SANS Technology Institute Research Review Journal 2026

SANS TECHNOLOGY INSTITUTE

Ed Skoudis

President, SANS Technology Institute

We live in times of immense technological change, where the pace of innovation, especially with advances in AI and cloud-scale systems, can feel as exhilarating as it is unsettling, introducing new vulnerabilities and attack surfaces as quickly as we learn to defend the last ones. That reality is precisely why I'm so proud to present this year's *SANS Technology Institute Research Review Journal*, which showcases the remarkable work of our student scholars and their research advisors as they push practical, defensible cybersecurity forward. The breadth and depth of this year's papers is genuinely breathtaking: from mapping attack surfaces in an agentic world and extracting offensive value from crash dumps, to improving forensic integrity across filesystems and cloud platforms, and so much more, with a special focus on AI and rigorously testing where AI both helps and harms, including prompt injection risk reduction, security-focused code review, threat modeling, telemetry normalization, and even the security implications of coding assistants. I encourage you to explore these topics and sample at least three to five papers (or more) to get a clear view of what's happening on the cutting edge of cybersecurity research, led by our inspiring SANS.edu students, sharpened by faculty mentorship, and driven by a shared mission: making cyberspace safer, more secure, and more trustworthy for everyone.

Dr. Johannes Ullrich

Dean of Research, SANS Technology Institute

It is no surprise that the artificial intelligence section doubled in size this year. Our students love to live at the cutting edge of information security research. Even papers outside our AI section cover adjacent topics. But I am equally excited by the variety of topics our students tackle. It is a testament to the quality of the classes they are taking. As new concepts are introduced in class, our unique instructors can demonstrate their applications, and our students expand what they learn into new and exciting research. Each paper represents a student who learned a concept, took it "home," applied it, and found a new way to solve a problem they had been faced with. I look forward to hearing from others how they applied the solutions presented in the papers.

Table of Contents

Artificial Intelligence (AI)

Cloud Security

Cyber Defense

Table of Contents
(continued)

Digital Forensics, Incident Response, and Threat Hunting

Industrial Control Systems Security

Leadership and Security Awareness

Offensive Operations, Red Teaming, and Penetration Testing

STUDENT **HIGHLIGHT**

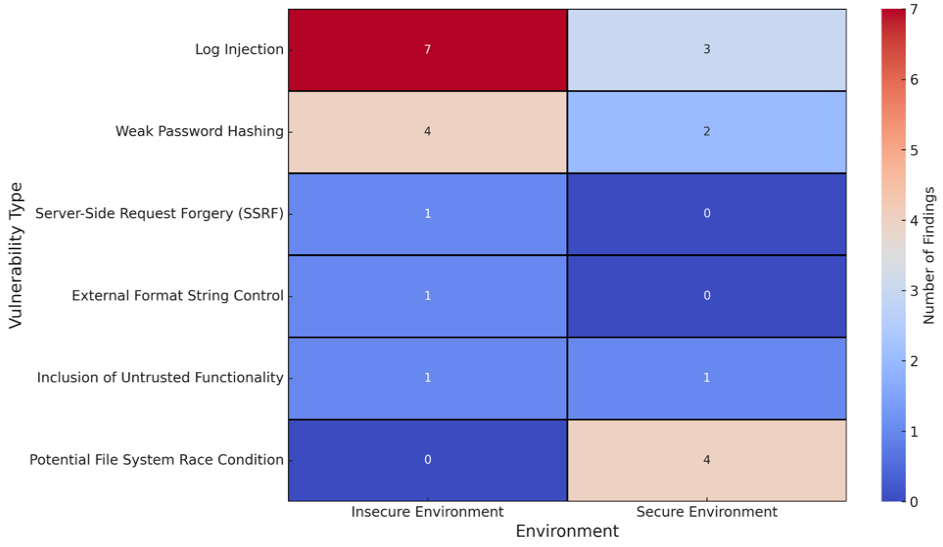
Do AI Coding Assistants Make Bad Coders Worse? A Security Evaluation of GitHub Copilot

by Andrew Hannaford **READ THE RESEARCH**

As AI coding assistants become increasingly integral to software development, the security of their generated outputs is under greater scrutiny. GitHub Copilot has raised concerns due to its potential to replicate or exacerbate existing vulnerabilities inadvertently. Previous research indicates that Copilot often generates insecure code (Pearce et al., 2022; Fu et al., 2024) and that its suggestions are influenced by the surrounding context (Asare, 2023). This study employs static application security testing (SAST) to evaluate the code produced in both environments.

This paper examines whether the overall security posture of a project affects the quality of the code produced by Copilot. It compares Copilot's output in two distinct environments: one that adheres to secure coding practices and another with known vulnerabilities. The objective is to determine whether Copilot perpetuates poor practices or adapts to more secure methodologies. The findings provide practical guidance for developers and emphasize strategies such as careful prompt design and secure project scaffolding to help mitigate the risk of introducing vulnerabilities through AI-assisted coding.

FIGURE 2: VULNERABILITY FINDINGS HEATMAP: SECURE VS INSECURE ENVIRONMENTS



"The use of AI tools in software development is no longer controversial; it is table stakes for anybody developing software. But many developers do not yet understand the full impact of these tools and how to use them securely. Andrew looked at one critical aspect of these tools: how existing code affects them. Most tools will use existing code to learn more about the code to which they contribute. But will they include mistakes and copy them? Andrew systematically investigated this important question. Before using coding tools like Copilot, read the paper. Or, if you already use them, read the paper! This is valuable insight for informing the policy and practice of modern AI-supported software development."

– **DR. JOHANNES ULLRICH, FACULTY RESEARCH ADVISOR**

Measuring Malware Obfuscation: Evaluating CNN-Based Detection for Real-World Resilience

by Michael Reglein READ THE RESEARCH

Static malware detection offers the speed and scalability necessary to process millions of files daily. However, it remains vulnerable when confronted with obfuscation, a deliberate modification of code designed to evade detection and analysis. This study examined how layered obfuscation affects image-based convolutional neural network (CNN) detectors and introduces a novel, reproducible framework for measuring obfuscation itself. Using greyscale byte-to-image representations of Windows portable executables, CNN detectors were tested against progressively more complex obfuscations, ranging from minor structural edits to multilayer overlays. This study highlights that effective detection requires not only capable models but also thoughtful sample exposure, calibration, and validation that reflect the realities of modern obfuscated malware.

Results demonstrated that exposure to obfuscated malware, rather than CNN complexity, governed model robustness. Obfuscation-naïve models, trained only on reference binaries, performed well on reference binaries; however, they degraded sharply when tested against obfuscated samples. Conversely, models trained with limited exposure to obfuscated variants maintained better detection rates and resilience. Preprocessing choices for the byte-to-image representation also proved influential on performance, preserving key obfuscation signals while minimizing computational cost. To verify that the preprocessing maintained meaningful variation between difficulty tiers, a Jensen-Shannon challenge score was introduced to measure divergence between reference and obfuscated binaries. This challenge score confirmed that higher obfuscation levels produced distinct statistical signatures, while explaining why intermediate tiers sometimes converged. Together, these findings offer instructive insights for researchers and practitioners seeking to enhance their understanding of the resilience of static AI-based malware detectors.

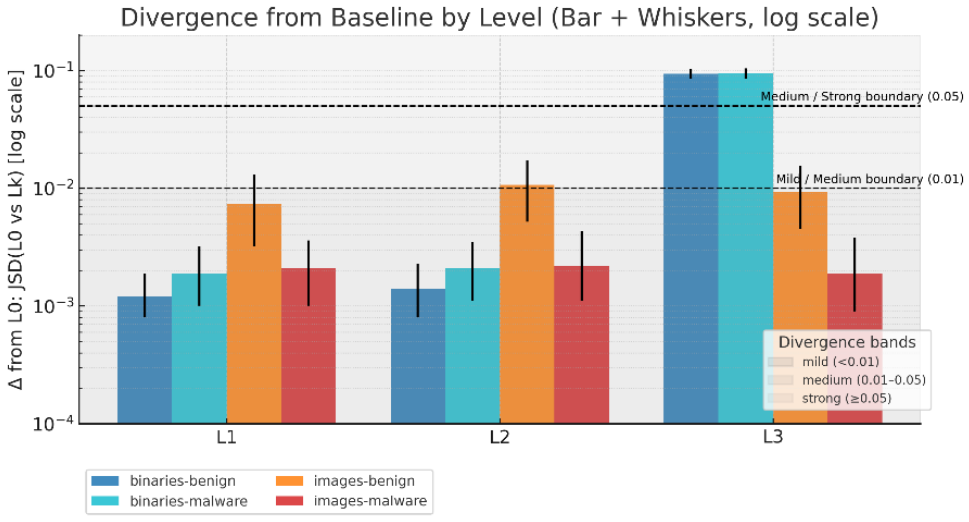


FIGURE 4: DIVERGENCE FROM L0 ACROSS L1–L3 FOR BENIGN AND MALWARE BINARIES/IMAGES. HORIZONTAL DASHED LINES MARK MILD (0.01), MEDIUM (0.05), AND STRONG (≥0.05) DIVERGENCE BANDS ON A LOG SCALE.

Automating Generative AI Guidelines: Reducing Prompt Injection Risk with "Shift-Left" MITRE ATLAS Mitigation Testing

by Adam Wilson READ THE RESEARCH

Automated testing during the build stage of the AI engineering life cycle can evaluate the effectiveness of generative AI guidelines against prompt injection attacks. This technique provides early feedback for developers and defenders when assessing the mitigation performance of an LLM-integrated application. This research combines prompt engineering techniques and automated policy violation checks in the GitHub Actions cloud-native build system to demonstrate a practical "shift-left" approach to securing apps based on foundation models.

Mitigation Technique(s)		Comparison to Control Group (Malicious Prompts + No Guidelines Mitigation Techniques)
CoT	RAG-Sourced Few-Shot	Cohen's *d* / Effect Size
✓		0.09 / N/A (< 0.20)
	✓	0.37 / Small (< 0.50)
✓	✓	0.48 / Small (< 0.50)

FIGURE 17: EFFECT SIZE OF GENERATIVE AI GUIDELINES MITIGATION TECHNIQUES

Leveraging Large Language Models for Security-Focused Code Reviews

by Daniel McQuade READ THE RESEARCH

This study investigates the potential application of Large Language Models (LLMs) in enhancing software security through automated vulnerability detection during the code review process. The research examines the efficacy of LLMs in identifying security vulnerabilities that human reviewers, particularly those without extensive security backgrounds, might overlook. Through analysis of historically significant Common Vulnerabilities and Exposures (CVEs) in popular open-source projects, including frameworks such as Django and Log4j, this research evaluates the capability of LLMs to detect subtle security flaws within complex codebases. The methodology employs a phased approach to LLM prompting, progressing from general code analysis to targeted vulnerability identification while maintaining controlled conditions by isolating vulnerable code segments. By comparing LLM performance against traditional human code reviews and automated security scanning tools, this study provides crucial insights into the potential role of artificial intelligence in augmenting software security practices. The findings suggest implications for the evolution of code review methodologies and the integration of AI-assisted security analysis within software development lifecycles.

