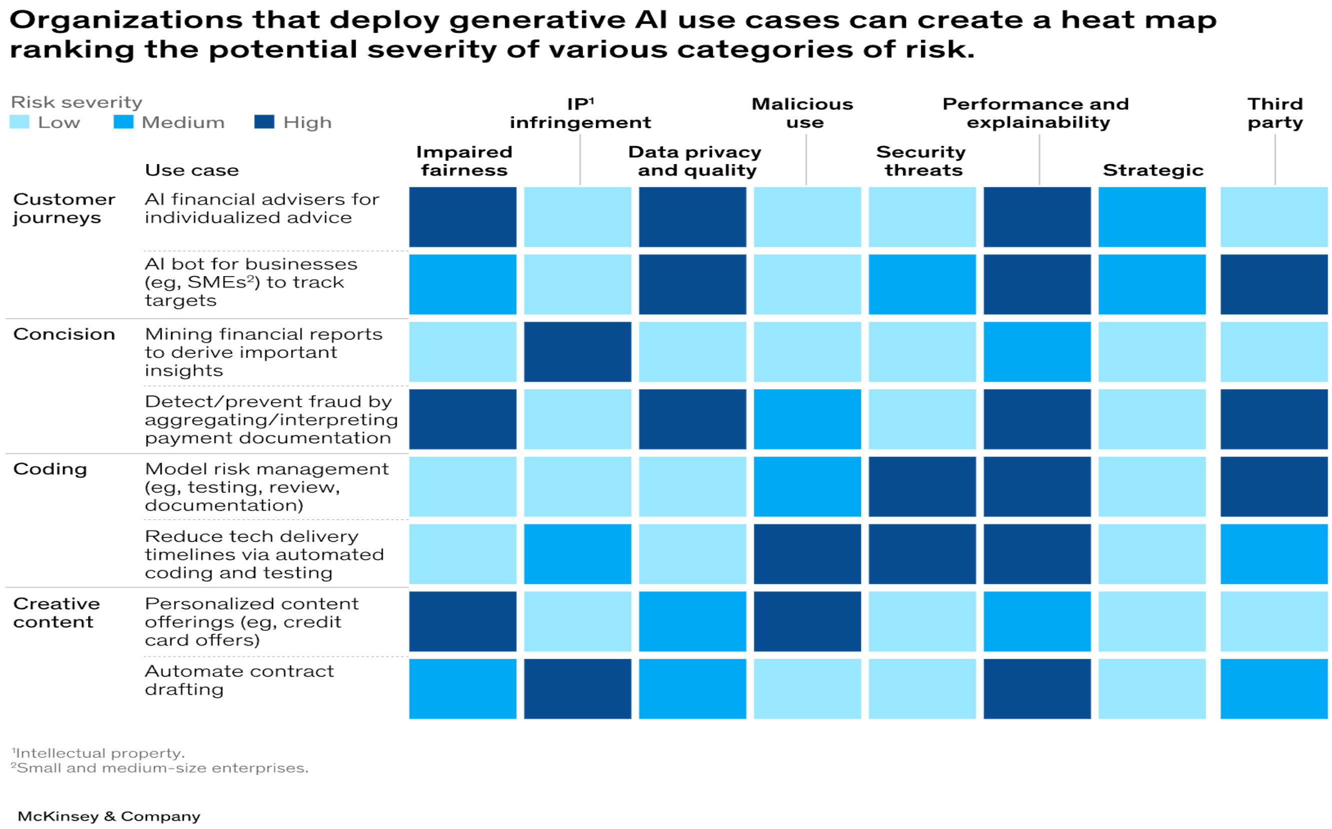
For Code

- Align SDLC processes to incorporate Gen-AI code generators.
- Use purpose built LLM platforms instead of open-source.
- Track IP infringement, copyrights and licensing issues.
- Implement Gen-AI based remediation tools to overcome the tech debt.

Adversaries

- Train users for identifying high fidelity Gen-AI based phishing or deepfake attacks.
- Limit any data ingestions by LLM platforms which allows adversaries to infer intelligence.
- Secure your LLM platforms from compromise, poisoning or data exfiltration attempts.
- Monitor malicious campaigns impacting the reputation & credibility of Organization.

Heatmap



Organizations that deploy generative AI use cases can create a heat map ranking the potential severity of various categories of risk.

Risk severity

Low Medium High

	Use case	Impaired fairness	IP[1] infringement	Data privacy and quality	Malicious use	Security threats	Performance and explainability	Strategic	Third party
Customer journeys	AI financial advisers for individualized advice	High	Low	High	Low	Low	High	Medium	Low
	AI bot for businesses (eg, SMEs[2]) to track targets	Medium	Low	High	Low	Medium	High	Medium	High
Concision	Mining financial reports to derive important insights	Low	High	Low	Low	Low	Medium	Low	Low
	Detect/prevent fraud by aggregating/interpreting payment documentation	High	Low	High	Medium	Low	High	Low	High
Coding	Model risk management (eg, testing, review, documentation)	Low	Low	Low	Medium	High	High	Low	High
	Reduce tech delivery timelines via automated coding and testing	Low	Medium	Low	High	High	High	Low	Medium
Creative content	Personalized content offerings (eg, credit card offers)	High	Low	Medium	High	Low	Medium	Low	Low
	Automate contract drafting	Medium	High	Medium	Low	Low	High	Low	Medium

[1] Intellectual property.
[2] Small and medium-size enterprises.

McKinsey & Company

VERACODE

Thank You