# Security Engineering

## INFR11208 (UG4) // NFR11228 (MSc)

**Daniel W. Woods* and Jingjie Li**
**Email: daniel.woods@ed.ac.uk and** jingjie.li@ed.ac.uk

# Assignment Feedback

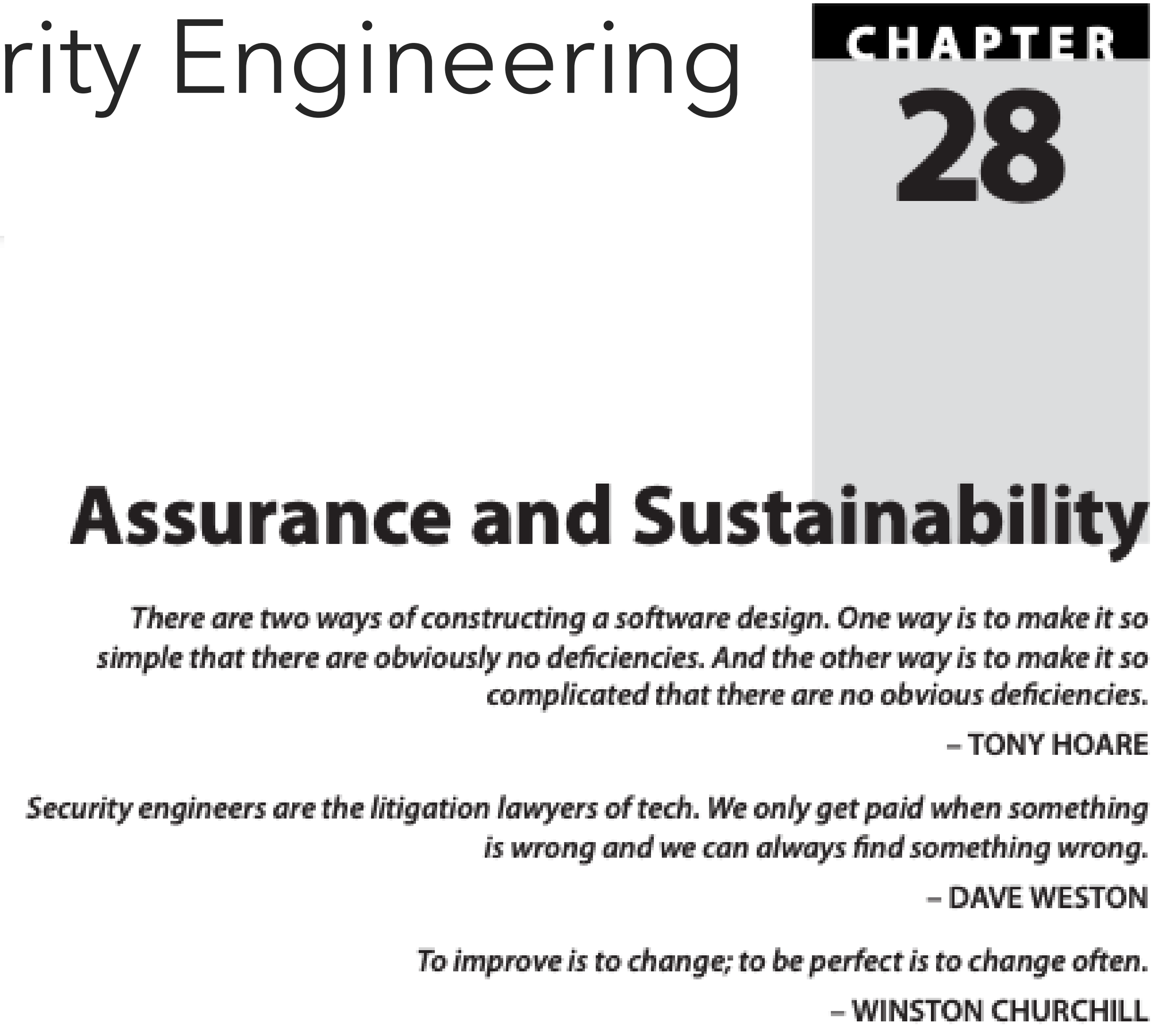
# The family

- Financial fraud & cyber bullying should be top 2
  - Investment fraud (grandparents), romance fraud (the lonely heart)
    - How do fraudsters make contact? Economically motivated using social engineering
    - Do any characteristics increase likelihood?
  - Cyberbullying v common, but mainly an extension of irl bullying. Which platform?
  - Grooming, online abuse, extortion but requires justification
- Too much focus on the router and phone without concrete threats
  - A lot said "traffic interception", but what about TLS? How frequent is DNS hijacking?
  - A lot said the phone will be compromised, but how will exploit be distributed?
  - How does technical infection lead to harm?

# The media company

- Tension between economic crime and targeted attacks
- For targeted attacks, have a clear motivation
  - Espionage is likely, but what's the harm? Targeting specific reporters? As they cross certain borders
  - Link compromise to actual campaigns (e.g. Volt Typhoon & edge devices)
  - DDoS needs a motivation? Also is the damage significant?
- But economically motivated is important too
  - Financial fraud/BEC the most common incident type
  - Ransomware too --> Fortinet device is a big risk vector
- High quality sources provide accurate insights into the world
  - This is really hard to evaluate, but you need to reason about this

# Chapter 8, Security Engineering



CHAPTER 28

## Assurance and Sustainability

*There are two ways of constructing a software design. One way is to make it so simple that there are obviously no deficiencies. And the other way is to make it so complicated that there are no obvious deficiencies.*

– TONY HOARE

*Security engineers are the litigation lawyers of tech. We only get paid when something is wrong and we can always find something wrong.*

– DAVE WESTON

*To improve is to change; to be perfect is to change often.*

– WINSTON CHURCHILL

https://www.cl.cam.ac.uk/archive/rja14/Papers/SEv3-ch28.pdf

# We'll cover three related concepts

- **Assurance:** whether a system will work, and how you’re sure of this.
- **Sustainability:** how long will it work for?
- **Compliance:** how you can satisfy other people of this.

Assurance: whether a system is secure, and **how you’re sure of this**.

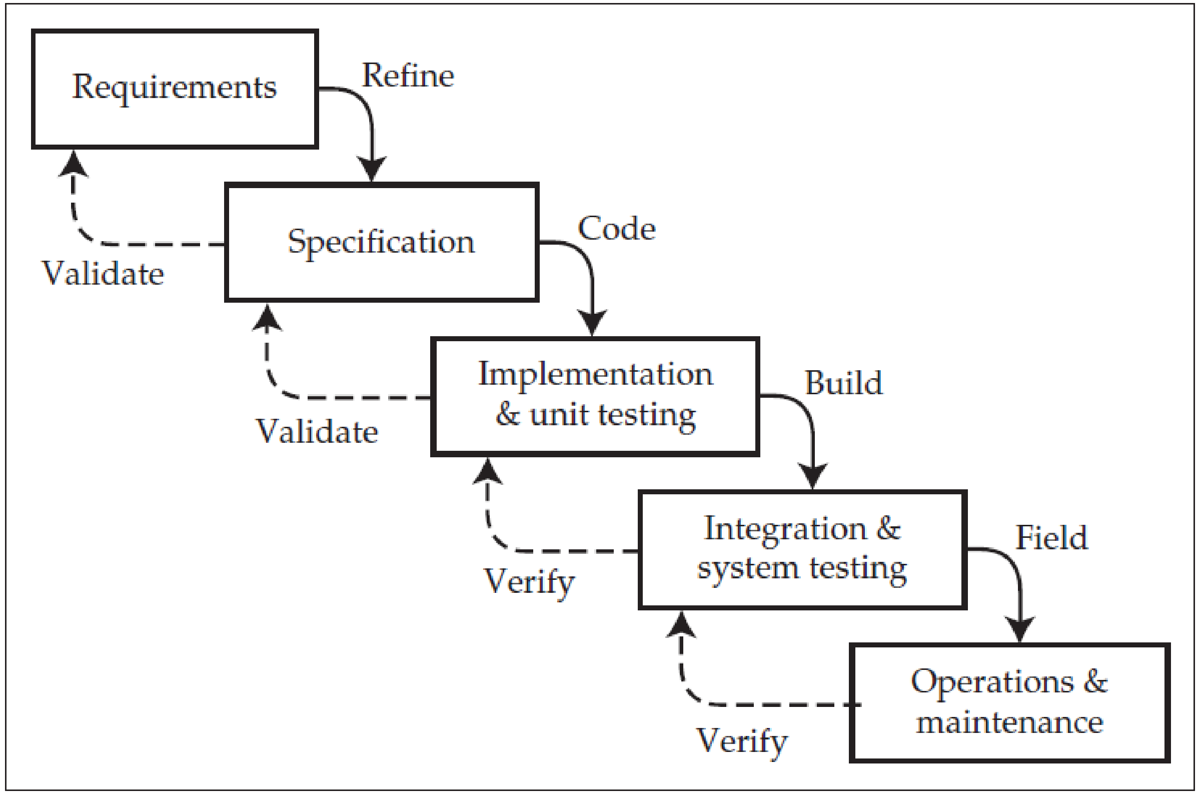
# Routes to assurance

- Building systems
  - Formal methods
  - Testing
  - Bug reports
- Operating systems
  - Monitoring + measuring
    - SOC, ASM, Vuln scanning, TPRM, phishing simulation etc
  - Pen testing/red teaming
  - Standards, checklists etc

# Building Assurance

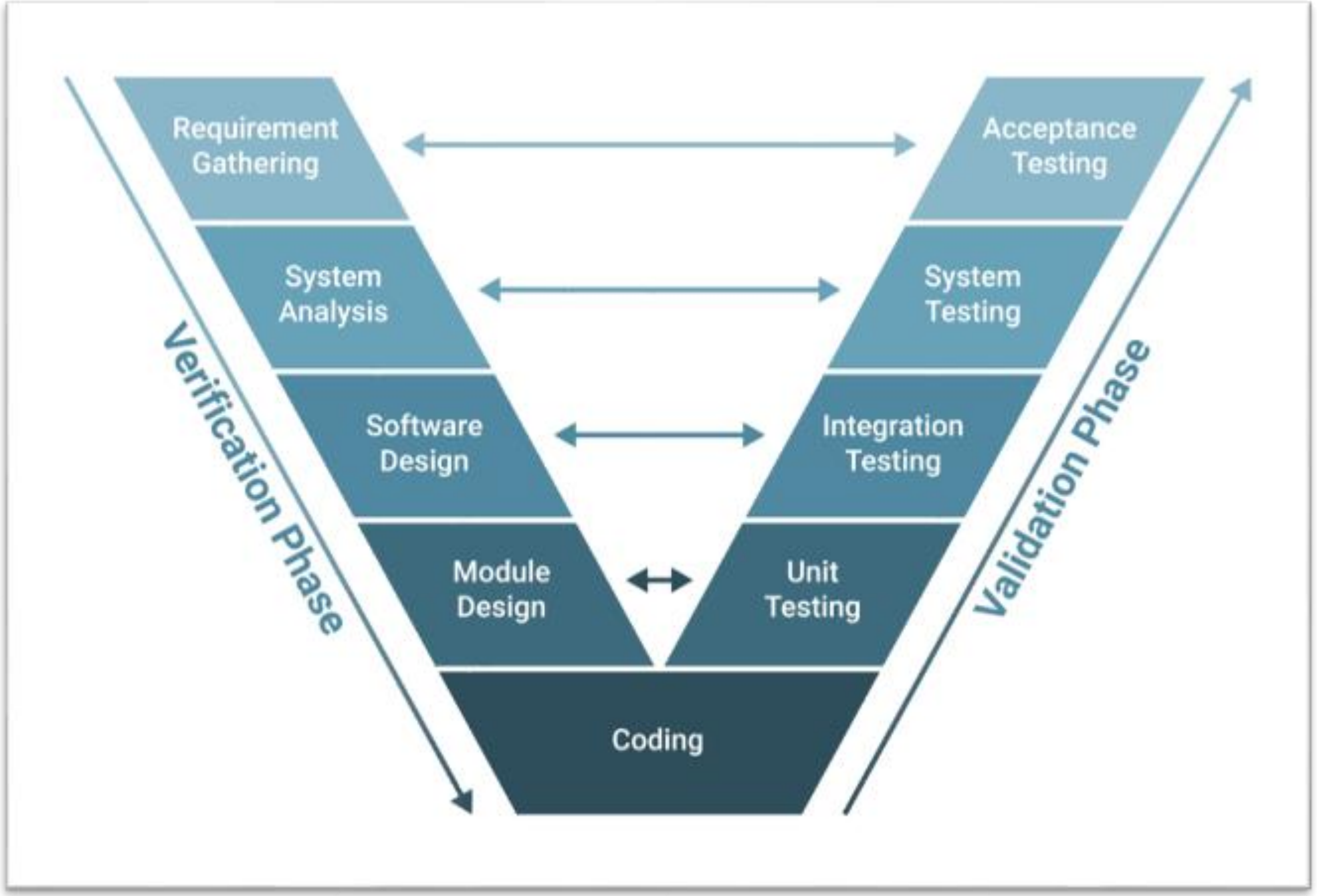
# Waterfall & Formal Methods

- Both share heavy investment in initial design
- Formally represent design and verify security properties
  - Focus on specification
- Used in safety and military systems
  - Also for security protocols (TLS)