Aspect	GitHub Copilot	Google Gemini	Claude
Technical Detail	High	Medium	High
Context Understanding	Medium	High	High
Remediation Guidance	High	Medium	High
False Positive Rate	Low	Low	Low

FIGURE 1: QUALITATIVE ASSESSMENT OF ANALYSIS DEPTH

SIEM Detection Logic Conversion with LLMs

by David Wolverton READ THE RESEARCH

Migrations of mature security information and event management (SIEMs) can be overwhelming due to the sheer volume of detection logic and log sources that must be translated between platforms and query languages. This research explores how Large Language Models (LLMs) and automation scripts can expedite the translation of detection logic between SIEMs, converting detections in minutes instead of hours. Multiple tests can be conducted to optimize translation results, test various LLM parameters, and increase the successful output of the conversion. This translation process can be automated by utilizing scripting and API integrations, significantly reducing the manual effort involved in SIEM migrations.

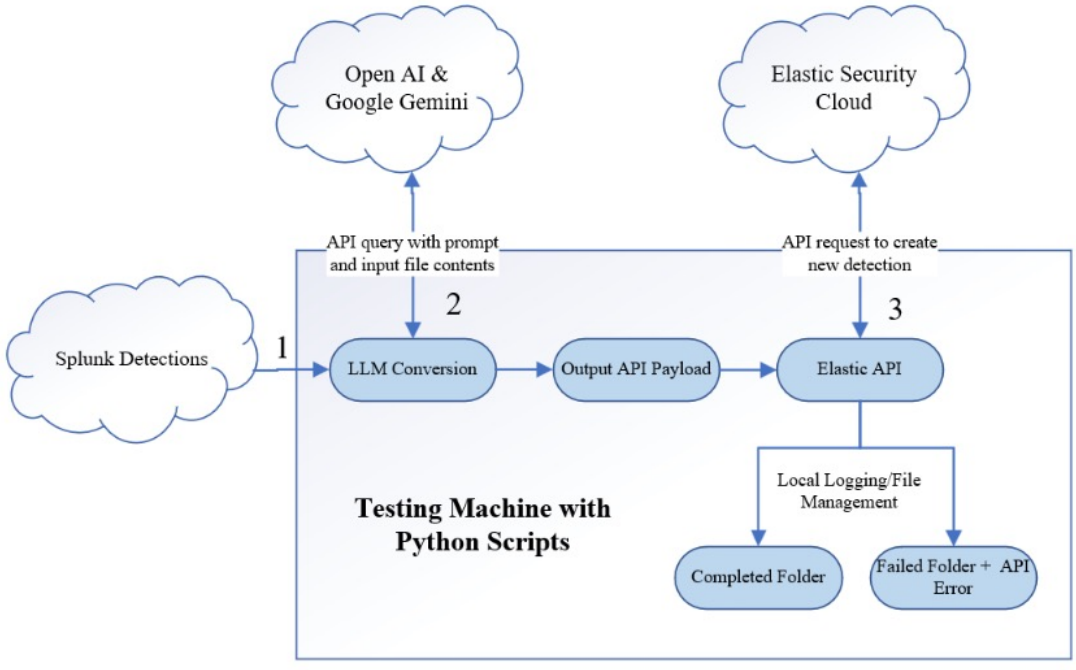


FIGURE 2: TESTING DATA FLOW

MITRE ATT&CK Labeling of Cyber Threat Intelligence via LLM

by Terence O'Brien READ THE RESEARCH

Cyber Threat Intelligence (CTI) is a critical component of a cybersecurity program, and the ability to effectively parse, categorize, and ingest it can significantly enhance an organization's cybersecurity. With the rise of Large Language Models (LLMs), there is an opportunity to automate much of this work and enrich a cybersecurity analyst's productivity. This paper explores the effectiveness of various online and locally hosted LLMs in classifying an arbitrary statement as containing an MITRE ATT&CK Framework (MAF) technique or not and then producing the technique number if it does. LLMs from OpenAI, Anthropic, and Meta, will be the models under test. The statements provided to the LLMs for classification were initially pre-labeled with an MAF technique, with the results being compared to these previously labeled statements. The results of this research will provide insight into the current capabilities of LLMs in classifying complex and nuanced cybersecurity attack techniques to provide a path to automating the labelling of MAF techniques.

FIGURE 6:
RESULTS FOR TECHNIQUE MATCHING WITH PROMPT 3

Model	Accuracy	Precision	Recall	F0.5
o1-preview	65.87%	57.00%	100.00%	49.89%
claude-3-5-sonnet-20241022	60.32%	50.00%	100.00%	44.44%
claude-3-opus-20240229	58.73%	48.00%	100.00%	42.86%
o1-mini	56.35%	45.54%	100.00%	40.89%
claude-3-5-haiku-20241022	54.76%	44.79%	91.49%	39.91%
gpt-4o	55.56%	44.44%	97.78%	39.91%
Llama 3.1 70b	51.59%	39.60%	100.00%	36.04%
gpt-4-turbo	50.00%	38.14%	92.50%	34.58%
gpt-4	47.62%	34.00%	100.00%	31.34%
gpt-4o-mini	38.89%	25.56%	69.70%	23.41%
Llama 3.1 8b	17.46%	0.99%	20.00%	0.98%

AI-Driven Insecurity: Assessing Security Gaps in AI-Generated IT Guidance

by Edward Abbott READ THE RESEARCH

The increasing reliance on AI-generated technical guidance for IT system configuration introduces significant security risks. This study assesses these risks through a case study: setting up an Apache web server on a Rocky Linux system using instructions from seven AI models. This inquiry also addresses the potential for over-reliance on AI and the possible erosion of cybersecurity skills among IT professionals. The research demonstrates the variability and potential security gaps in AI-generated instructions by analyzing responses to two carefully designed prompts. The findings highlight that AI models, in their native state, often do not adequately account for cybersecurity best practices, and that security-focused prompts are essential to elicit more secure configuration guidance. These results emphasize the critical need for human oversight, validation, and security expertise in AI-driven IT operations.

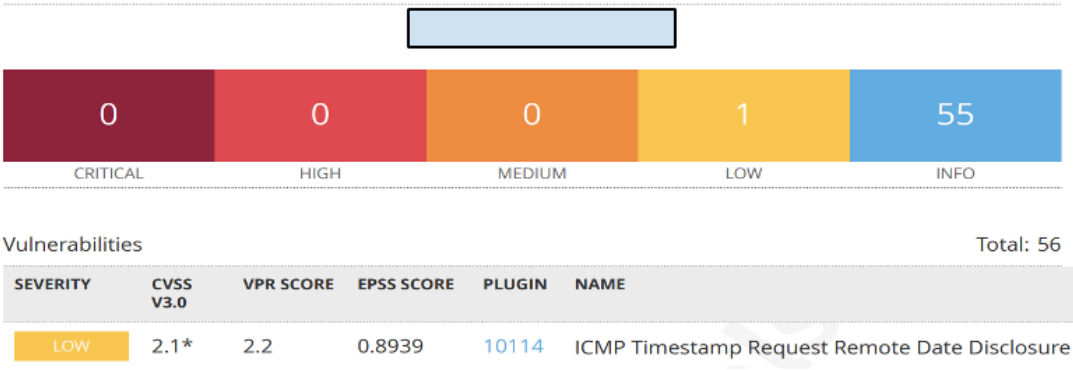


FIGURE 2: ROCKY BASELINE SCAN RESULT

Trust But Verify: Evaluating the Accuracy of LLMs in Normalizing Threat Data Feeds

by Nicholas Peterson READ THE RESEARCH

This paper examines whether Large Language Models (LLMs) can be reliably applied to the normalization of Indicators of Compromise (IOCs) into Structured Threat Information Expression (STIX) format. Using benchmark datasets of 200 IOCs across three types (MD5 hashes, URLs, and IPv4 addresses), the performance of Google's Gemini 2.0 Flash and OpenAI's ChatGPT-4o will be evaluated. While both models achieved 100% validity in generating syntactically correct STIX outputs, their fidelity in accurately preserving IOC values varied significantly. Gemini outperformed ChatGPT overall, though both models struggled with hash values, exhibiting frequent omissions and erroneous pattern translations. The inconsistencies in these errors pose a major obstacle to the reliable use of LLMs in operational security and data engineering pipelines.

Model	IPv4	URL	Hash
Gemini 2.0 Flash	100.0	100.0	99.975
ChatGPT 4o	100.0	98.7	99.075

FIGURE 2: PERCENTAGE OF ACCURATE INDICATOR VALUES BY INDICATOR TYPE

Evaluating Large Language Models for Automated Threat Modeling: A Comparative Analysis

by Eric Sekercan READ THE RESEARCH

This study investigates the use of Large Language Models (LLMs) as an assistant to conduct threat models of systems or applications. It researches the efficacy of a sample of modern LLMs against a constant system, a WordPress application deployed in Kubernetes. It compares the results based on four key metrics: threat coverage, completeness & depth of explanation, consistency, and false positive rate. The methodology involved codifying the inputs and the outputs of the LLMs to provide a consistent base to draw clear comparisons. By analyzing the results, this study yields essential implications for cybersecurity practitioners seeking to scale their threat modeling and security assessment programs. Furthermore, the findings suggest implications for AI-augmented security reviews, which will likely lead to more autonomous agentic workflows.

Model	Consistency	Unique Threats/ Techniques	Depth of Explanation	False Positives
claude-3-5-haiku-latest	88%	18	3/5	1
claude-3-7-sonnet	87.5%	22	4/5	2
claude-opus-4-0	87.5%	17	4.5/5	1
chatgpt-4o-latest	85.7%	18	4/5	1
o4-mini	84%	24	4.5/5	2

TABLE 1: COMPARATIVE RESULTS SUMMARY

Can Your Security Stack Handle AI? An Empirical Assessment of Enterprise Controls Versus Generative AI Risks

by Blake Roth

READ THE RESEARCH

Enterprise security teams face a critical dilemma. Executives want AI productivity gains, but it remains uncertain if existing security controls can handle the risks. Traditional security frameworks were not designed for systems that can generate convincing misinformation, leak sensitive data through creative prompting, or be manipulated into providing attack guidance. Through systematic adversarial testing across multiple AI platforms, this study reveals which enterprise controls work against AI-specific threats and which leave dangerous gaps. Testing data loss prevention, access controls, logging, and monitoring against five threat categories shows strong protection for privacy and data leakage, but alarming vulnerabilities to social engineering attacks that bypass AI safeguards entirely. Organizations can implement AI safely using enhanced versions of existing security frameworks, but they must simultaneously address critical gaps that traditional controls cannot handle.

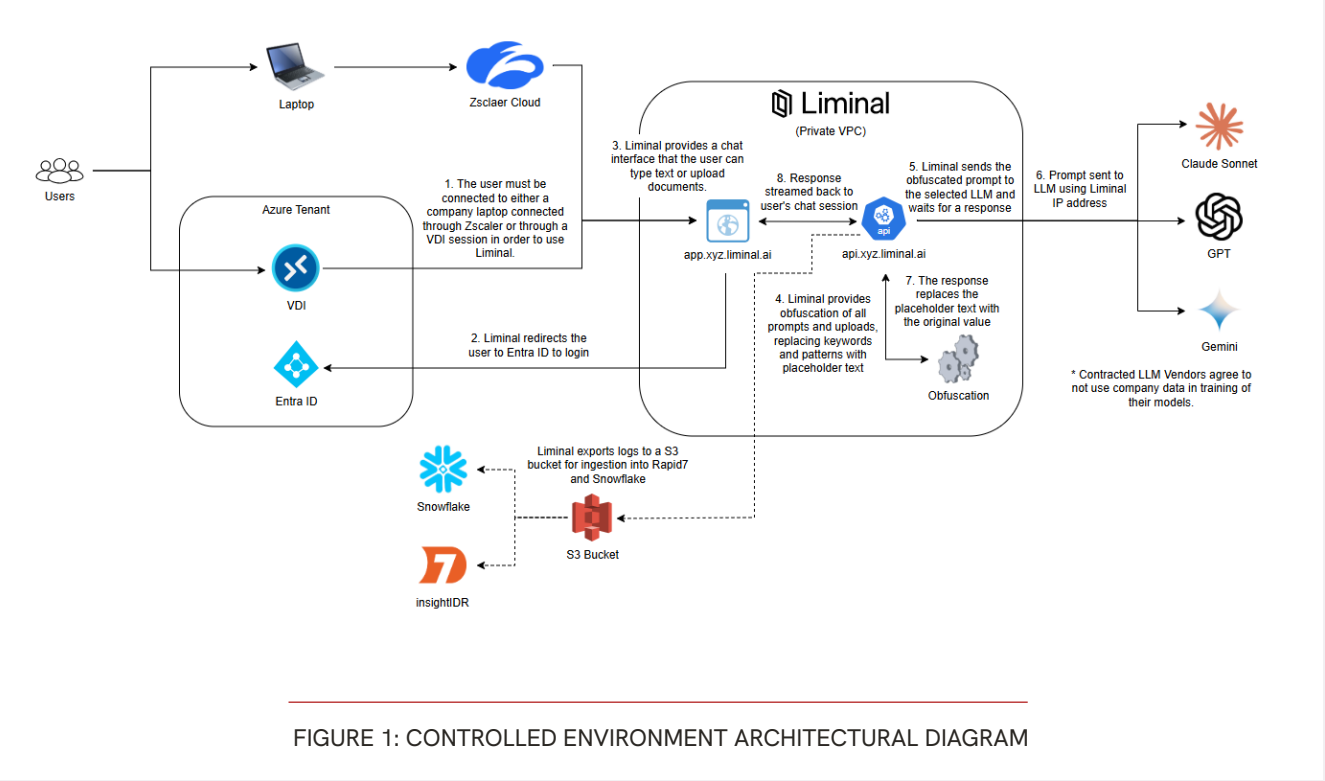


FIGURE 1: CONTROLLED ENVIRONMENT ARCHITECTURAL DIAGRAM

Fixing What You Broke: Can AI Be Used to Thwart AI-Generated Malware?

by Owen Slubowski READ THE RESEARCH

AI has increased accessibility to beneficial technological capabilities for organizations and ushered in a new era of advanced threats. With the help of AI, the generation of new and dangerous malware is faster and more easily developed than ever before. This paper investigates whether legacy tools such as VirusTotal, MetaDefender, and Hybrid Analysis remain effective in detecting modern threats and explores the alternative of using AI as a detection technique. This paper will compare the results of AI-generated malware analysis using legacy tools and various AI models and prompts to develop best practices to protect organizations of all sizes. The results show that AI-assisted malware analysis is significantly more effective at detecting these new threats than legacy approaches and provides vital analysis. Throughout testing, ChatGPT has been proven to be the most effective model for malware analysis. This paper also explores how AI file analysis can be automated using the low-code automation solution N8n to further augment detection. The implications of this research can help organizations defend their interests more cost-effectively amid rapid technological change.

TABLE 1:
SUMMARY OF
EXPERIMENT FINDINGS

Detection Method	Detection Rate
AI total	20/36
Legacy total	2/12
ChatGPT total	8/12
Grok total	6/12
Gemini total	6/12
VT total	1/4
HA total	0/4
MetaDefender total	1/4

STUDENT **HIGHLIGHT**

Securing Azure with PIM: A Just-in-Time Access Study

by Dustin Bourgois READ THE RESEARCH

Misconfigured privileges in Azure environments present serious risks to organizations, including privilege escalation, data breaches, and financial losses, especially as cloud adoption increases. This study assesses Azure Privileged Identity Management (PIM) and its Just-in-Time access model within a controlled Azure environment, simulating enterprise scenarios across Azure Subscription Roles. Findings show that PIM's time-bound privileges significantly reduce the attack surface by limiting unauthorized access outside approved periods. These insights provide organizations with actionable strategies to strengthen cloud security amid evolving cyber threats.

TABLE 4:
TEST RESULTS FOR AZURE PIM ACCESS ENFORCEMENT

Test Scenario	Role	Outcome	User Tested
Activate Owner Role Without MFA	Owner	Blocked	Sys_admin
Add Permissions as Contributor	Contributor	Blocked	Sys_admin
Add Permissions as Owner	Contributor	Allowed	Sys_admin
Add Resources as Owner	Owner	Allowed	Sys_admin
Add Resources as Contributor	Contributor	Allowed	Sys_admin
Add Resources as Reader	Reader	Blocked	Sys_admin
Add Tag to Resource	Contributor	Allowed	Sys_admin
Add Tag to Resource	Reader	Blocked	Sys_admin
Outside Window Access	All Roles Eligible	Blocked	All PIM Roles
Modify Resources as Reader	Reader	Blocked	Sys_admin

"As the faculty research advisor for Dustin Bourgois's study on Azure PIM, I am proud of its approach to the critical vulnerability of persistent admin privileges in cloud environments, a key enabler of the 22% of breaches tied to credential misuse in the 2025 Verizon DBIR. The paper's strength lies in its rigorous, lab-based testing of PIM's Just-in-Time model across Owner, Contributor, and Reader roles, yielding fresh empirical proof that time-bound, MFA-secured elevations shrink the attack surface. This is done while candidly highlighting gaps, such as logging shortfalls and service principal constraints, that theoretical analyses often miss. This hands-on approach makes it a practical guide for real-world deployment."

– **RUSSELL EUBANKS, FACULTY RESEARCH ADVISOR**

The Flavor of Clouds: Are Some Cloud Platforms More Attractive to Attackers?

by James Smith READ THE RESEARCH

Financial loss and sensitive data exposure continue to be a significant risk for entities that host systems in the cloud. Identifying if attackers prefer attacking systems hosted in one cloud provider over another could assist architects and engineers in selecting a provider. Honeypots were deployed to Amazon Web Services (AWS), Azure, and Google Cloud Platform, incorporating a Social Engineering lure to measure human interaction and bot interactions to determine if attackers preferred one cloud provider over another. The data analysis did not identify human interactions, leaving only bot interactions for further examination. Hosting providers that hosted the bots were identified by enriching the data during analysis.

The results showed that the SSH server hosted in AWS experienced significantly fewer attacks, and far fewer attacks originated from AWS. Determining causation from this metric alone was not possible. AWS is likely employing undocumented mitigation strategies, attackers may prefer other clouds over AWS, or resources are allocated based on the number of usernames used in the attacks against SSH. The data also showed that a very low percentage of bots attacked all three cloud providers overlapped with one another, indicating that bot herders are configuring attack infrastructure to focus on particular clouds rather than directing bots to crawl the internet mindlessly. Bots were tailored to the environments they attacked based on analysis of how they interacted with the web servers. Defenders, engineers, and architects should not deviate from required and selected security frameworks regardless of attacker preferences that may be identified.

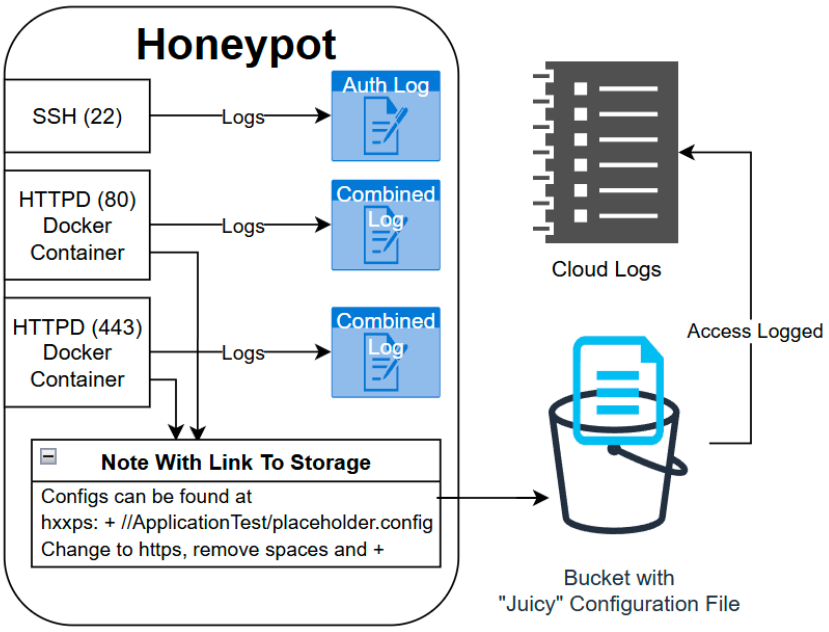


FIGURE 1:
A DIAGRAM OF THE HONEYPOT DESIGNED FOR THIS EXPERIMENT

Detecting Azure Hybrid Machine Attack Paths with Graph Theory

by Shawn Woods READ THE RESEARCH

Today's on-premises and cloud environments are ever-growing and becoming increasingly complex. Attackers know this and can and will exploit this fact, pivoting from network to network. Identity and access management is more critical than ever with hybrid cloud environments. Proper privileges must be assigned according to least privilege principles; if they are not, this is where the problem starts. Attack path mapping and graph databases offer a solution that can highlight potential paths to compromise. Through simple Cypher queries, defenders can observe the potential risks within their environments and mitigate them as needed. This research extends the data collected by the security tool BloodHound to uncover hidden connections between on-premises devices and their cloud identities within an Azure environment. The research offers insights into how organizations can utilize standard tools to add context to their attack maps.