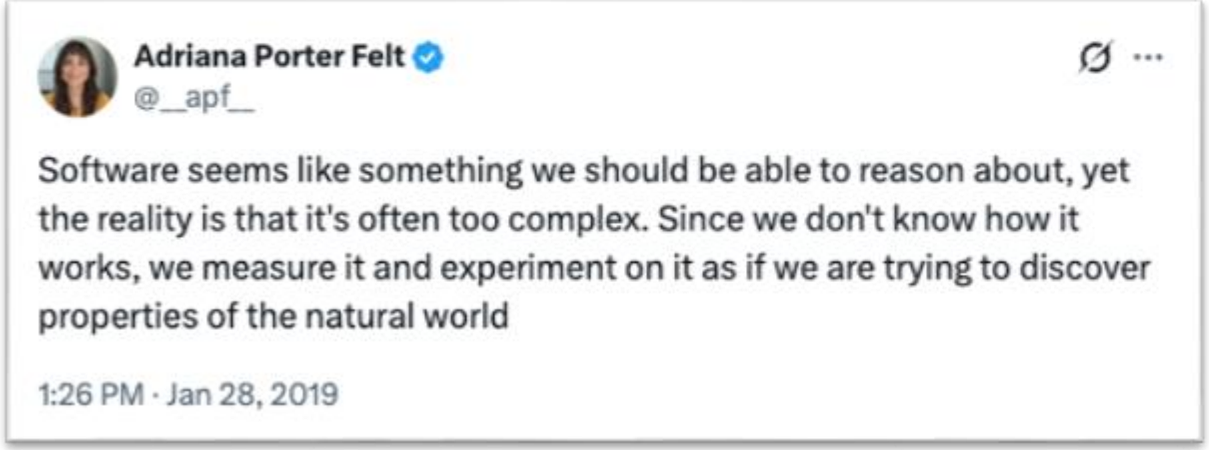
# Testing

- Unit testing for reliability, fuzzing for security
- Static/dynamic analysis
- Internal security audits
- LLM code reviews??
  - See Anthropic
- Finally, user testing



Verification and validation model

# User Testing

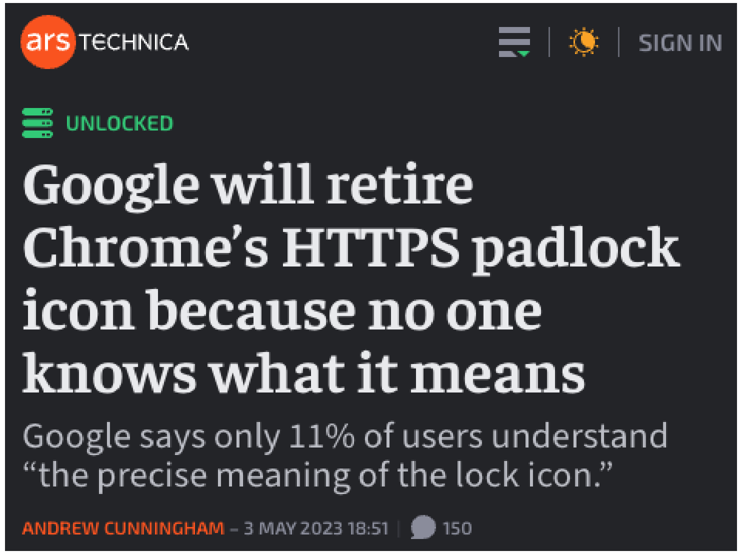


Adriana Porter Felt
@__apf__

Software seems like something we should be able to reason about, yet the reality is that it's often too complex. Since we don't know how it works, we measure it and experiment on it as if we are trying to discover properties of the natural world

1:26 PM · Jan 28, 2019

DEVELOPMENT
ITERATION #1
DESIGN
TESTING
ITERATION #2
ITERATION #3
ITERATION #4
ITERATION #5
ITERATION #6



ars TECHNICA
SIGN IN

UNLOCKED

Google will retire Chrome's HTTPS padlock icon because no one knows what it means

Google says only 11% of users understand "the precise meaning of the lock icon."

ANDREW CUNNINGHAM – 3 MAY 2023 18:51 | 150

"Our research in 2021 showed that **only 11% of study participants correctly understood the precise meaning of the lock icon**. This misunderstanding is not harmless — nearly all phishing sites use HTTPS, and therefore also display the lock icon."

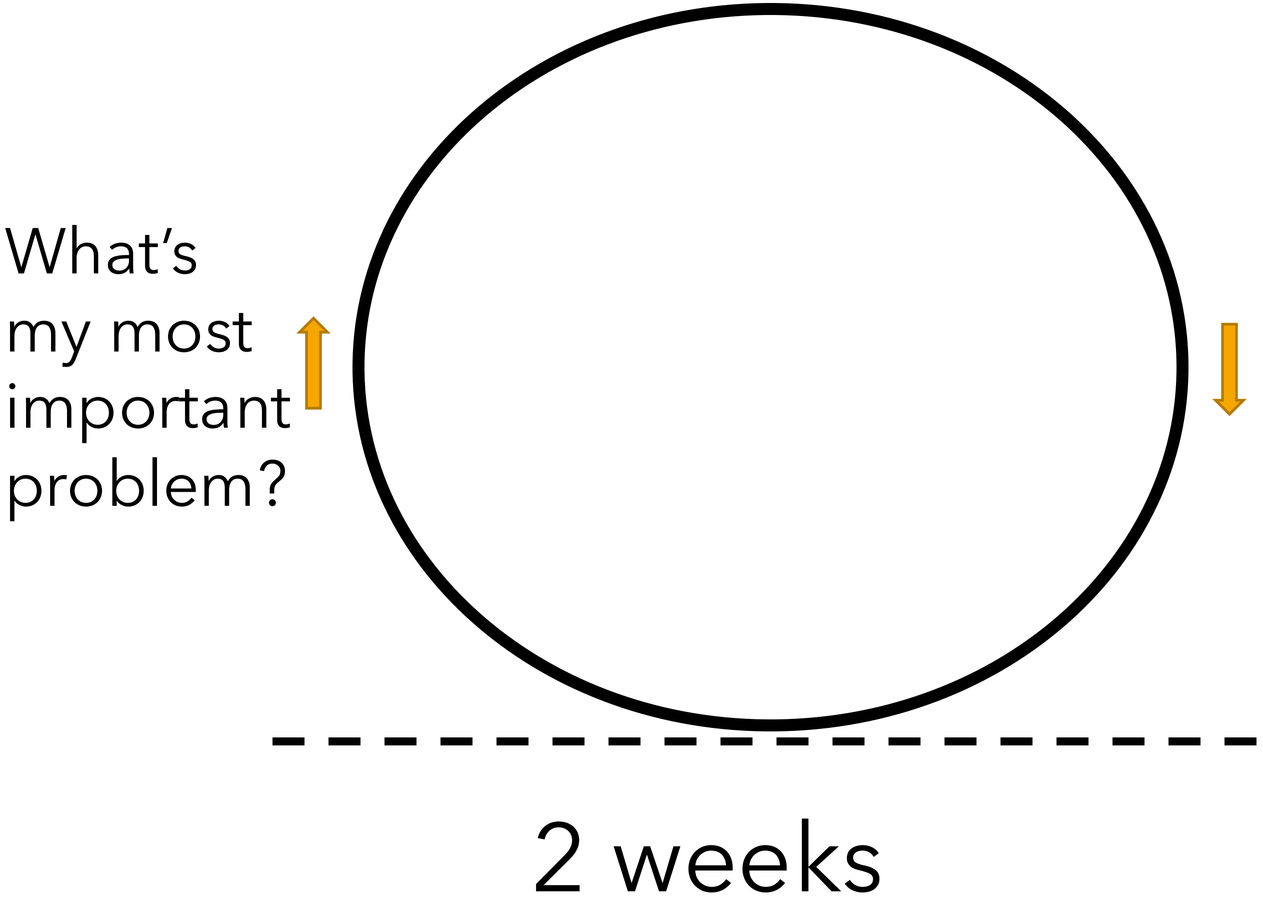
# Assurance in agile developments

**Cons**

- Past assurance breaks as you add new requirements
- Hard to maintain verification and validation
  - Need automation

**Pros**

- Easier to integrate feedback
  - Vulnerabilities reported to you!
  - Insights from user testing

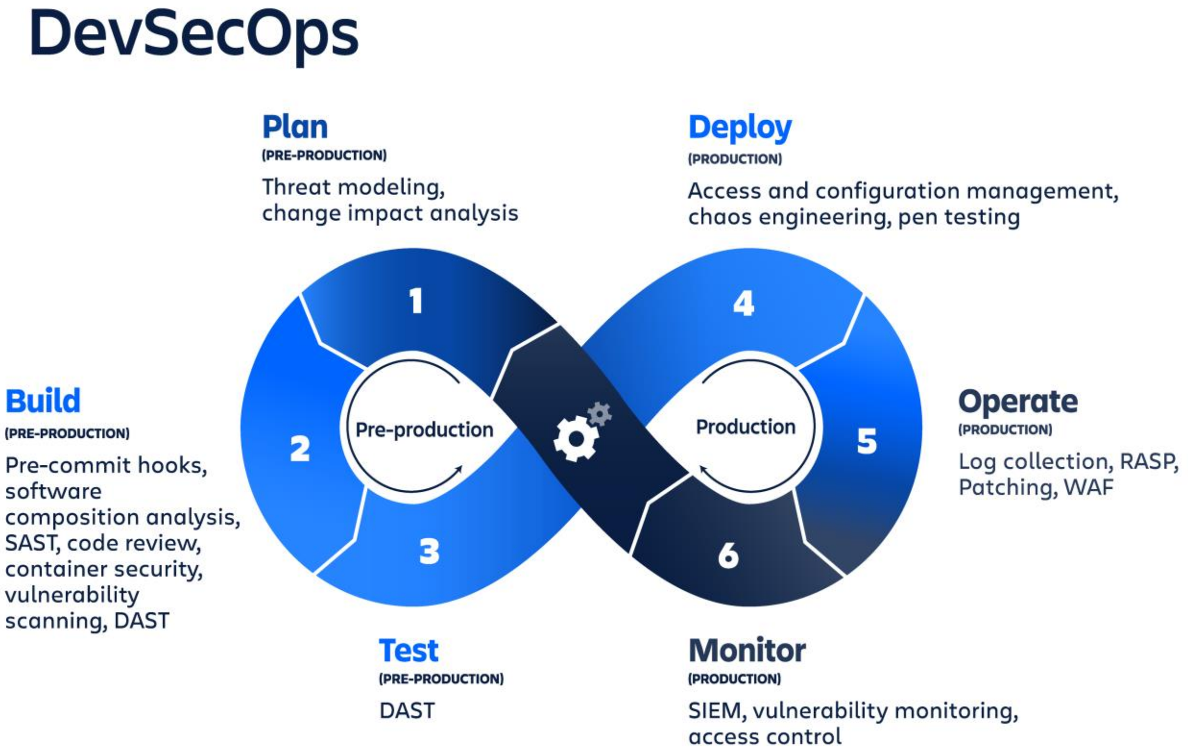


What's my most important problem?

2 weeks

# "Shift left" in secure software development

- Secure by design
  - Memory safe languages, secure templates
  - Train developers on secure coding
- Automated code review
  - SAST scans for patterns like creds or SQL injection
  - SCA generate Software Bill of Materials and check for CVEs
- Continuous monitoring
  - DAST & fuzzing to identify run-time issues
  - Vulnerability scanning, SIEM etc
- Test deployment
  - Cloud security scans, patch management etc
- Vulnerability disclosure policy



DevSecOps

**Plan** (PRE-PRODUCTION)
Threat modeling, change impact analysis

**Deploy** (PRODUCTION)
Access and configuration management, chaos engineering, pen testing

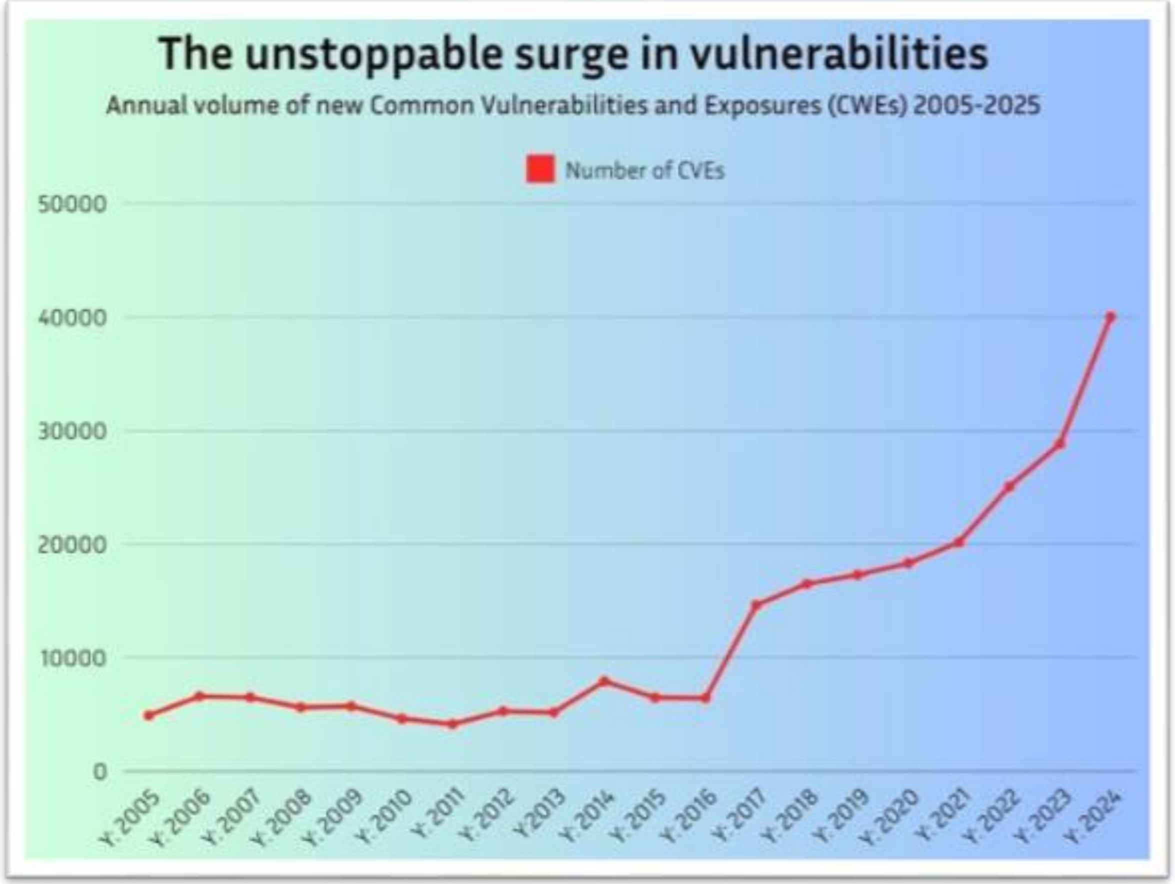
**Build** (PRE-PRODUCTION)
Pre-commit hooks, software composition analysis, SAST, code review, container security, vulnerability scanning, DAST

**Operate** (PRODUCTION)
Log collection, RASP, Patching, WAF

**Test** (PRE-PRODUCTION)
DAST

**Monitor** (PRODUCTION)
SIEM, vulnerability monitoring, access control

# The challenge of bug reports



# Patching is non-trivial

"17 out of the 41 in-the-wild 0-days from 2022 are variants of previously reported vulnerabilities."

**Did you fix the root cause?**

**Source:** https://security.googleblog.com/2023/07/the-ups-and-downs-of-0-days-year-in.html

# Getting Your Disclosure Policy Right

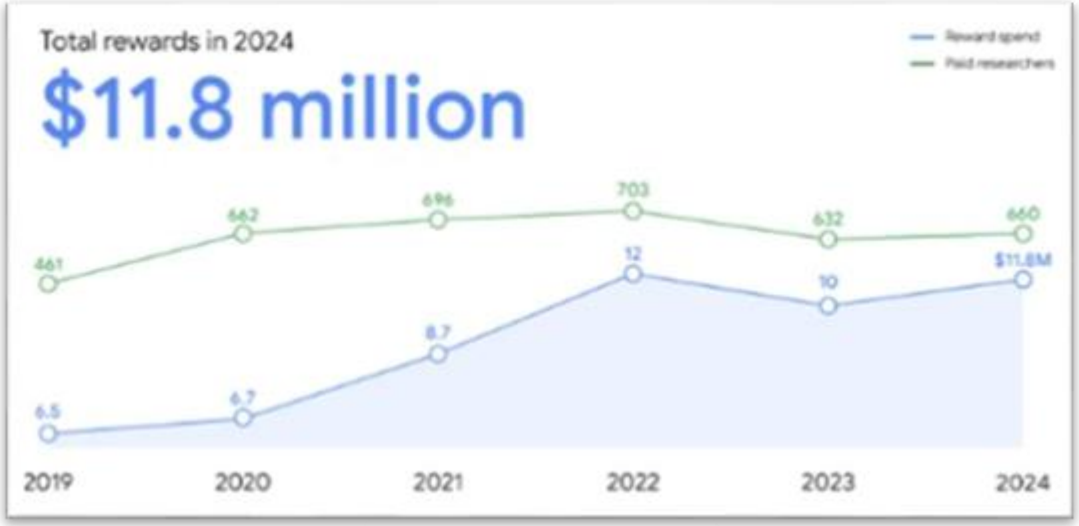
- How do you get people to report bugs to you before they disclose publicly?
- How do you avoid researchers selling your bugs to someone else?
- How do you get your CEO to not deny/deflect?

# Types of Disclosure Policies

- Bug bounty policy
  - offer financial rewards
- Vulnerability disclosure policy
  - offer legal guarantees you won't sue
- Security contact via security.txt

**Empirical Study of AI Vendors:**

- 40% have BBP
- 16% have VDP
- 8% have security contact
- 36% have no policy

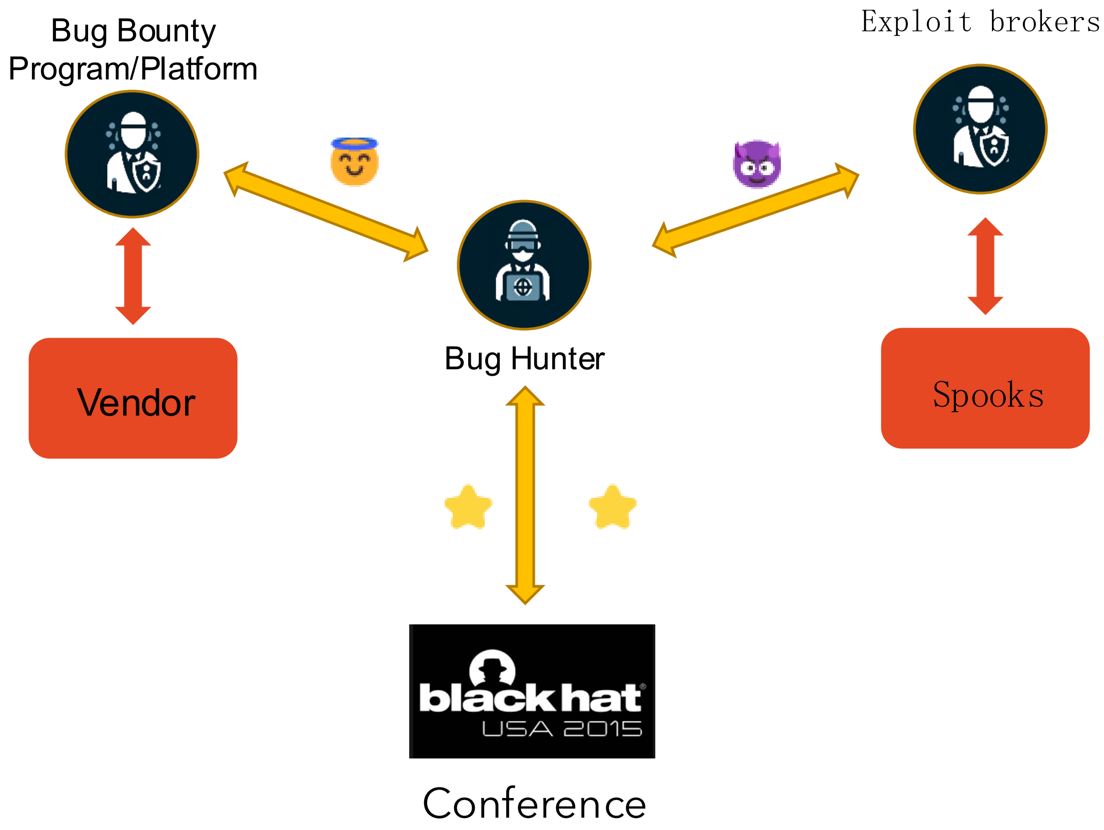


**Source:** https://www.mediapost.com/publications/article/404028/google-paid-118m-in-bug-bounties-in-2024.html

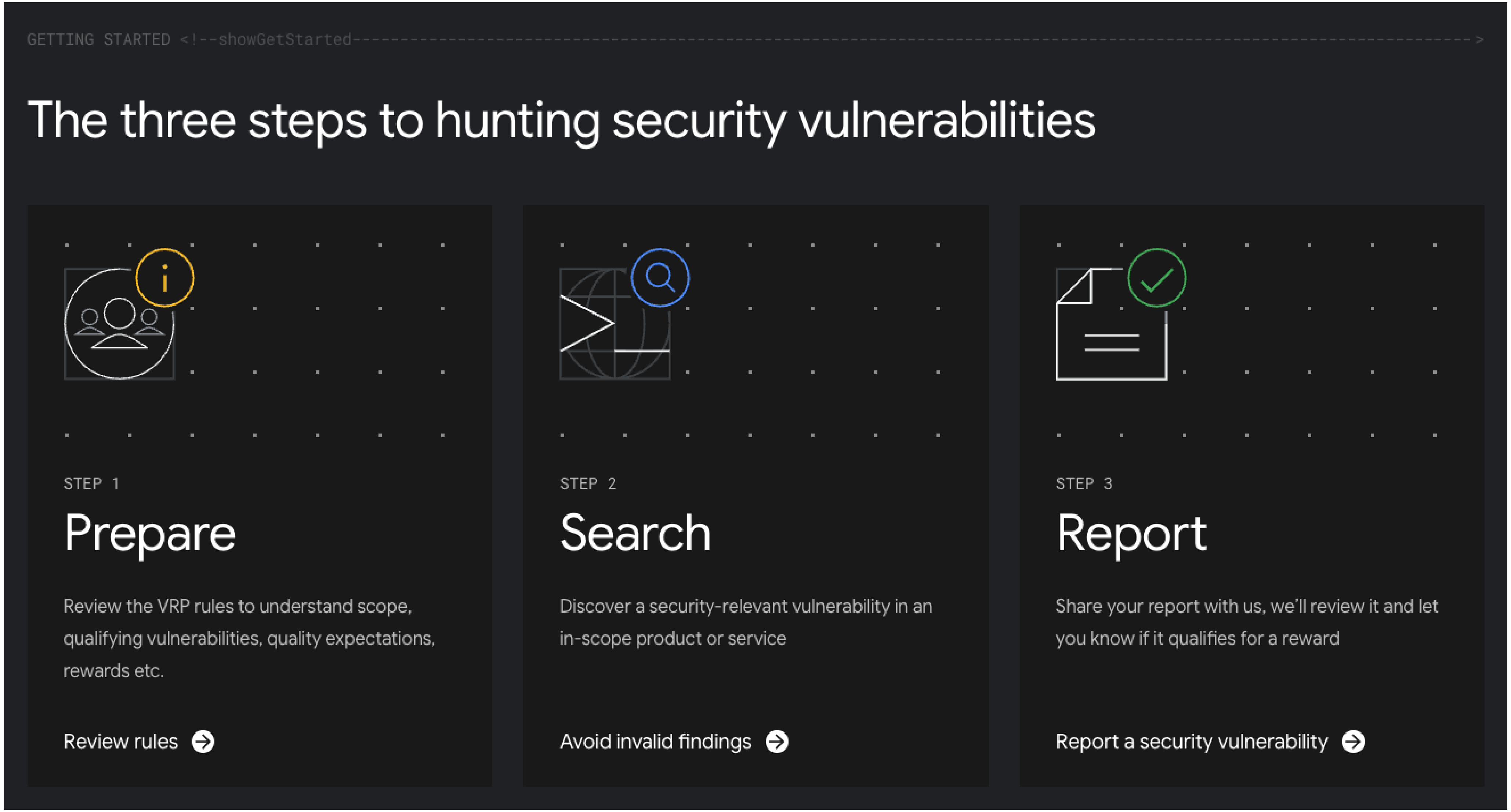
**Source:** Piao, Yangheran, Jingjie Li, and Daniel W. Woods. "" Abuse Risks are Often Inherent to Product Features": Exploring AI Vendors' Bug Bounty and Responsible Disclosure Policies." arXiv preprint arXiv:2509.06136 (2025).

# How to make sure the report comes to you

- Pay enough
- Give them credit too
- Don't require a working exploit
- Reduce legal risk



# https://bughunters.google.com/



GETTING STARTED <!--showGetStarted------------------------------------------------------------------------------------->

## The three steps to hunting security vulnerabilities

STEP 1

### Prepare

Review the VRP rules to understand scope, qualifying vulnerabilities, quality expectations, rewards etc.