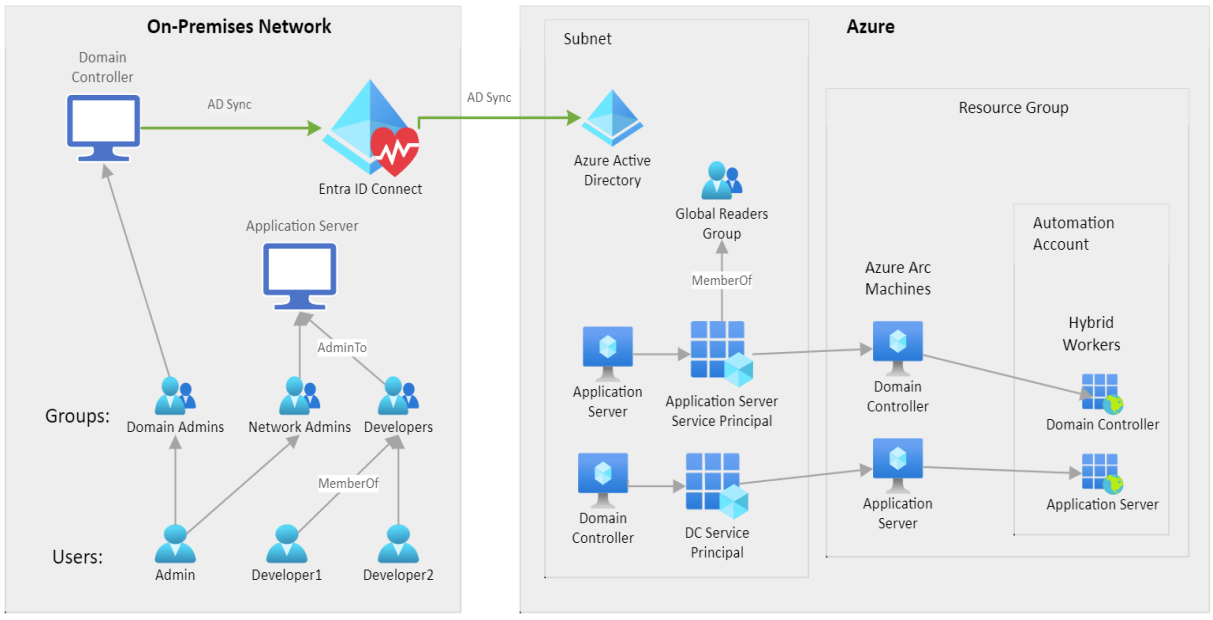


FIGURE 2: ON-PREMISES AND CLOUD ENVIRONMENT

Out-of-Band Defense: Securing VPNs from Password-Spray Attacks with Cloud Automation

by Ryan Wehe READ THE RESEARCH

This research examines an out-of-band solution to detect and block password-spray attacks on Remote Access VPN services, addressing vulnerabilities like Cisco's CVE-2024-20481 amid rising threats post-COVID-19. A virtual Cisco Adaptive Security Appliance (ASAv) was deployed in AWS, using a Python script to monitor logs and block malicious IPs via AWS Network Access Control Lists (NACLs). While effective against controlled brute-force tests, the approach faltered against distributed real-world attacks, prompting a shift to Cisco Security Cloud Control (SCC) and an extended detection window. Limitations persisted with low-volume, multi-source attacks, underscoring the need for scalable cloud defenses, pattern correlation, and threat intelligence integration. Offering an MIT-licensed vendor-agnostic tool, this research enhances VPN security and guides future strategies against sophisticated cyber threats.

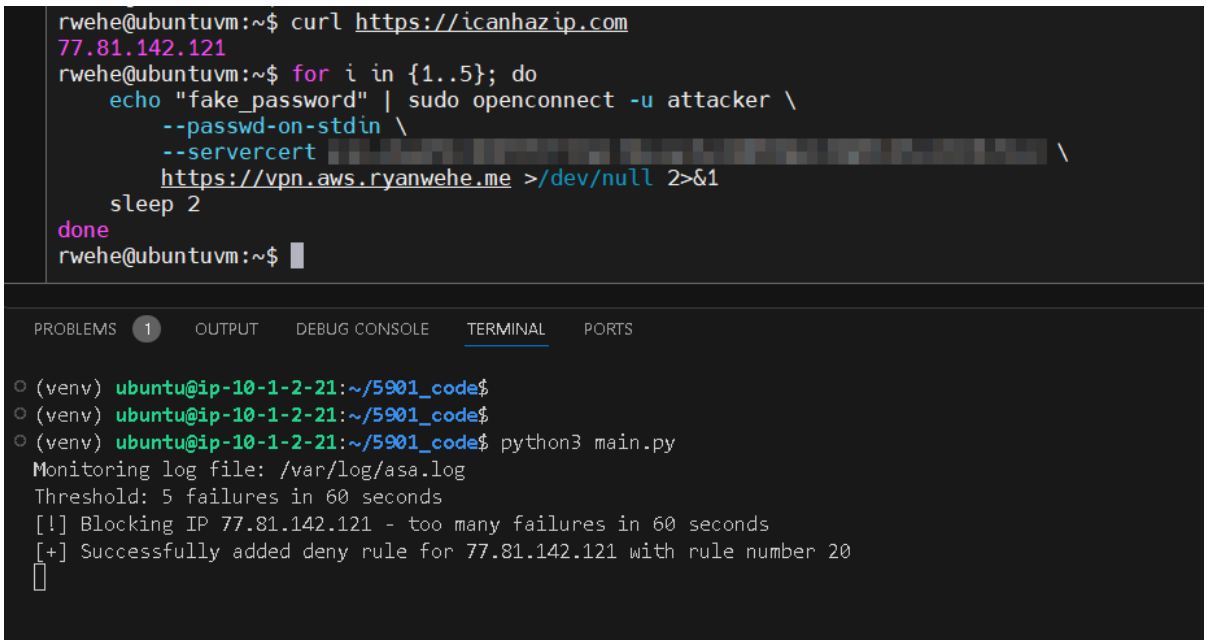


FIGURE 6: ATTACK SUCCESSFULLY BLOCKED

Harnessing Entra ID Snapshots for Effective Post-Security Incident Detection and Containment

by Henry Lopez READ THE RESEARCH

The techniques employed by advanced persistent threats (APTs) necessitate an adaptive incident response strategy. With cyber-attacks increasing in complexity and frequency, security incident responders must detect and contain threats swiftly. This study explores the incorporation of identity snapshots as part of a defense-in-depth strategy and the effectiveness of snapshots in assisting incident responders. Periodic identity snapshots enhanced the accuracy and confidence of incident responders when detecting and containing threats within an organization's identity provider. By maintaining a scheduled log of identities, incident responders can compare the state of identities before and after a compromise, facilitating more targeted investigations. Although identity snapshots are not a comprehensive solution for all threat actors, they serve as a valuable supplementary tool for investigating identity providers such as Entra ID. This research focuses on implementing identity snapshots within Microsoft's Azure Entra ID, demonstrating their potential to significantly enhance the efficiency and effectiveness of post-incident detection and containment.

Detected Threat Accounts					
Group	Sum of Undetected Threats	Sum of Detected Threats	Average of Undetected Threats	Average of Detected Threats	Accuracy of Detected Threats
1	4	5	1.3	1.7	55.6%
2	0	9	0.0	3.0	100.0%

FIGURE 11: DATA BY DETECTED THREATS

Marketing or Added Value? The Truth About Purpose-Built Detection and Response for Containers

by Jeffrey Everling READ THE RESEARCH

Containers are at the frontline of modern organizations. Protecting them is of utmost importance as they support critical business processes. The popular shift-left security approach for containers is adding value for short-lived containers; however, as containers persist over time, runtime security becomes essential to limit the impact of any successful attack. With the rise of Cloud Detection and Response (CDR), this paper dives deeper into the added value and gaps of these solutions compared to the traditional pillar, Endpoint Detection and Response (EDR).

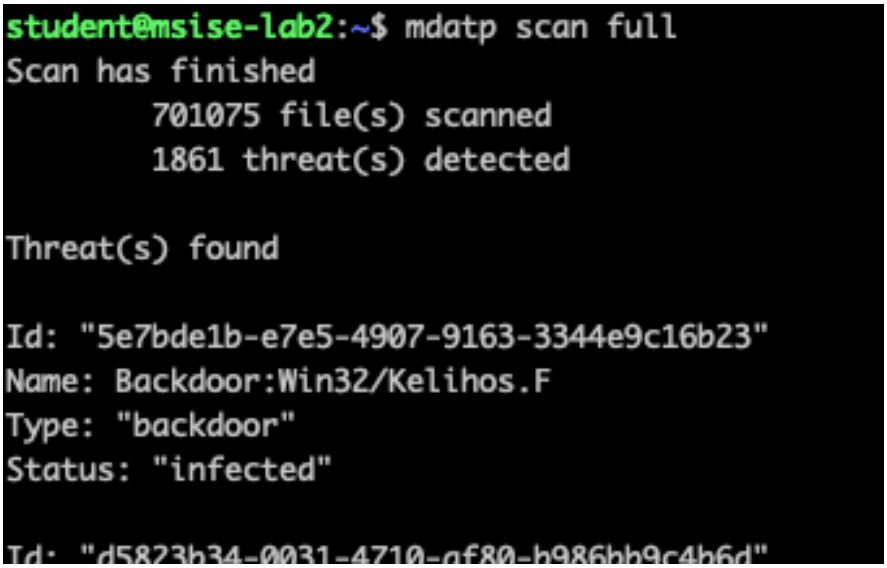


FIGURE 9: RESULTS OF THE FULL SCAN

STUDENT HIGHLIGHT

Unveiling the Dependency on Network Telemetry: Optimizing Lateral Movement Detection

by Kyu Jin Therrien READ THE RESEARCH

Lateral movement is a critical phase of adversarial activity during cyberattacks, enabling attackers to traverse a network, escalate privileges, and exfiltrate sensitive data. Identifying adversaries in complex networks presents significant challenges due to adversaries' use of legitimate tools and processes to evade signature database detection. This study investigates the dependency on network and endpoint telemetry for identifying lateral movement attacks, focusing on the Remote Services technique from MITRE ATT&CK.

The findings emphasize the importance of leveraging anomaly behavior analysis, whether applied to network or endpoint telemetry, to unveil adversarial activities that might otherwise blend with legitimate operations. Lateral movement techniques, identified in MITRE ATT&CK and this study, highlight the need for robust network visibility tools and micro-segmentation strategies to limit adversaries' network propagation while ensuring comprehensive threat visibility and correlation.

TABLE 7: RESEARCH TEST SCENARIO RESULT SUMMARY

	Network Logs	Endpoint Logs
Nmap Scanning	TP: Yes MTTD: 4.5 hour	None
Service Creation Lateral Movement	TP: X MTTD: 4.5 hour	TP: Yes MTTD: 0.01 hour
Remote Services: SSH	TP: Yes MTTD: 0.5 hour	None
Remote Services: WinRM	TP: Yes MTTD: 0.5 hour	None

"As the faculty research advisor for this project, I commend its use of empirical measurements to replace intuition. The study simulated remote services-driven lateral movement in a controlled lab environment using MITRE Caldera, rigorously testing a key defender debate: the efficacy of network versus endpoint telemetry against adversaries leveraging legitimate tools and processes.

The key finding is unequivocal: network telemetry delivered accurate true-positive detections across all techniques, reliably identifying 'legitimate-looking' movement, whereas endpoint telemetry excelled only with known IOCs but missed critical remote-service tactics like SSH and WinRM.

This is timely, as organizations grapple with telemetry overload yet fail to detect internal pivots. The paper equips the industry with a clear directive: prioritize network visibility and anomaly detection, followed by cross-source correlation, to bridge persistent gaps."

– RUSSELL EUBANKS, FACULTY RESEARCH ADVISOR

Beneath the Mask: Can Contribution Data Unveil Malicious Personas in Open-Source Projects?

by Ruby Nealon

READ THE RESEARCH

In February 2024, after building trust over two years with project maintainers by making a significant volume of legitimate contributions, GitHub user "JiaT75" self-merged a version of the XZ Utils project containing a highly sophisticated well-disguised backdoor targeting *sshd* processes running on systems with the backdoored package installed. A month later, this package began to be distributed with popular Linux distributions until a Microsoft employee discovered the backdoor while investigating how a recent system upgrade impacted the performance of SSH authentication. Despite the potential impact this backdoor could have had globally, no tooling has been created for monitoring and identifying anomalous behavior by personas contributing to other open-source projects. This paper demonstrates how using graph databases and theory with OSINT contribution data gathered from GitHub can efficiently identify anomalous behaviors exhibited by the "JiaT75" persona across other open-source projects.

FIGURE 3: Query Demonstrating *IS_GIT_REPOSITORY* AND *IS_GIT_COMMIT*

Beyond Detection: Using Real Phishing Data to Gauge Security Training Program Success

by Cory Keller READ THE RESEARCH

Identification of phishing emails can be cumbersome, accomplished by rule-based filters, machine learning, user-submitted, and other automated analyses. User submission is the cheapest and easiest to implement but a much more time-intensive process, adding overhead to already burdened staff. Analysts digging through these reported emails are likely overwhelmed with ticket work, often leading to missed opportunities to find a malicious email and remediate the email's threat before any user replies, clicks on URLs, and submits credentials. Usually, the only course of action is scheduled or remedial user awareness training. By searching for malicious indicators in phishing emails, these metrics can be built based on phishing tactics seen in organizations. This paper defines one method of network security monitoring in an organization to find these existing indicators. It covers the tools utilized, assuming organizational prerequisites are met to analyze decrypted packet captures with network security monitoring.

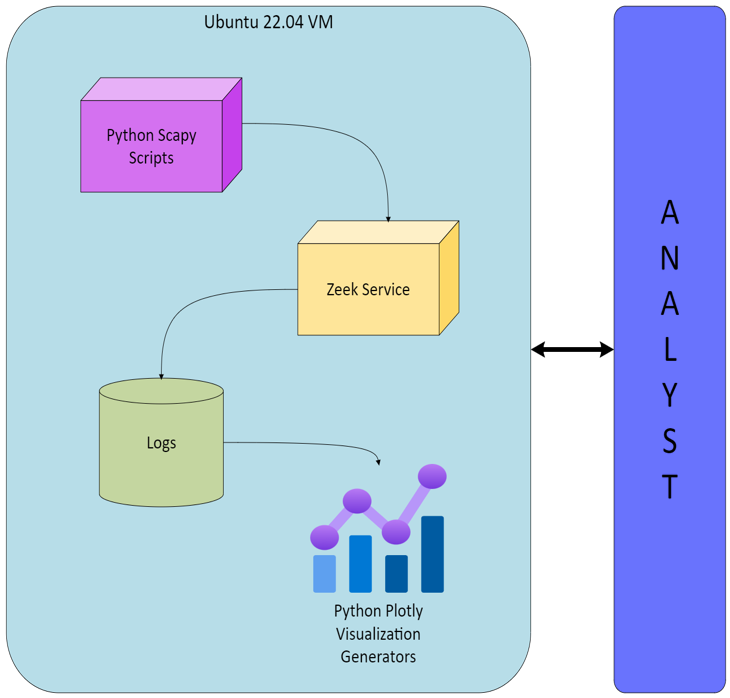


FIGURE 1: VIRTUAL MACHINE SETUP AND WORKFLOW

Shift Left the Awareness and Detection of Developers Using Vulnerable Open-Source Software Components

by Wellington Rampazo

READ THE RESEARCH

The number of open-source software components, as well as the number of existing security vulnerabilities, has increased over the years. Although many vulnerabilities have been published in public data sources like the GitHub Advisories Database, the usage of vulnerable components is substantial, leading to security incidents with catastrophic consequences. Development teams tend to prioritize software releases with new features to achieve business goals over fixing issues or upgrading their software to more secure dependencies, mainly when the software has been released already. The research presented in this paper demonstrates that companies can shift the detection and awareness of developers using vulnerable components left in the early development stages. Implementing network monitoring added to a solution capable of identifying and querying for open-source software components with existing vulnerabilities allows developers to measure the risk and evolve into a secure solution in the earliest stage.

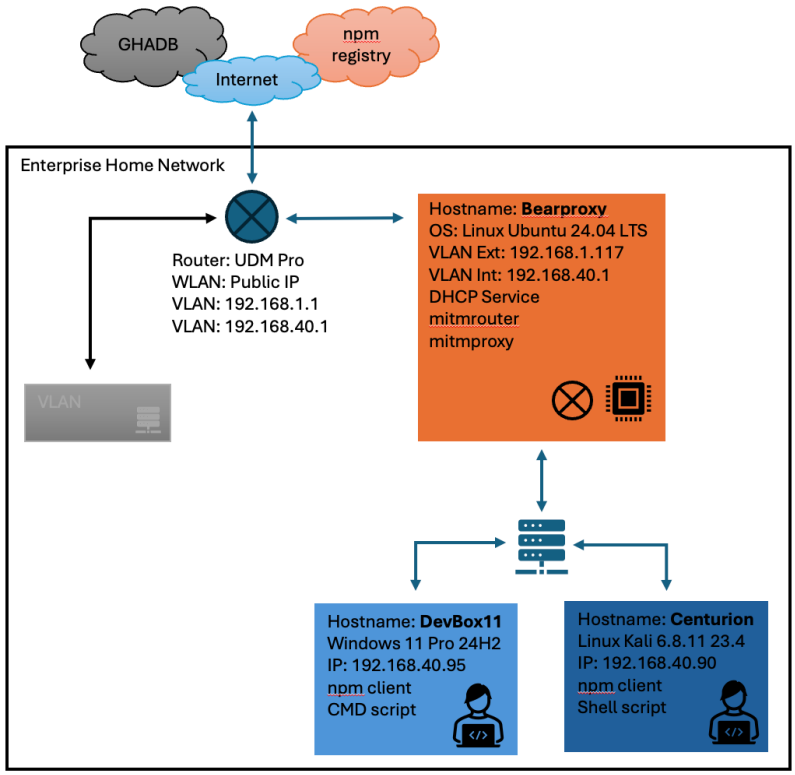


FIGURE 2:
RESEARCH ENVIRONMENT

Persistence Busters: High Impact Methods for Adversary and Threat Detection

by Clark Crisp READ THE RESEARCH

Adversary persistence is a cornerstone of modern cyberattacks, allowing attackers to maintain covert access to systems and evade detection over extended periods. This research investigates the top persistence techniques targeting Windows systems as documented in the MITRE ATT&CK framework and how to detect them. The research uses free and open-source software (FOSS) tools and SIEM detection rules to analyze persistence mechanisms, assess their defensive capabilities, and identify detection opportunities. Key findings reveal gaps in default detection configurations, with advanced evasion tactics bypassing baseline analysis methods and tools. However, combining endpoint detection and response (EDR), ScriptBlock logging, and custom detection rules significantly enhance detection capabilities. This research underscores the importance of a layered defense model and tailored configurations to counter persistence mechanisms effectively, offering actionable insights and recommendations for improving detection strategies.