Review rules →

STEP 2

### Search

Discover a security-relevant vulnerability in an in-scope product or service

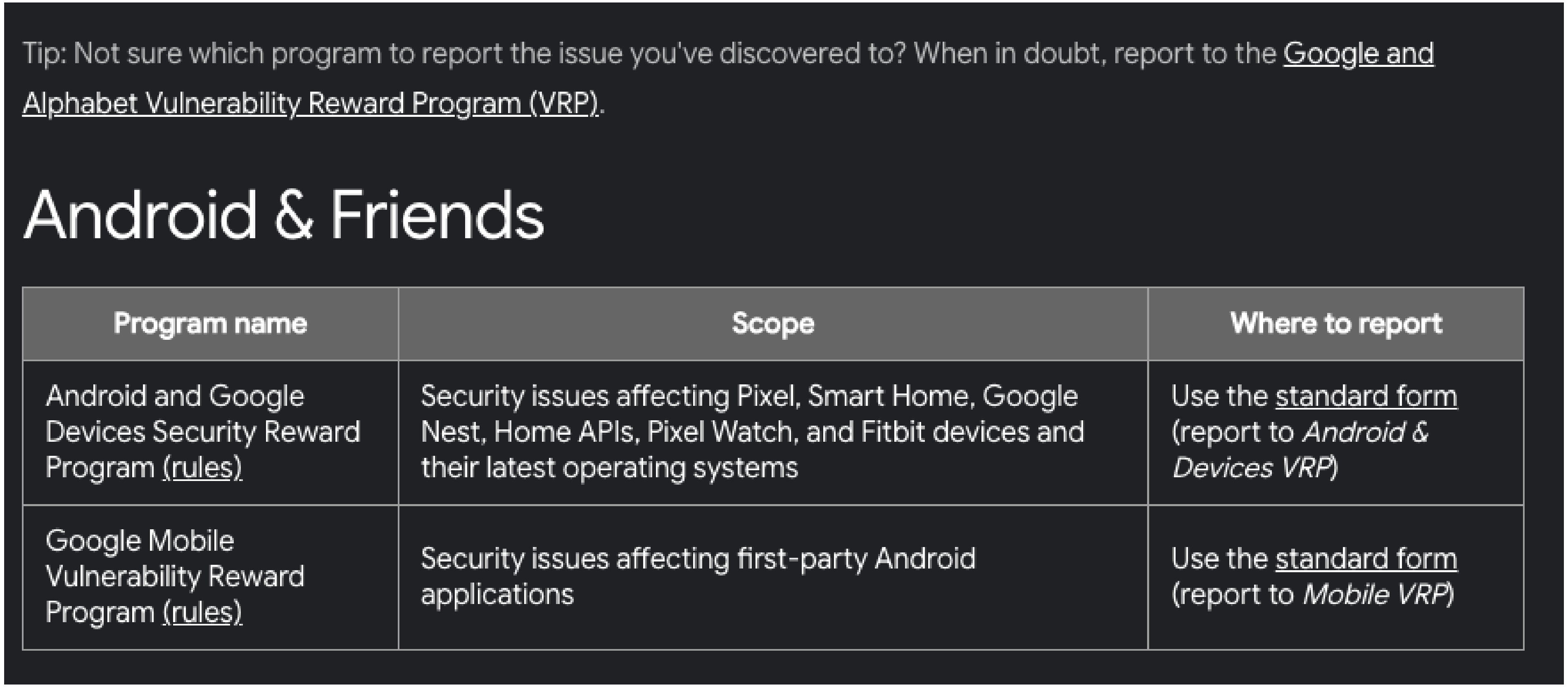
Avoid invalid findings →

STEP 3

### Report

Share your report with us, we'll review it and let you know if it qualifies for a reward

Report a security vulnerability →



Tip: Not sure which program to report the issue you've discovered to? When in doubt, report to the Google and Alphabet Vulnerability Reward Program (VRP).

## Android & Friends

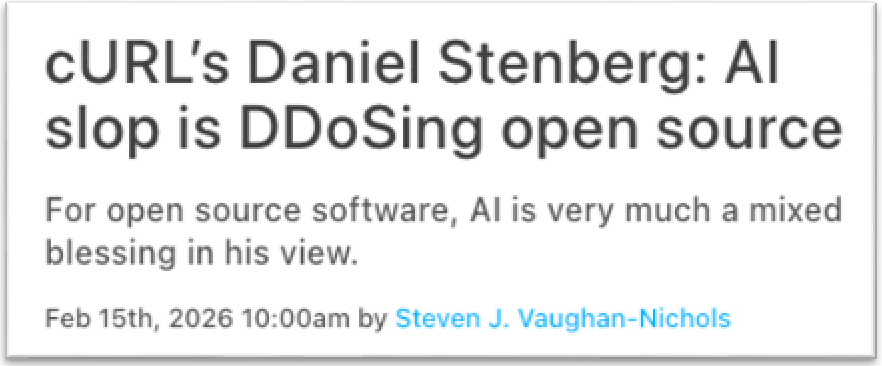
| Program name | Scope | Where to report |
|---|---|---|
| Android and Google Devices Security Reward Program (rules) | Security issues affecting Pixel, Smart Home, Google Nest, Home APIs, Pixel Watch, and Fitbit devices and their latest operating systems | Use the standard form (report to *Android & Devices VRP*) |
| Google Mobile Vulnerability Reward Program (rules) | Security issues affecting first-party Android applications | Use the standard form (report to *Mobile VRP*) |

"In general, we reward critical and high severity vulnerabilities ...
A few classes of vulnerabilities exist that generally do not qualify for a reward:
•Phishing attacks that involve tricking the user into entering credentials.
•Issues that only affect userdebug builds.
•Bugs that simply cause an app to crash.
•Issues with negligible security impact, as described in Bug Hunter University, with some exceptions."

# Managing bug reports is hard

"In ride sharing, the task is clear (drive the customer from location A to location B), and the driver can reliably estimate the time to completion. In contrast, ambiguity in bug bounty policies–like the description of qualifying bugs–leaves hackers uncertain about how long it will take to find a specific vulnerability or how much they might receive"

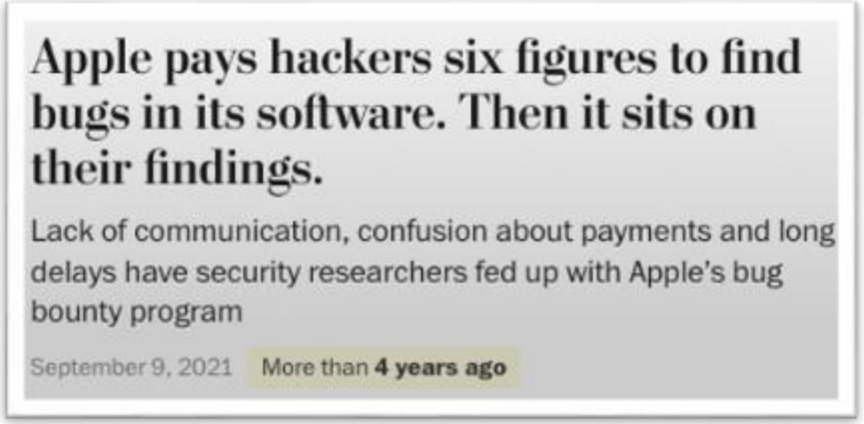
**Source:** Piao, Y & Woods, DW 2025, The Bug Bounty Manifesto: Collectivization and Fairness. in Security Protocols XXIX: 29th International Workshop. Lecture Notes in Computer Science, Springer, Twenty-ninth International Workshop on Security Protocols, Cambridge, United Kingdom, 26/03/25.



cURL's Daniel Stenberg: AI slop is DDoSing open source

For open source software, AI is very much a mixed blessing in his view.

Feb 15th, 2026 10:00am by Steven J. Vaughan-Nichols



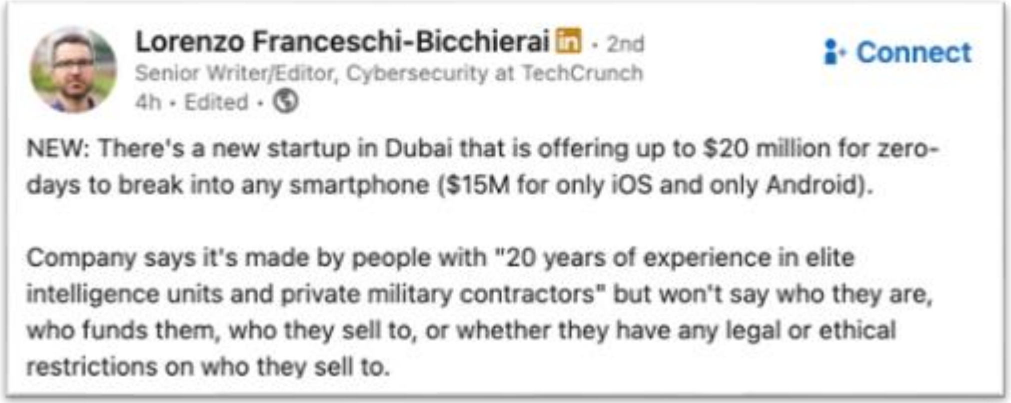
Apple pays hackers six figures to find bugs in its software. Then it sits on their findings.

Lack of communication, confusion about payments and long delays have security researchers fed up with Apple's bug bounty program

September 9, 2021 More than **4 years ago**

# So then what is assurance for builders?

- System is so complex you know there are vulnerabilities
- … maybe it's enough to say exploits are getting more expensive?
- Most vendors fall a long way short
  - Lots of room for improvement here!
- What does it matter if your users will be spear phished anyway



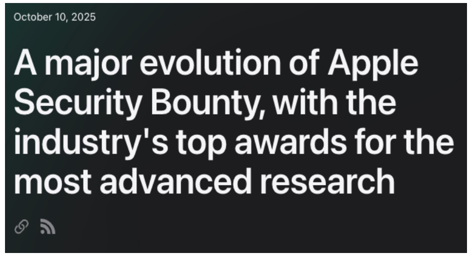
Lorenzo Franceschi-Bicchierai · 2nd

Senior Writer/Editor, Cybersecurity at TechCrunch

4h · Edited ·

NEW: There's a new startup in Dubai that is offering up to $20 million for zero-days to break into any smartphone ($15M for only iOS and only Android).

Company says it's made by people with "20 years of experience in elite intelligence units and private military contractors" but won't say who they are, who funds them, who they sell to, or whether they have any legal or ethical restrictions on who they sell to.



October 10, 2025

A major evolution of Apple Security Bounty, with the industry's top awards for the most advanced research

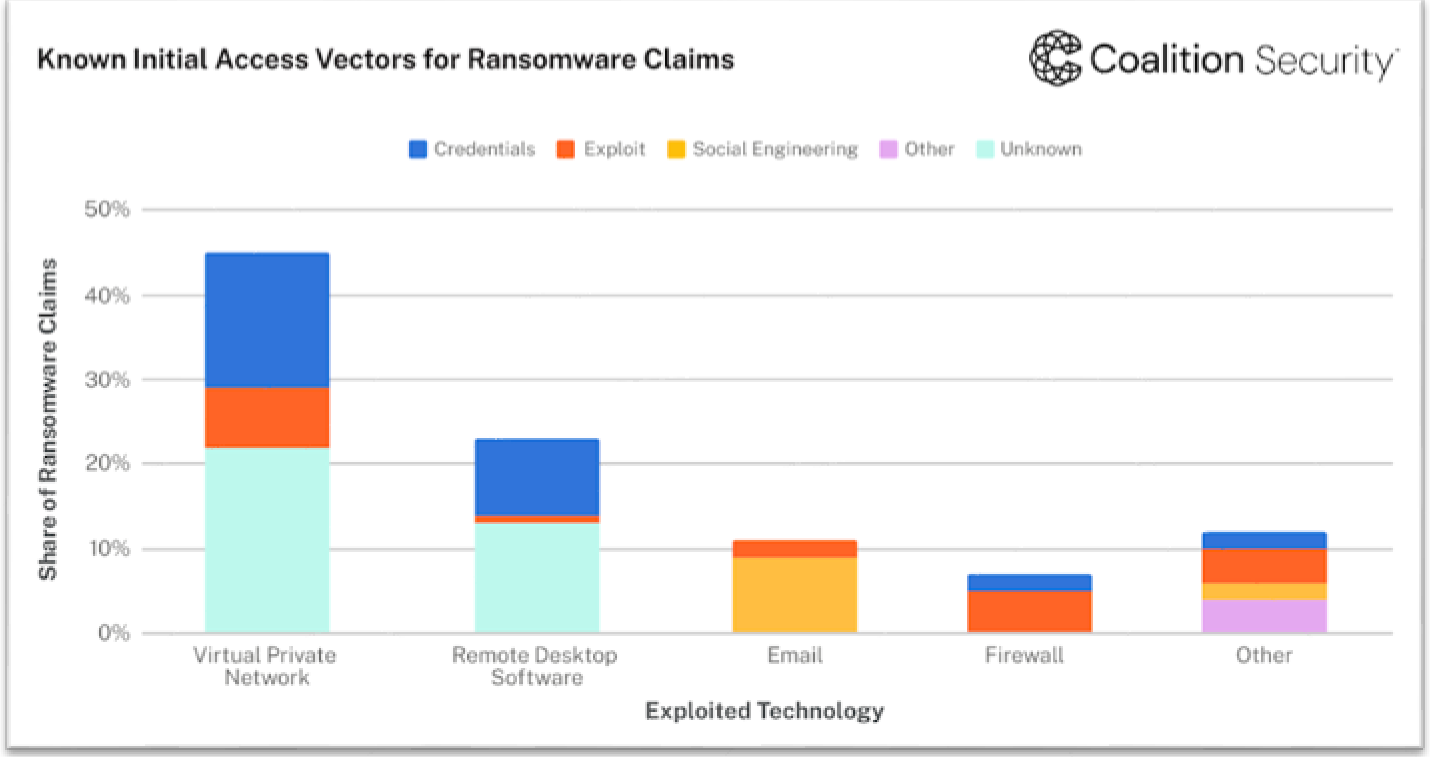
"We're doubling our top award to **$2 million** for exploit chains that can achieve similar goals as sophisticated mercenary spyware attacks."

Follows announcement of Memory Integrity Enforcement.

# Deploying Assurance

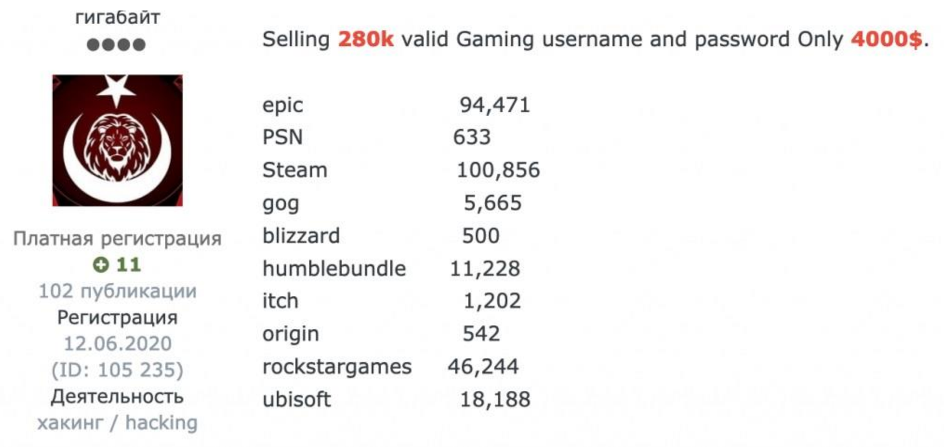
# What to verify, test, and monitor

1. **Credential security**
2. Vulnerabilities
3. User awareness
4. Signs of intrusion
   - Malware
   - Data exfiltration
   - Access requests
   - Etc



# How to be sure passwords are secure

- Passwords are inherently **hard to control**
  - Set and "stored" by the user
  - Could be re-used on other insecure apps or on personal devices
  - Could be leaked
- Policies try to address often lead to **unintended consequences**
  - Write down passwords
  - Re-use same password on multiple services



гигабайт

●●●●

Платная регистрация

⊕ 11

102 публикации

Регистрация

12.06.2020

(ID: 105 235)

Деятельность

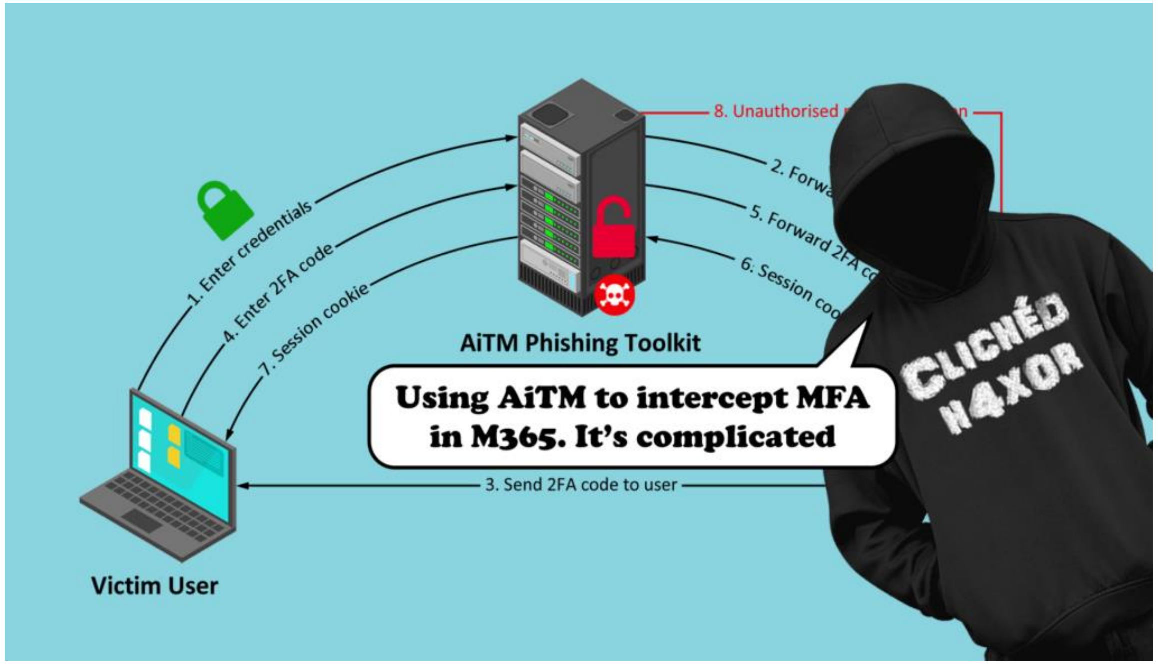
хакинг / hacking

Selling **280k** valid Gaming username and password Only **4000$**.

| | |
|---|---|
| epic | 94,471 |
| PSN | 633 |
| Steam | 100,856 |
| gog | 5,665 |
| blizzard | 500 |
| humblebundle | 11,228 |
| itch | 1,202 |
| origin | 542 |
| rockstargames | 46,244 |
| ubisoft | 18,188 |

# How to be sure 2nd factors are secure

- SMS-based better than passwords
  - Fresh codes better than stale passwords
  - ... maybe the code is phished or SIM swapped
- App-based is better still
  - At least you don't rely on a Telecomm firm
  - … maybe the code was phished
- FIDO designed to avoid this
  - Tied to a device
  - If user logs into go0gle.com, won't auth
- **Better MFA tends to be costlier**
  - Onboarding + credential reset

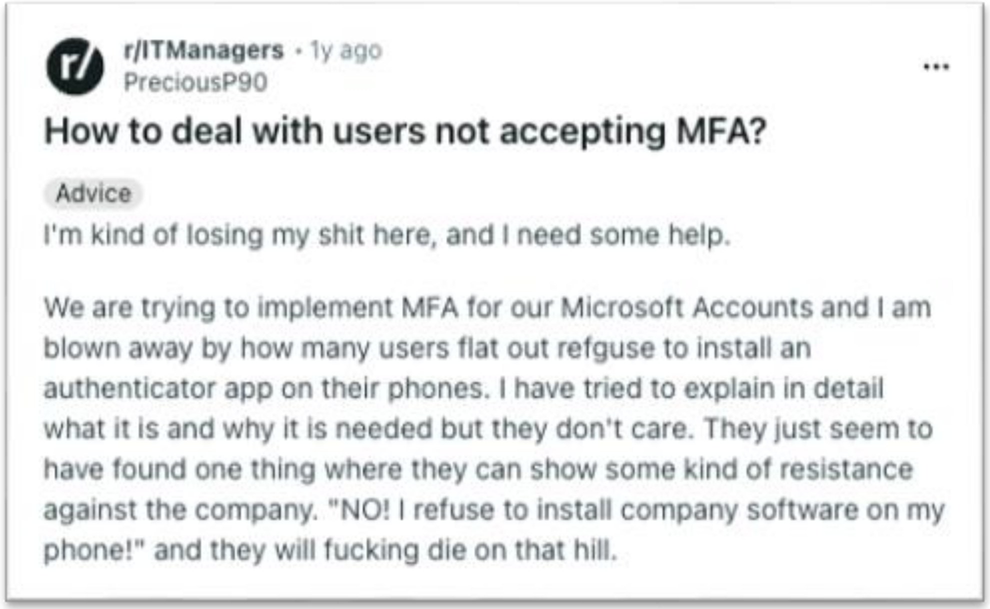


**Source:** https://www.pentestpartners.com/security-blog/intercepting-mfa-phishing-and-attackers-in-the-middle/

# Configuring identity security is hard

- Not every service has MFA enabled
  - "63% of organizations that stated they used MFA for protecting privileged access were still impacted by a ransomware event linked to coverage gaps in MFA configuration/installation method"
- Not every account enforces MFA
  - Service accounts
  - Users who refuse
  - Before users enrol

Source: Identity Has Become the Prime Target of Threat Actors



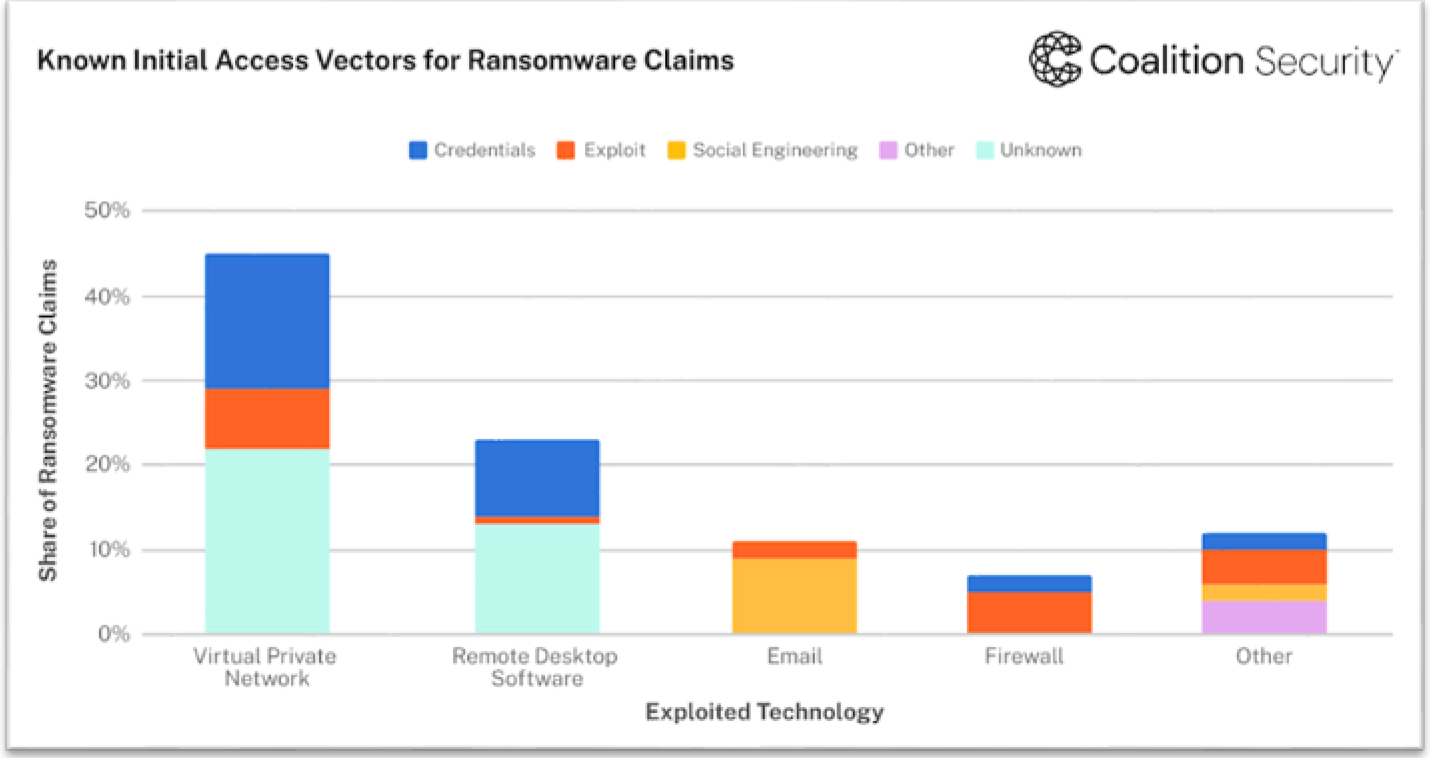
**How to deal with users not accepting MFA?**

I'm kind of losing my shit here, and I need some help.

We are trying to implement MFA for our Microsoft Accounts and I am blown away by how many users flat out refguse to install an authenticator app on their phones. I have tried to explain in detail what it is and why it is needed but they don't care. They just seem to have found one thing where they can show some kind of resistance against the company. "NO! I refuse to install company software on my phone!" and they will fucking die on that hill.