Summary of Persistence Detection and Visibility				
Technique	**Kibana**	**ScriptBlock Logging**	**EDR Telemetry**	**Baseline Analysis**
T1547 – Boot or Autostart Execution	Two Alerts	Yes	Yes	Detected
T1543 – Create or Modify System Process	Two Alerts	Yes	Yes	Detected
T1053 – Scheduled Task/Job	One Alert	Yes	Yes	Not detected due to T1112

FIGURE 18: DETECTION SYSTEM RESULTS

Evaluating Modern Network Protocol Fingerprinting: Defending Bastion Hosts in Hostile Networks

by Christopher Carroll READ THE RESEARCH

Adversaries continue to attack the network perimeter and trusted user workstations to gain access to sensitive networks. Modern networks are designed and often mandated to use encrypted communication paths everywhere. Once inside the trusted network, credential theft can enable adversaries to penetrate further and gain access to sensitive data. Bastion hosts can be used to strengthen security and prevent unauthorized access. Bastion hosts may be the next target once an adversary gains initial access to a network. Many organizations rely on hardened Bastion host configurations and host-based security solutions to detect, deny, and disrupt adversarial activity. Modern network protocol fingerprinting can provide meaningful out-of-band insight into encrypted connections, shifting the advantage to network defenders.

```
analyst@watchingu:~/pcap$ ja4 scenario05.pcap -J
{
    "stream": 0,
    "src": "192.168.91.132",
    "dst": "192.168.91.129",
    "srcport": "50149",
    "dstport": "22",
    "client_ttl": "128",
    "server_ttl": "64",
    "JA4L-S": "9_64",
    "JA4L-C": "1101_128",
    "ssh_extras": {
        "hassh": "ec7378c1a92f5a8dde7e8b7a1ddf33d1",
        "hassh_server": "a65c3b91f743d3f246e72172e77288f1",
        "ssh_protocol_client": "SSH-2.0-OpenSSH_for_Windows_8.1",
        "ssh_protocol_server": "SSH-2.0-OpenSSH_9.2p1 Debian-2+deb12u3",
        "encryption_algorithm": "chacha20-poly1305@openssh.com"
    },
    "JA4SSH.1": "c1460s36_c187s13_c4s76",
    "JA4SSH.2": "c1460s36_c198s2_c0s100",
    "JA4SSH.3": "c1460s36_c198s2_c0s68"
}
```

FIGURE 15: JA4+SSH Fingerprint of an Unauthorized File Copy to the Bastion

Validating the Effectiveness of MITRE Engage and Active Defense

by Mark Stephens

READ THE RESEARCH

The phrase "Know Thy Enemy," derived from Sun Tzu's *The Art of War*, underscores a fundamental principle in cybersecurity— understanding an adversary's tactics, motivations, and weaknesses is key to staying ahead of their attacks. Traditionally, defenders are forced into a reactive stance, responding to threats only after they emerge. However, by leveraging Active Defense strategies and MITRE's Engage Framework, security teams can flip the script, forcing attackers into unfamiliar territory where they are more likely to make mistakes—mistakes that can be exploited for detection, attribution, and strategic countermeasures. This research examines the impact of Active Defense compared to a traditional security posture when an adversary employs common tactics and techniques to identify high-value targets or exfiltrate sensitive data. By shifting from passive protection to Active Defense, defenders can fundamentally alter cyber conflict dynamics, gaining security from and intelligence about the attacker.

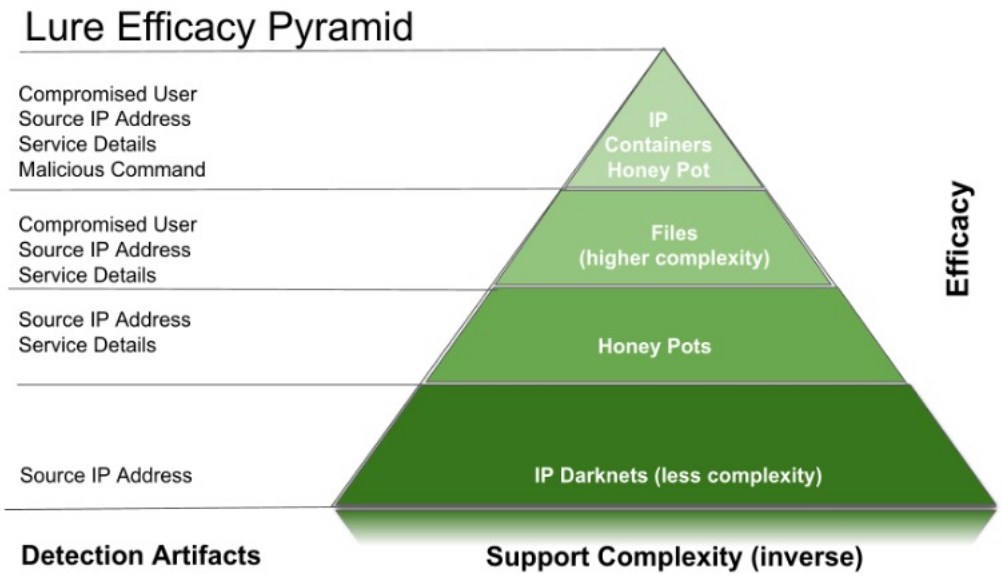


FIGURE 26: LURE SELECTION AND OPERATIONAL OVERHEAD

Building Scalable Detection-as-Code Pipelines with Agentic Validation and Refinement

by Benjamin Opel READ THE RESEARCH

The proposed DaC pipeline uses large language models (LLMs) for logic conversion, variant analysis, and simulation testing via Atomic Red Team, with queries executed against Splunk to measure true positives and false negatives. This paper addresses the pressing need for scalable automation in detection engineering to counter evolving cyber threats. It proposes evaluating an agentic validation and refinement pipeline integrated with Sigma rules and MITRE AT&CK techniques within a Detection-as-Code (DaC) framework. Recognizing that security teams often struggle with manual rule development, which limits throughput and therefore coverage against sophisticated adversaries, a quantitative analysis was conducted to assess the impact of agentic processes on key metrics such as analyst productivity, LLM rule creation efficacy, and MITRE technique coverage. Findings indicate that agentic systems produce dependably valid detection rules in coordination with sound threat emulation design, significantly enhancing throughput and coverage with minimal overhead.

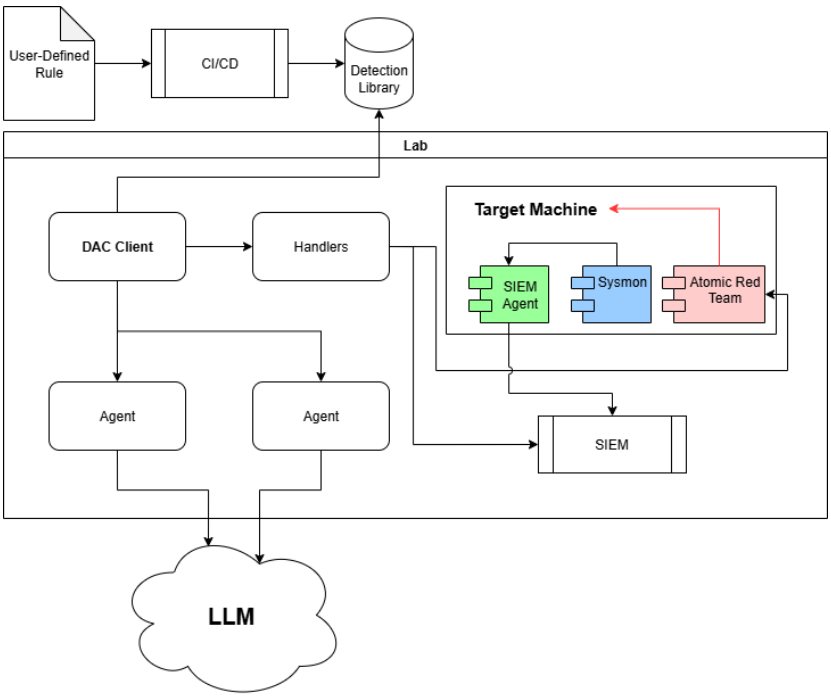


FIGURE 1:
DAC-VAN Concept Illustration

"You Again": Fingerprinting and Tracking Mechanisms of Malicious Sites

by Erin Kuffel READ THE RESEARCH

Browsers provide many APIs for any visited site to perform stateful and stateless tracking, and legitimate websites utilize these capabilities. Yet little is widely known about what tracking, if any, malicious sites perform. Without this insight, malicious websites may be tracking users via their browsers, while security solutions remain blind and ineffective to such activity. Unlike prior research that has focused solely on tracking mechanisms on benign sites, this study examines both malicious and legitimate web pages and compares their use of a combined six stateful tracking mechanisms and stateless fingerprinting techniques. Findings reveal a common dependence on stateful tracking and navigator-based queries, with over 90% originating from third-party sources. These insights underscore the need for browser security solutions that can distinguish and mitigate malicious tracking activity without disabling legitimate website functionality.

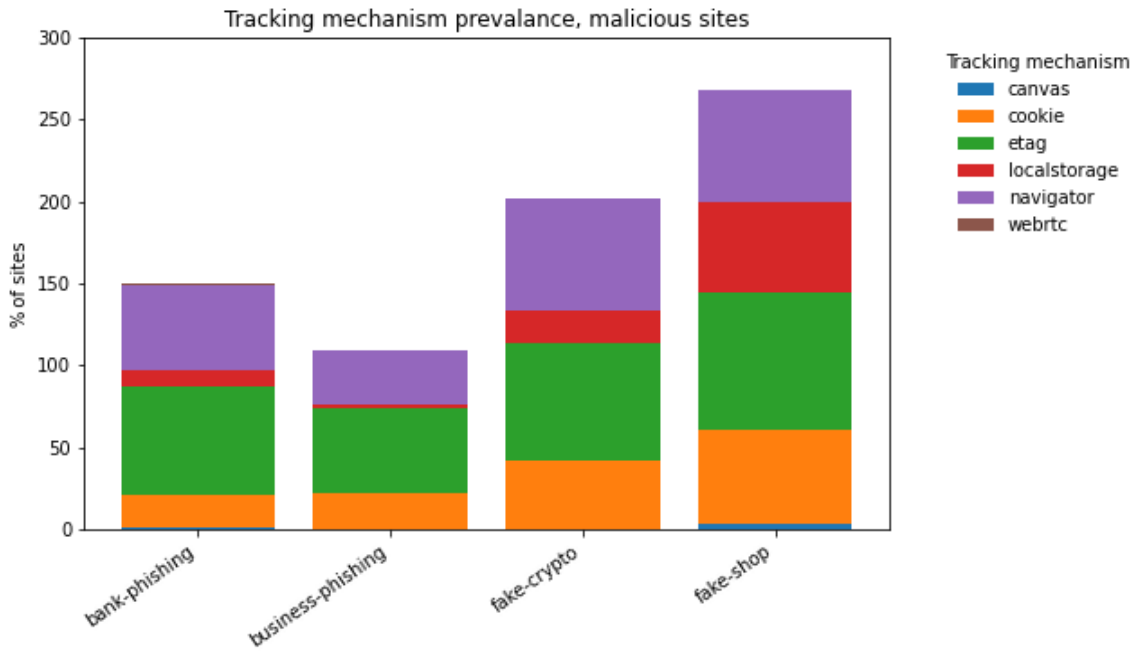


FIGURE 3:
A STACKED BAR CHART BREAKING DOWN THE TRACKING MECHANISM COMPOSITION FOR EACH MALICIOUS SITE CATEGORY

Identifying Advanced Persistent Threat Activity Through Threat-Informed Detection Engineering: Enhancing Alert Visibility in Enterprises

by Eric LeBlanc

READ THE RESEARCH

Advanced Persistent Threats (APTs) are among the most challenging to detect in enterprise environments, often mimicking authorized privileged access prior to their actions on objectives. Moving within the environment slowly and quietly, APTs can often persist within the environment for months before detection. There are several approaches to detecting these adversaries, with many mature enterprises utilizing some combination of User-Entity Behavior Analytics (UEBA), Risk-Based Alerting (RBA), and traditional detection engineering practices. However, even these advanced approaches can have gaps. While they may show anomalous behavior, they can result in false positives, leading to wasted analyst cycles and potential alert fatigue. To combat this, the question is asked: does threat modeling prior to detection engineering generate more robust detections than traditional detection engineering alone? By leveraging the threat modeling process, enterprises can leverage their existing detection strategies differently, using information gained from the threat modeling process to alert them with detections aligning to Tactics, Techniques, and Procedures (TTPs) commonly used together as part of an intrusion.

FIGURE 3: Baseline Detections by MITRE ATT&CK Tactic

Isolated Trust: Zero Trust in Standalone Systems

by Brian Crowley

READ THE RESEARCH

The use of air-gapped, isolated systems remains an essential tool for organizations that require high confidentiality or integrity, including those in the government, industrial control systems, and the banking industry. When these systems are compromised, insider threats are often to blame. This paper evaluates the use of zero-trust principles, initially designed for large enterprise networks, to strengthen systems secured under the Joint Special Access Program Implementation Guide (JSIG). Many zero-trust concepts cannot be implemented in an isolated environment, due to the requirement for enterprise-level connectivity, identity and access management (IdAM), system monitoring, and other requirements that are too costly to implement in very small environments. Those controls that do apply provide incremental improvements to the JSIG baseline. System owners must resolve the tension between additional security tools and maintaining a minimal footprint and attack surface.

Test	Result
USB Mass Storage	Blocked, no alert
HID input, writing EICAR file from Notepad	Succeeded, insofar as the file was written. EICAR was removed by Antivirus, with a visible alert only when attempting to run it. AV removed the test file before AppLocker could assess if it should be allowed to execute. Executable was blocked by default AppLocker rules, except in the Administrator profile
Exfiltrate to Network	Succeeded. No alert. PowerShell logging makes forensic review of this far easier.

TABLE 1: THE BASH BUNNY WAS PARTIALLY SUCCESSFUL

Breaking Through Deception: Addressing Barriers in the Adoption of Cyber Deception Technologies

by Dakota Campbell

READ THE RESEARCH

Despite the increasing sophistication of cyber threats and the need for organizations to employ innovative defense strategies, cyber deception technologies, tools designed to mislead attackers and gain a defensive advantage, remain significantly underutilized across organizational cybersecurity programs. This underuse presents a compelling question: why do tools with such potential struggle to gain widespread adoption? Although much of the prior research on cyber deception technologies has highlighted the theoretical value of the tools, little has addressed real-world deployment concerns, such as lack of awareness, perceived complexity, unclear return on investment, integration challenges, and legal uncertainties. This white paper uses a mixed-method approach to explore the underlying barriers limiting implementation and identifies targeted mitigation strategies. The findings aim to provide organizations and security professionals with evidence-based, actionable insights that reduce uncertainty, support strategic decision-making, and foster greater adoption of cyber deception as a core component of modern defense-in-depth strategies.

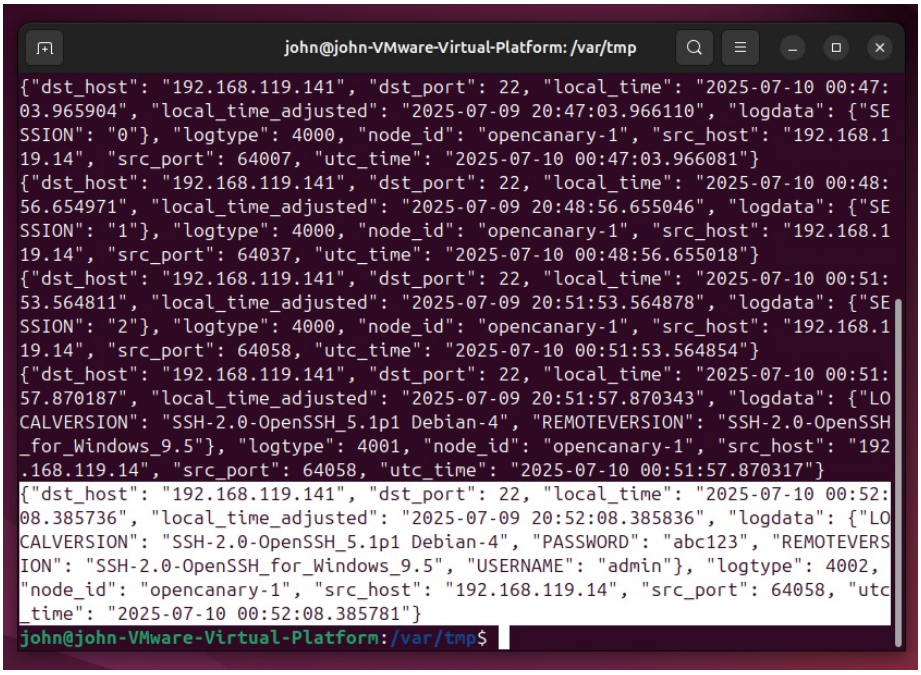


FIGURE 6:
OpenCanary Log Entry Showing That the SSH Connection Attempt Was Successfully Captured

Evaluating Zero Trust Network Access: A Framework for Comparative Security Testing

by Derron Carstensen READ THE RESEARCH

Not all Zero Trust Network Access (ZTNA) solutions are created equal, and despite bold marketing claims, many fall short of delivering proper Zero Trust security. In a crowded market with vague standards and feature parity buzzwords, organizations are left guessing which solutions enforce the principles of Zero Trust. While most evaluations rely on vendor checklists and surface-level comparisons, this white paper takes a different approach: building and applying a hands-on testing framework grounded in NIST SP 800-207 and the CISA Zero Trust Maturity Model. This research uncovers gaps, clear differentiators, and actionable insights by testing five leading ZTNA products across the core pillars of Identity, Devices, Networks, Applications, and Data. The result is a methodology for objectively comparing ZTNA solutions based on their alignment with core Zero Trust principles and organizational security requirements.

		Market Leaders			Niche Vendor	SMB Vendor
	Tests	A	B	C	D	E
Identity	User-Based Access Differentiation	Success	Success	Success	Success	Success
	Step-Up Multi-Factor Authentication	Success	Success	Fail	Success	Success
Device	Disk Encryption	Success	Success	Success	Fail	Success
	Endpoint Protection is Installed	Success	Success	Success	Fail	Success
	Endpoint Protection is Running	Success	Success	Success	Fail	Success
	Endpoint Protection Definitions are Up-to-Date	Success	Success	Success	Fail	Fail
Network	Network Segmentation (Untrusted Locations)	Success	Success	Success	Fail	Fail
	Network Segmentation (Trusted Locations)	Success	Success	Success	Fail	Fail
	Service Cloaking	Fail	Partial	Success	Partial	Fail
Application	Application Visibility	Success	Fail	Fail	Fail	Fail
	Application-Based Policies	Success	Fail	Fail	Fail	Fail
	Protection Against Vulnerable Application Exploitation	Success	Fail	Fail	Fail	Fail
	Detection and Prevention of Local File Inclusion Attacks	Success	Success	Fail	Fail	Fail
	Detection and Prevention of Command Injection Attacks	Success	Partial	Fail	Fail	Fail
	Detection and Prevention of SQL Injection Attacks	Partial	Success	Fail	Fail	Fail
Data	Personally Identifiable Information (PII) Data Loss Prevention	Fail	Success	Fail	Fail	Fail
	Payment Card Industry (PCI) Data Loss Prevention	Fail	Success	Fail	Fail	Fail
	Detection and Prevention of Malicious Data Transfers	Partial	Fail	Fail	Fail	Fail

TABLE 6: TEST RESULTS

Defending Vulnerable Populations Against Scams: Effectiveness of Browser Extensions in Mitigating Scammer Attack Chains

by Thomas Gorman

READ THE RESEARCH

Online scams and phishing campaigns increasingly target vulnerable individuals at home, exploiting trust and limited technical awareness to execute multi-stage attack chains and steal thousands of dollars from victims. These users generally have few defenses—such as free email spam filters, admin credentials, and limited user education—offering up opportunities scammers will take advantage of. This research evaluates the effectiveness of a browser extension as a security control—*Grandma's Guardian*—designed for simplicity and accessibility so that even non-technical home users can benefit from enterprise-grade protection. The extension integrates domain allowlisting, MIME type allowlisting for web downloads, and semantic content filtering to proactively block malicious interactions from view of the user. In controlled simulations of realistic scam scenarios, *Grandma's Guardian* consistently prevented all tested attacks when occurring within a Chromium-based browser. These findings demonstrate that targeted, multi-layered browser-level controls can meaningfully reduce scam exposure for at-risk populations at home, with strong potential for broader application within any organization.

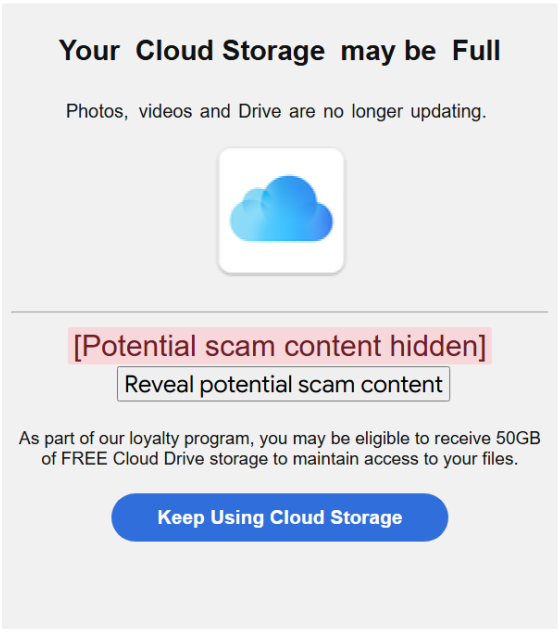


FIGURE 16:
REAL PHISHING ATTEMPT

STUDENT **HIGHLIGHT**

Scrutinizing A Web-Based LLM in Private Browsing Mode: An Analysis of Memory Artifacts and Privacy Implications

by Chris Kosmas

READ THE RESEARCH

Using web-based LLMs such as ChatGPT has changed the web browsing landscape to become part of the typical everyday experience. Web browser memory forensics has been a staple technique used by investigators to gather information about a user's device and behavior, and to predict their preferences. The advent of private browsing furthered security and privacy while attempting to limit the collection of personally identifiable information. Private browsing has impacted the ability to gather complete data using typical browsing methods, such as surfing social media, shopping, visiting academic websites, downloading files, and logging into email accounts. Numerous studies have been conducted on memory-based forensics. These studies uncovered artifacts of interest to forensic analysts and provided privacy insights to the everyday user interested in safeguarding their data. This project aims to apply the same memory forensics techniques in previously explored browsing scenarios to a popular web-based LLM technology, ChatGPT. It attempts to determine if there are any artifacts worth exploring in more detail. It highlights the strength of privacy-focused browsers, such as Brave, compared to mass-market browsers, like Edge.