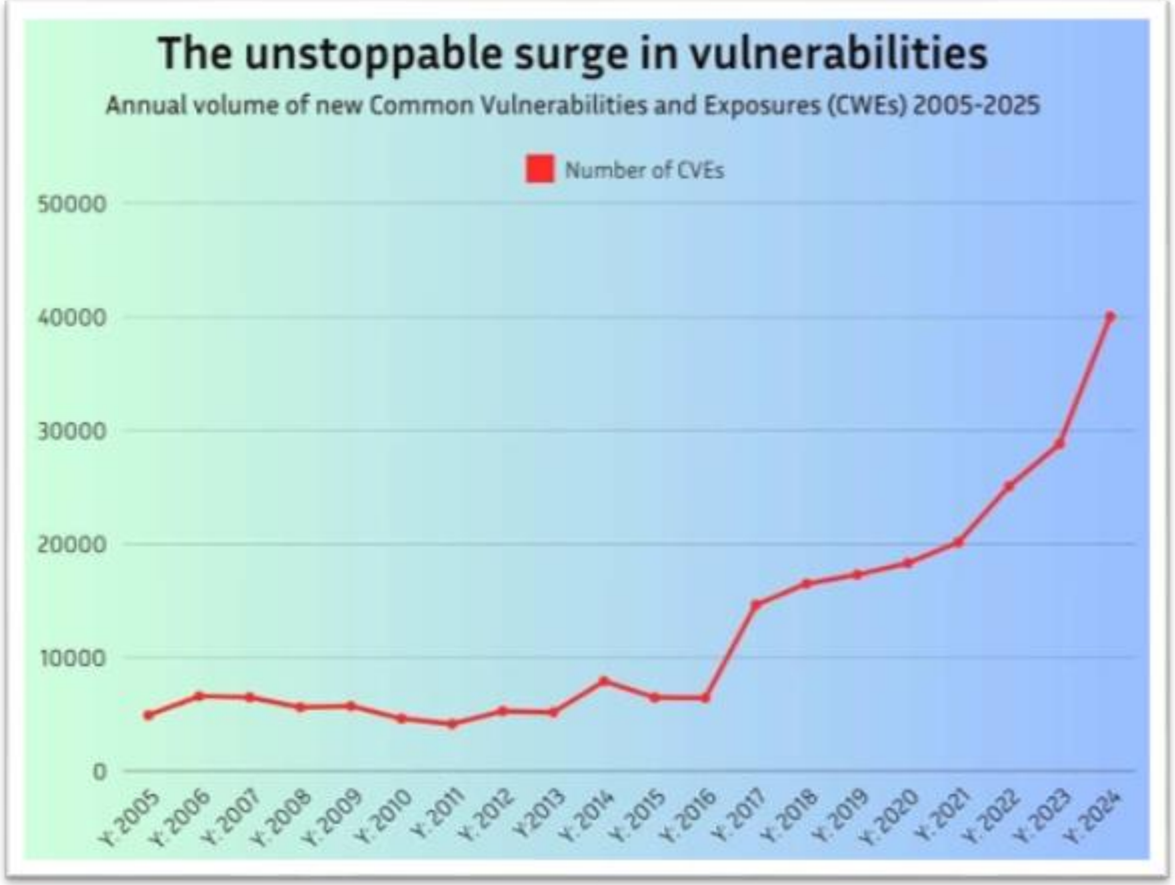
# What to verify, test, and monitor

1. Credential security
2. **Vulnerabilities**
3. User awareness
4. Signs of intrusion
   - Malware
   - Data exfiltration
   - Access requests
   - Etc



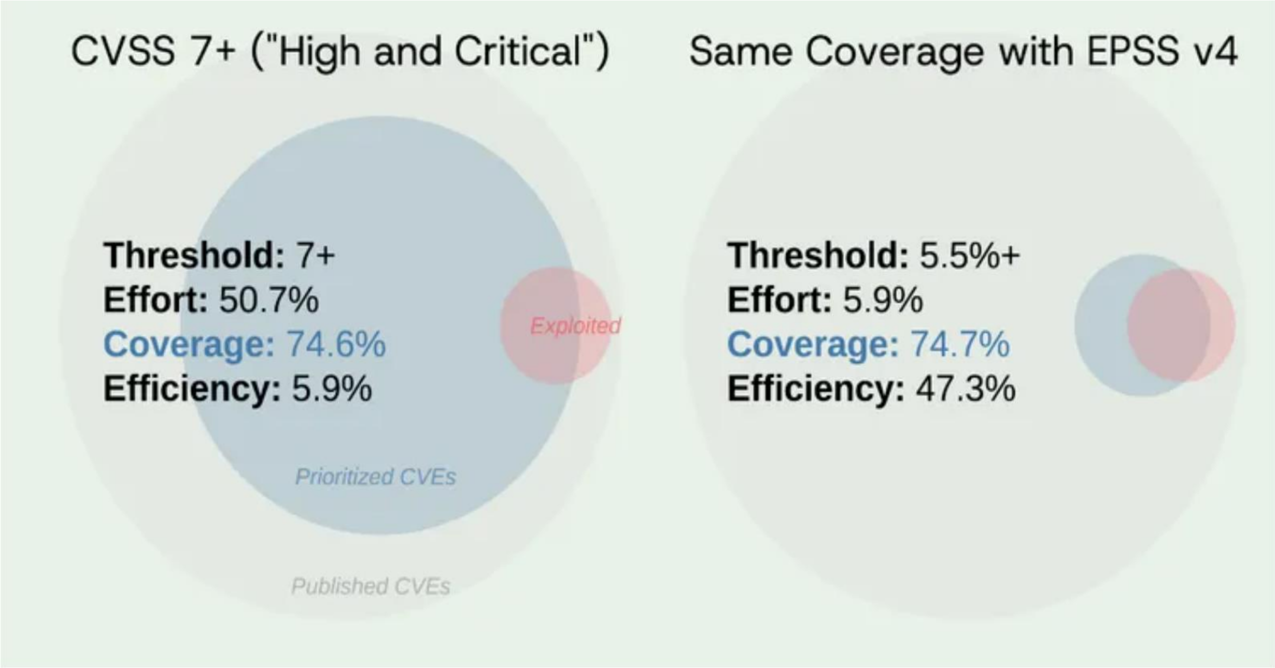
# Patch Management

- If vulnerability reports are bad for vendors, they're even worse for customers
  - o Most use **far more** products than they build
  - o Microsoft produces 1 patch, and 1 billion plus customers need to apply it



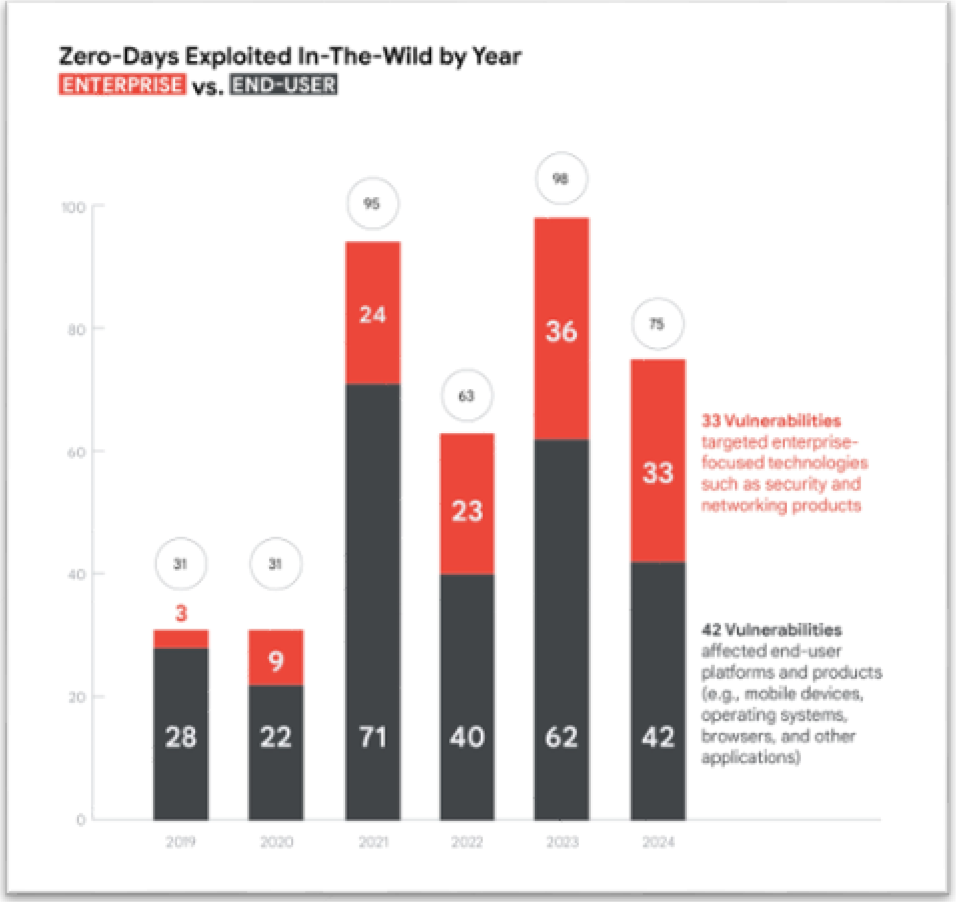
# How sure are you about your patch strategy

1. **PCI DSS:** Update high and critical vulnerabilities (according to CVSS) within one month of release
   - Miss many exploited vulns
2. **EPSS:** Patch what ML says is most likely to be exploited.
   - Miss less exploited vulns, maybe patch before n-day exploits.
3. **KEV:** Patch what has already been exploited.
   - Efficient but maybe you are the threat intelligence!
4. **Cloud infra:** Rely on professionals.
   - Until there's an 0day



# Uncertainties in patch management

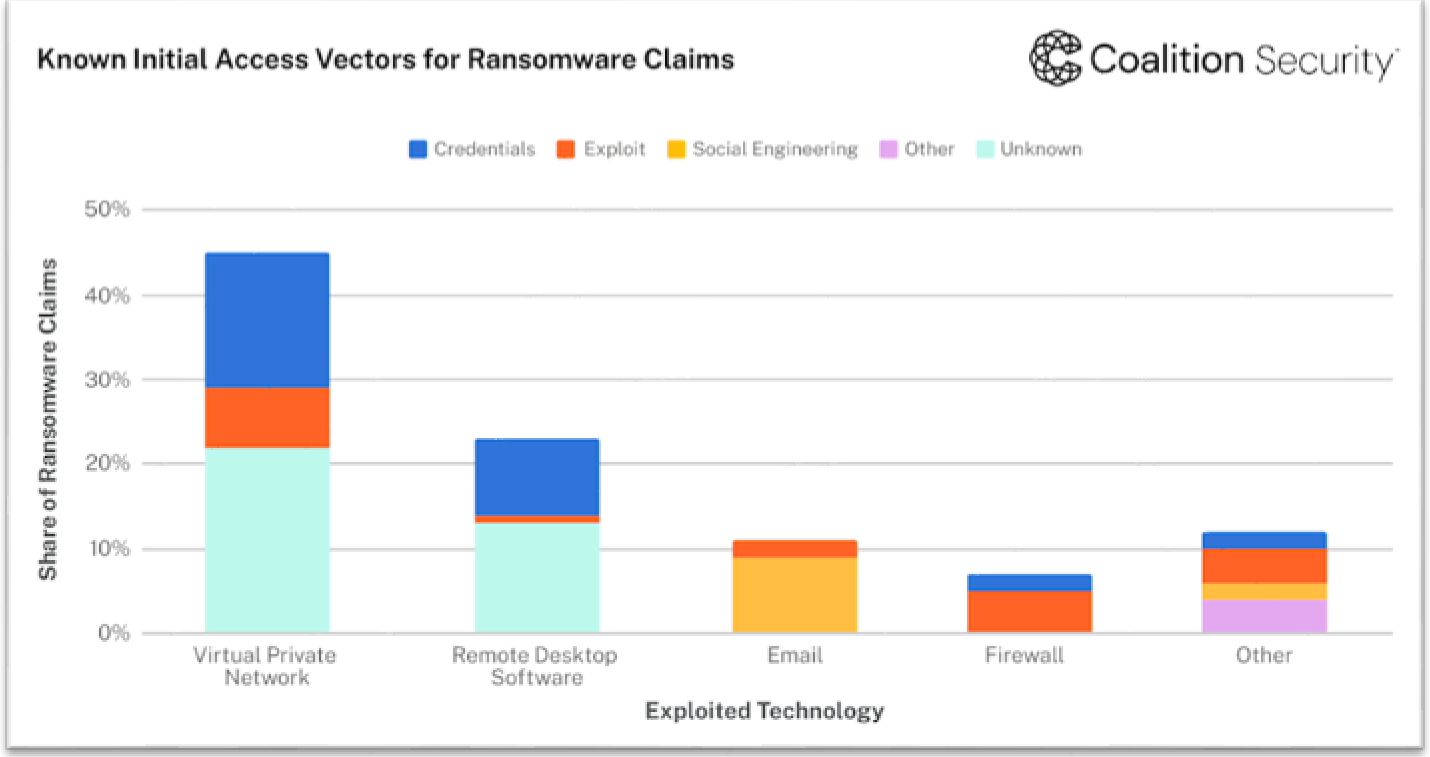
- 0-days
  - Google Project Zero are tracking
- The patch doesn't work
  - 41% of 0days in 2022 were variants on known vulns
- Shadow IT + user-managed devices
  - ASM and vulnerability scanning



**Source:** https://security.googleblog.com/2023/07/the-ups-and-downs-of-0-days-year-in.html

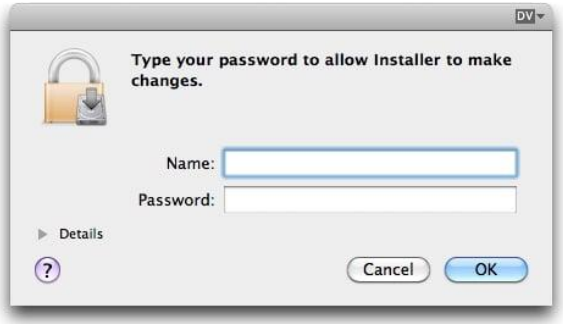
# What to verify, test, and monitor

1. Credential security
2. Vulnerabilities
3. **User awareness**
4. Signs of intrusion
   - Malware
   - Data exfiltration
   - Access requests
   - Etc



# How are you sure about your users?

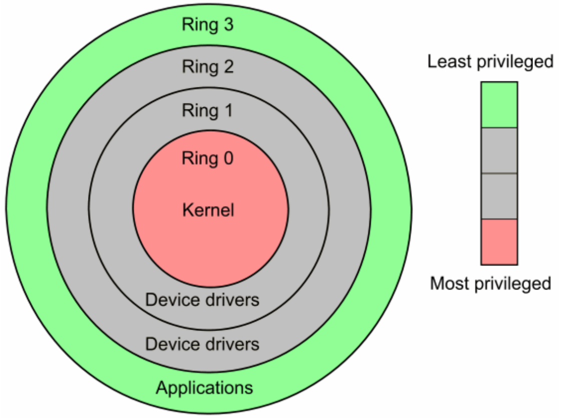
- You can't be!
  - Errors at every point of the stack!
  - Training helps with knowledge-based behaviour
- Reduce the "blast radius"
  - Application white lists
  - Disable autorun
  - Principle of least privilege
  - Phishing resistant MFA
  - Etc etc

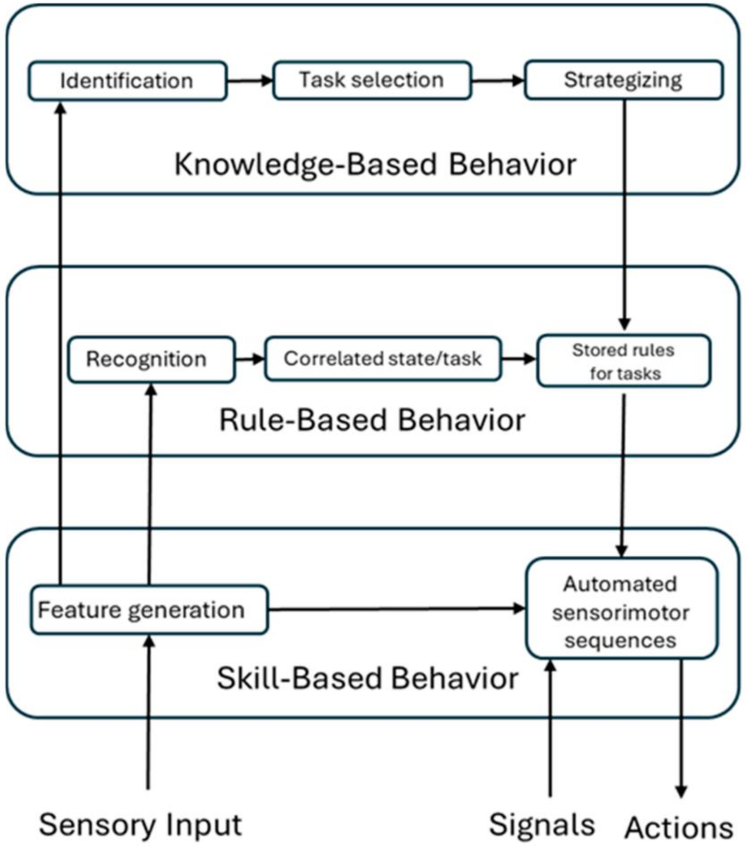


Type your password to allow Installer to make changes.

Name:

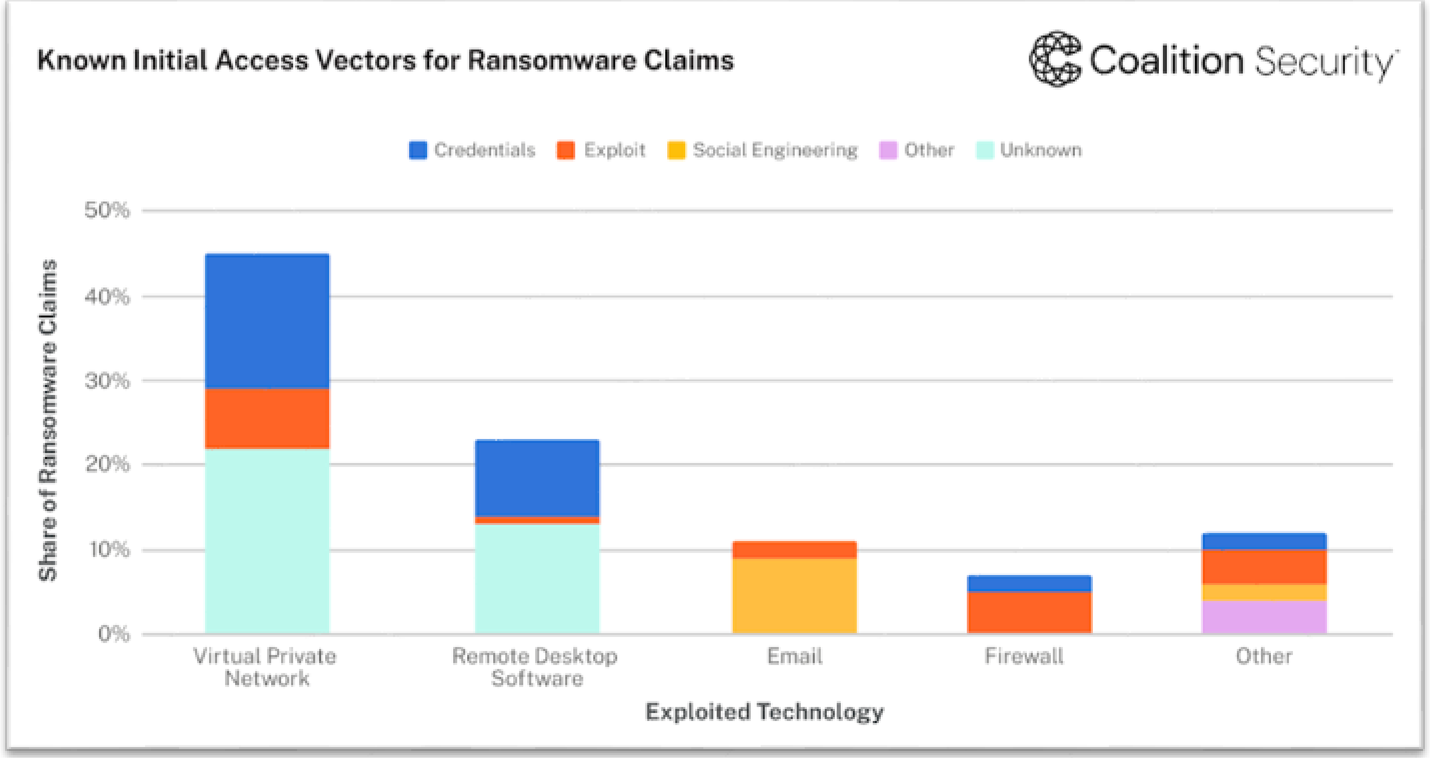
Password:





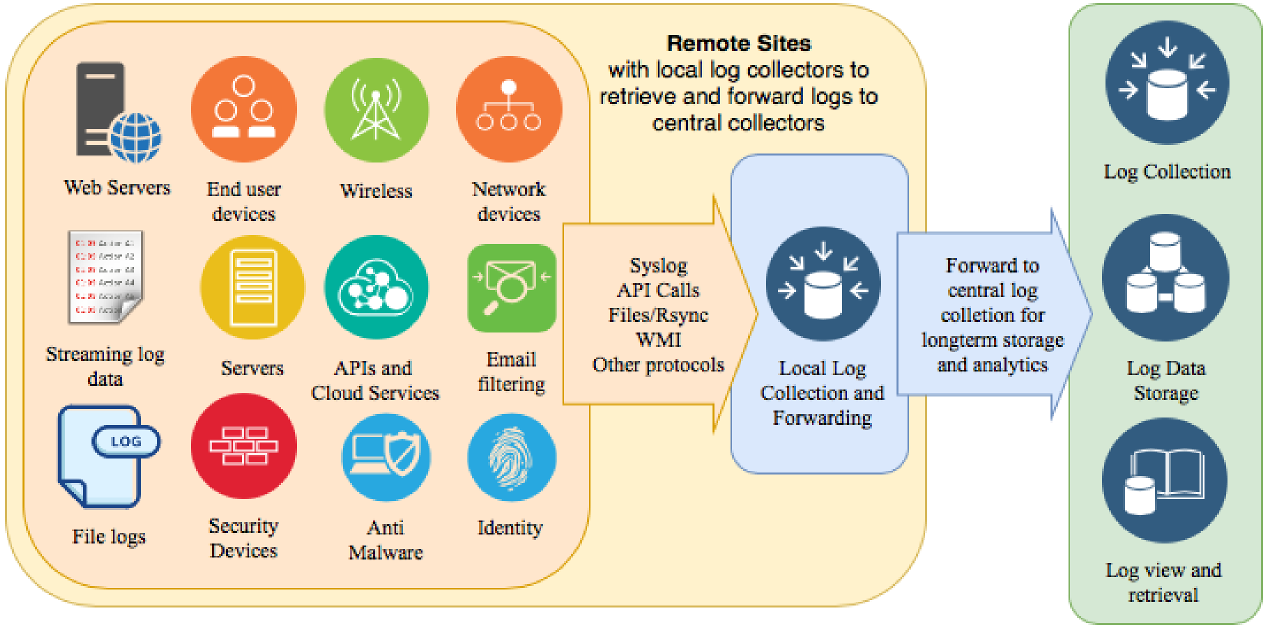
# What to verify, test, and monitor

1. Credential security
2. Vulnerabilities
3. User awareness
4. **Signs of intrusion**
    - Malware
    - Data exfiltration
    - Access requests
    - Etc

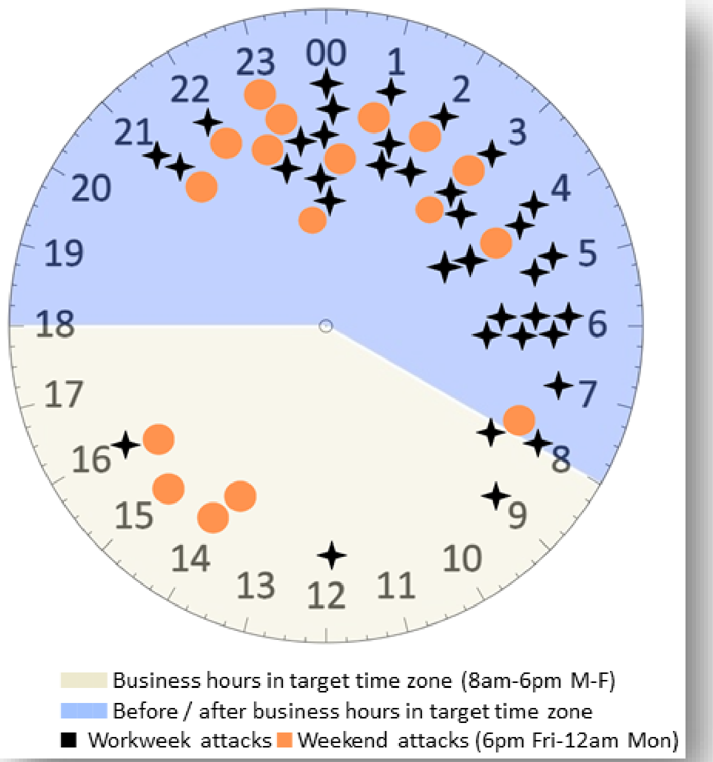


# Monitoring and Response

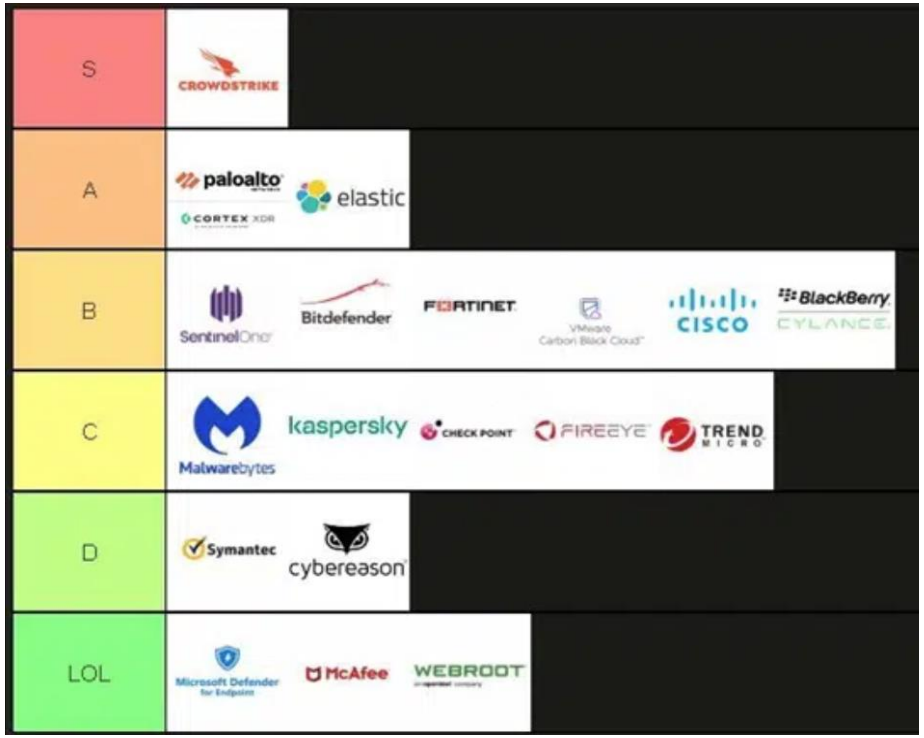
- HUGE amount if InfoSec focuses on **monitoring logs**:
  - Security Operations Centre
  - Endpoint Detection & Response
    - MDR, ITDR etc
  - Data loss prevention
  - Threat Hunting
- Huge cat and mouse game
- Provides an **extra "safety net"**
  - In addition to assurance in prevention controls, you can detect and contain compromise



# Evidence that monitoring works



**Source:** https://www.darktrace.com/blog/im-sorry-were-closed-why-most-ransomware-attacks-happen-out-of-hours



| | |
|---|---|
| S | CROWDSTRIKE |
| A | paloalto CORTEX XDR<br>elastic |
| B | SentinelOne<br>Bitdefender<br>FORTINET<br>VMware Carbon Black Cloud<br>CISCO<br>BlackBerry CYLANCE |
| C | Malwarebytes<br>kaspersky<br>CHECK POINT<br>FIREEYE<br>TREND MICRO |
| D | Symantec<br>cybereason |
| LOL | Microsoft Defender for Endpoint<br>McAfee<br>WEBROOT |

Attackers discuss EDR tools, and invest in bypasses.

**Source:** Conti leaks.

**MALWARE-FREE ACTIVITY**

»

| | |
|---|---|
| 75% | 2023 |
| 71% | 2022 |
| 62% | 2021 |
| 51% | 2020 |
| 40% | 2019 |

**Source:** https://go.crowdstrike.com/rs/281-OBQ-266/images/GlobalThreatReport2024.pdf

Sustainability: **how long** will the system be secure for?

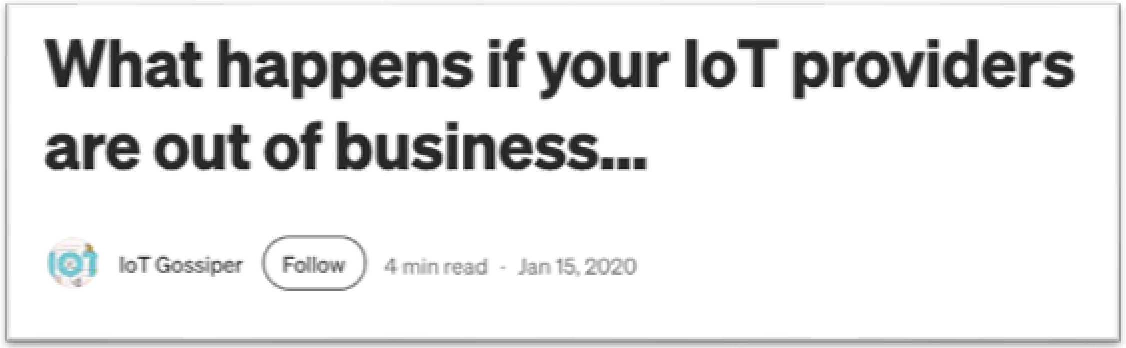
# Return on Investment of releasing patches

**Why vendors release patches**

- Your customers demand security updates
- May be reputation backlash

**Why they might not**

- Planned obsolescence helps business
- Running a patch program is expensive
  - Triage takes time
  - Paying researchers ($10m+ a year for main vendors)
  - Developers need to work on security patches



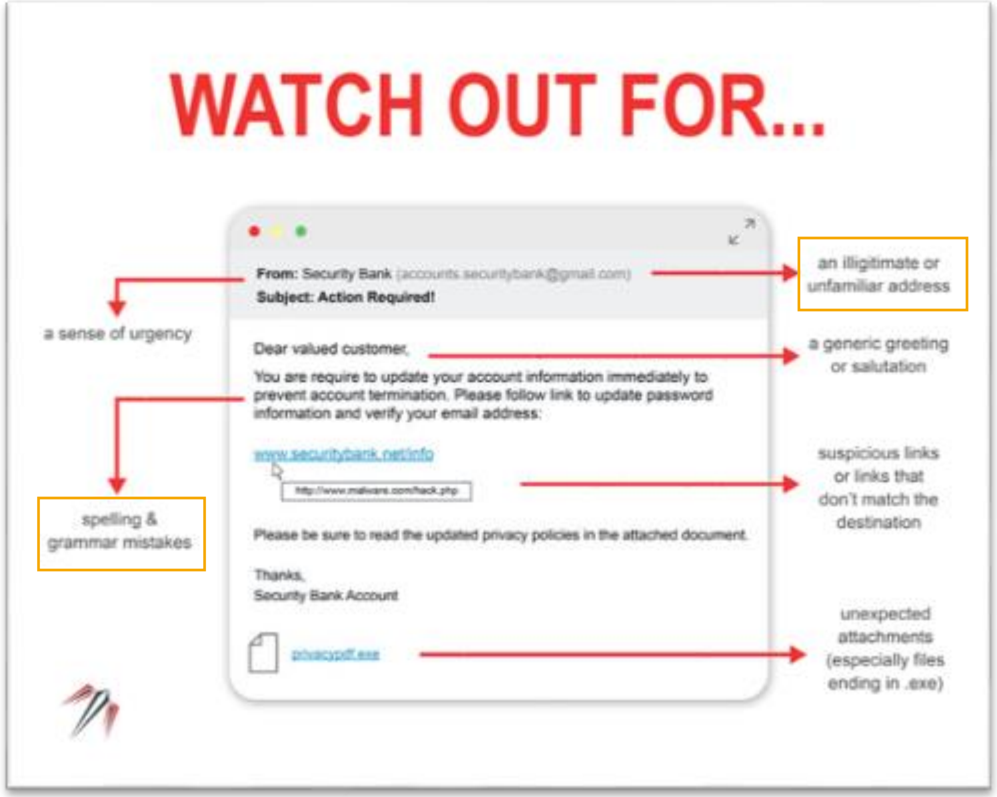
**What happens if your IoT providers are out of business…**

IoT Gossiper Follow 4 min read · Jan 15, 2020

**Windows 10 Case Study**

- Windows 10 used by 42.8% of all Windows computer users worldwide.
- Windows 8 ended in Jan 2016, only 3.7% of Windows users were still using it.
- Only 2.2% were still using Windows 8.1 when support ended in January 2023.
- Many of the computers still running Windows 10 can’t upgrade to Windows 11.

# Changing Threat Actor Capabilities



WATCH OUT FOR...

**From:** Security Bank (accounts.securitybank@gmail.com)
**Subject: Action Required!**

Dear valued customer,

You are require to update your account information immediately to prevent account termination. Please follow link to update password information and verify your email address:

www.securitybank.net/info

http://www.malware.com/hack.php

Please be sure to read the updated privacy policies in the attached document.

Thanks,
Security Bank Account

privacypdf.exe

an illigitimate or unfamiliar address

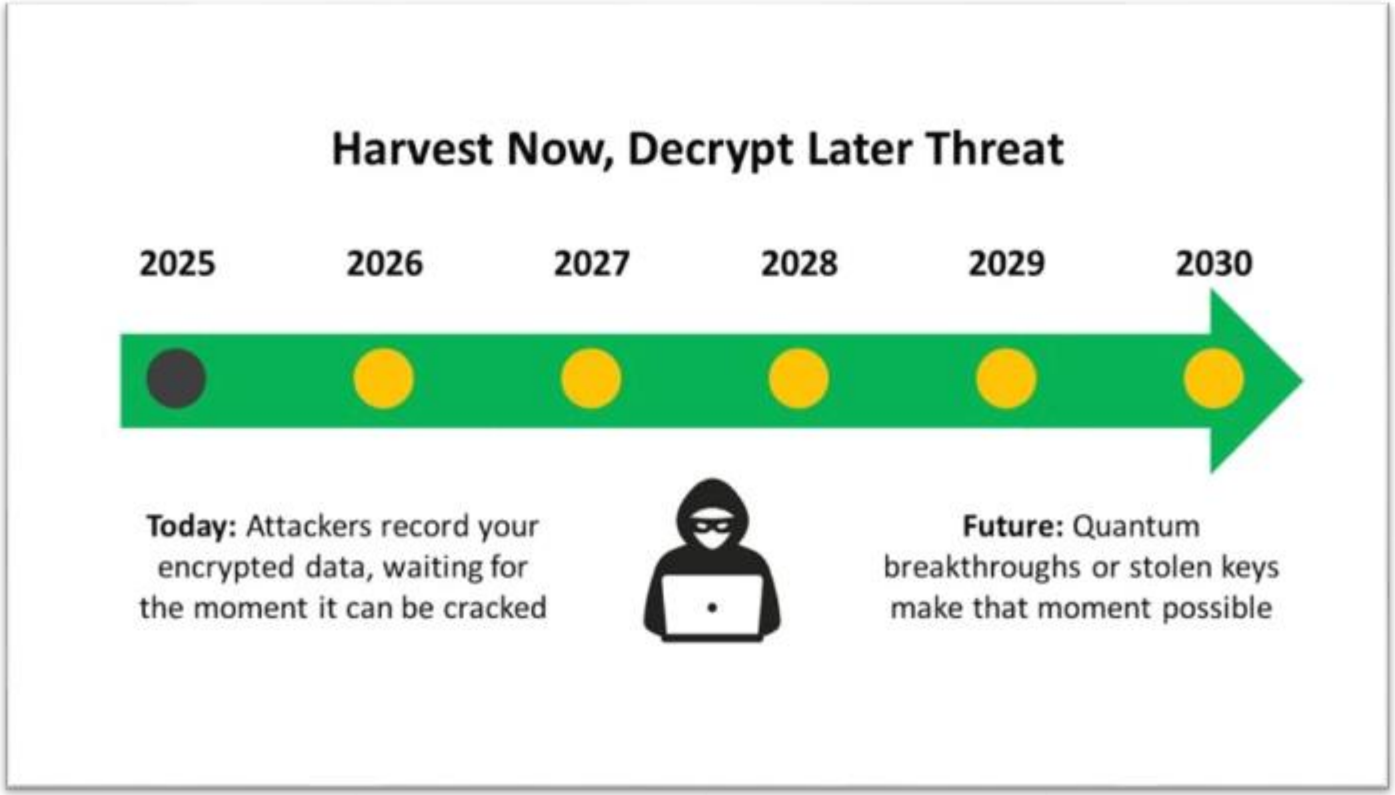
a sense of urgency

a generic greeting or salutation

spelling & grammar mistakes

suspicious links or links that don't match the destination

unexpected attachments (especially files ending in .exe)



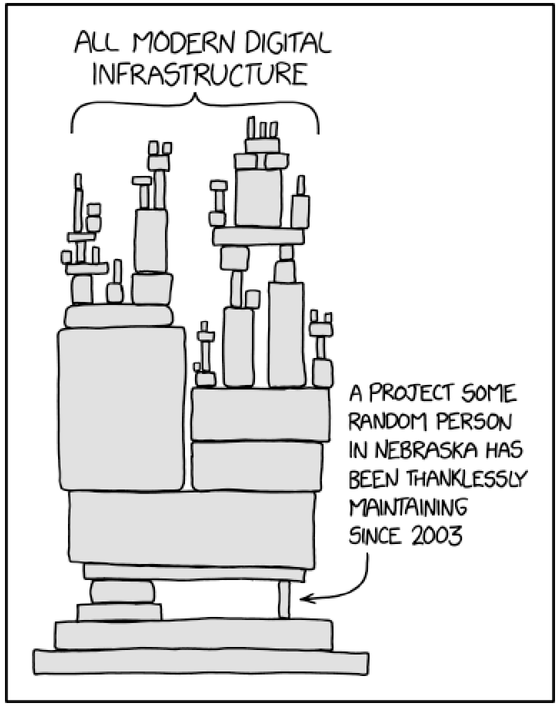
Harvest Now, Decrypt Later Threat

**Today:** Attackers record your encrypted data, waiting for the moment it can be cracked

**Future:** Quantum breakthroughs or stolen keys make that moment possible

# Ecosystem as strength and vulnerability

Increases the number of stakeholders who can fix an issue.


ALL MODERN DIGITAL INFRASTRUCTURE
A PROJECT SOME RANDOM PERSON IN NEBRASKA HAS BEEN THANKLESSLY MAINTAINING SINCE 2003

But can you rely on stakeholders you don't control to fix an issue?