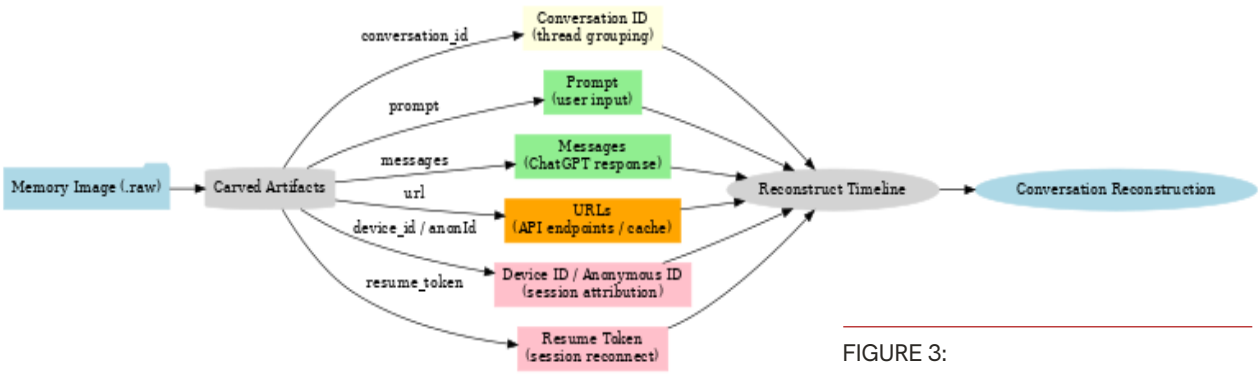


FIGURE 3:
LIST OF CARVABLE STRING ARTIFACTS

"As adoption rates of generative AI tools and AI chatbots become mainstream, it is imperative that we investigate methods to ensure sensitive data is not leaked and confirm techniques to collect and analyze the forensic artifacts. Many users understand the risks of exposing sensitive data to AI assistants, which in turn leads them to access these tools using modern private browsing capabilities, which are designed specifically not to retain evidence of use. This research explores the ability to recover private browsing artifacts attributed to AI usage when the normal private browsing safeguards are followed by the user. It highlights the effectiveness of memory analysis for uncovering artifacts that otherwise would not be saved in typical private browsing sessions, while also introducing a simple but effective script for searching through acquired memory images, highlighting this crucial piece of evidence that can be used to supplement digital forensic analysis."

– **DOMENICA (LEE) CROGNALE, FACULTY RESEARCH ADVISOR**

No-Cost Detection of Endpoint Hard Drive Removal

by Ryan Graham

READ THE RESEARCH

Most organizations today cannot detect if an end user removes their laptop's hard drive, connects it externally, steals data, then reinstalls the drive. While solutions exist today, such as tamper-evident tape, tamper protection BIOS settings, or forensic tools to identify drive activity artifacts, none provide real-time alerts when a drive has been removed and externally connected. This paper analyzes low-cost detection methods, using existing hard drive counters from Self-Monitoring, Analysis, and Reporting Technology (S.M.A.R.T.) and the Windows Registry, for their fidelity in detecting hard drive removal. Since these counters capture the number of times a drive is powered on, they could be monitored for anomalies upon system boot, providing a cost-effective detection mechanism against insider threats.

TABLE 27:
SUMMARY OF FINDINGS FOR COMMON USER SHUTDOWN SCENARIOS WITH FAST STARTUP DISABLED

Fast Startup Disabled				
User Shutdown Scenario	S.M.A.R.T. power count	BootID	Change Value (power count – BootID)	Detection Result
Reboot	0	+1	-1	True Negative
Windows Stop Error (BSOD)	0	+1	-1	True Negative
Reboot into BIOS	0	+1	-1	True Negative
Normal Cold Boot	+1	+1	0	True Negative
Forced Shutdown from Windows	+1	+1	0	True Negative
Power Loss	+1	+1	0	True Negative
Cold Boot into BIOS	+1	+1	0	True Negative
Reboot into Flash Drive	+1	+1	0	True Negative
Cold Boot After External Power-on	+2	+1	+1	True Positive
Forced Shutdown at Boot	+2	+1	+1	False Positive
Cold Boot into Flash Drive	+2	+1	+1	False Positive

Breaking Time: Methods, Artifacts, and Forensic Detection of Timestomping on FAT32, Ext3, and Ext4 File Systems

by Allan Korol

READ THE RESEARCH

Hiding malicious files is imperative to breach a computer system successfully. To conceal malicious files among legitimate ones and complicate forensic investigations, adversaries often employ timestomping, which is the manipulation of file timestamps, as a defense evasion technique. This paper explores the diverse methods used to timestomp files on FAT, Ext3, and Ext4 file systems, focusing on how adversaries adapt their approaches based on available system access and permissions. The direct and indirect forensic artifacts left behind by these methods are analyzed, providing a framework to help investigators identify likely timestomping techniques. By correlating artifacts with specific adversary capabilities, this research supports faster and more accurate detection, enhancing incident response efforts on Linux systems.

```
struct directory {
      // Short 8.3 names
      unsigned char name[8]; // file name
      unsigned char ext[3]; // file extension
      unsigned char attr; // attribute byte
      unsigned char lcase; // Case for base and extension
      unsigned char ctime_ms; // Creation time, milliseconds
      unsigned char ctime[2]; // Creation time
      unsigned char cdate[2]; // Creation date
      unsigned char adate[2]; // Last access date
      unsigned char reserved[2]; // reserved values (ignored)
      unsigned char time[2]; // time stamp
      unsigned char date[2]; // date stamp
      unsigned char start[2]; // starting cluster number
      unsigned char size[4]; // size of the file
};
```

LISTING 1: FAT32 File System Record Structure

Digital Forensics and Incident Response in the Cloud: Addressing GCP Challenges

by Mark Nakamura

READ THE RESEARCH

Many digital forensics and incident response (DFIR) practitioners, as well as aspiring cybersecurity analysts, often gravitate towards AWS and Azure as their first forays into cloud security. This is mainly due to the availability of learning resources as well as existing methodologies, tools, and guides for performing forensics within those cloud service providers (CSPs). Google Cloud Platform (GCP), while not incredibly new, can be overshadowed by the other CSPs and lacks as much research in the field (as opposed to the others), and many vendors tend to add features and capabilities designed for AWS and Azure based on market sentiment. The goal of this paper is to research cloud forensic capabilities, identifying challenges and potential solutions unique to GCP.

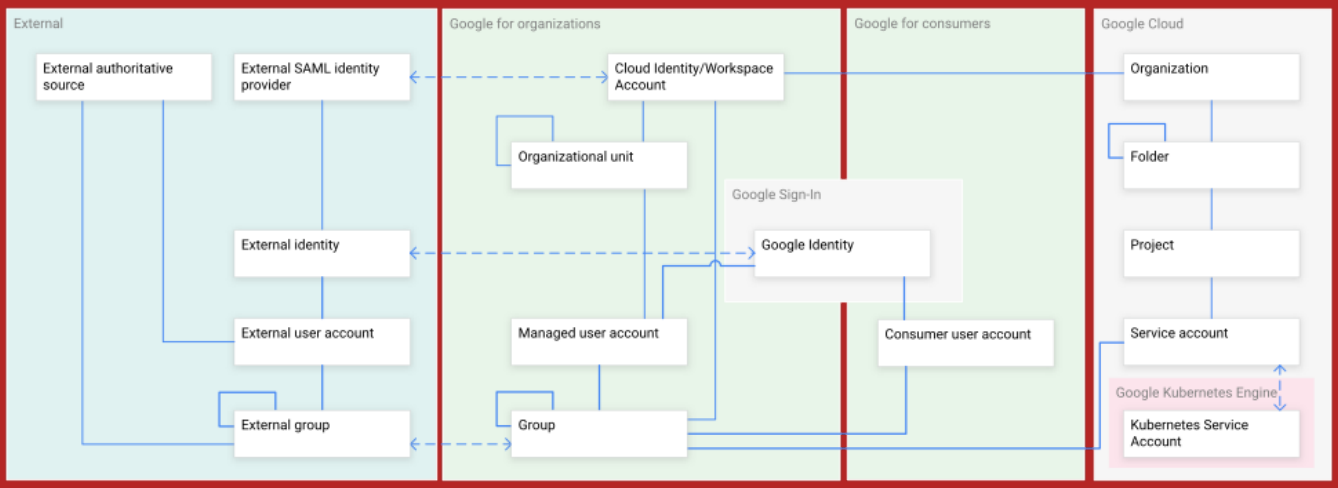


FIGURE 1: IDENTITY ARCHITECTURE

Adversary-Aware IOC Retention: Analyzing Time-to-Live Patterns by Threat Actor Attribution

by Nathaniel Jakusz

READ THE RESEARCH

It is well established that not all threat actors operate similarly. Still, security teams continue to waste storage, processing, and opportunity costs on bloated threat intelligence feeds containing stale IOCs. Early research into this topic compared the price of retaining IOCs over a set time against the price of responding to an incident, while later research evolved to create decay models that reduce the relevance of an IOC over time. Unfortunately, current decay models apply a uniform approach and do not account for individual threat actor Tactics, Techniques, and Procedures (TTPs). After analyzing hundreds of IOCs across three unique Advanced Persistent Threats (APTs) from disparate regions, it can be confirmed that not only do threat actors cycle their IOCs at different rates, but those rates can be tracked. This paper introduces an enhanced decay model incorporating a threat actor variable that accounts for these differences in sophistication and hygiene. This optimized approach to IOC retention will lead to more accurate IOC prioritization, reducing processing and storage costs and time spent responding to false positives.

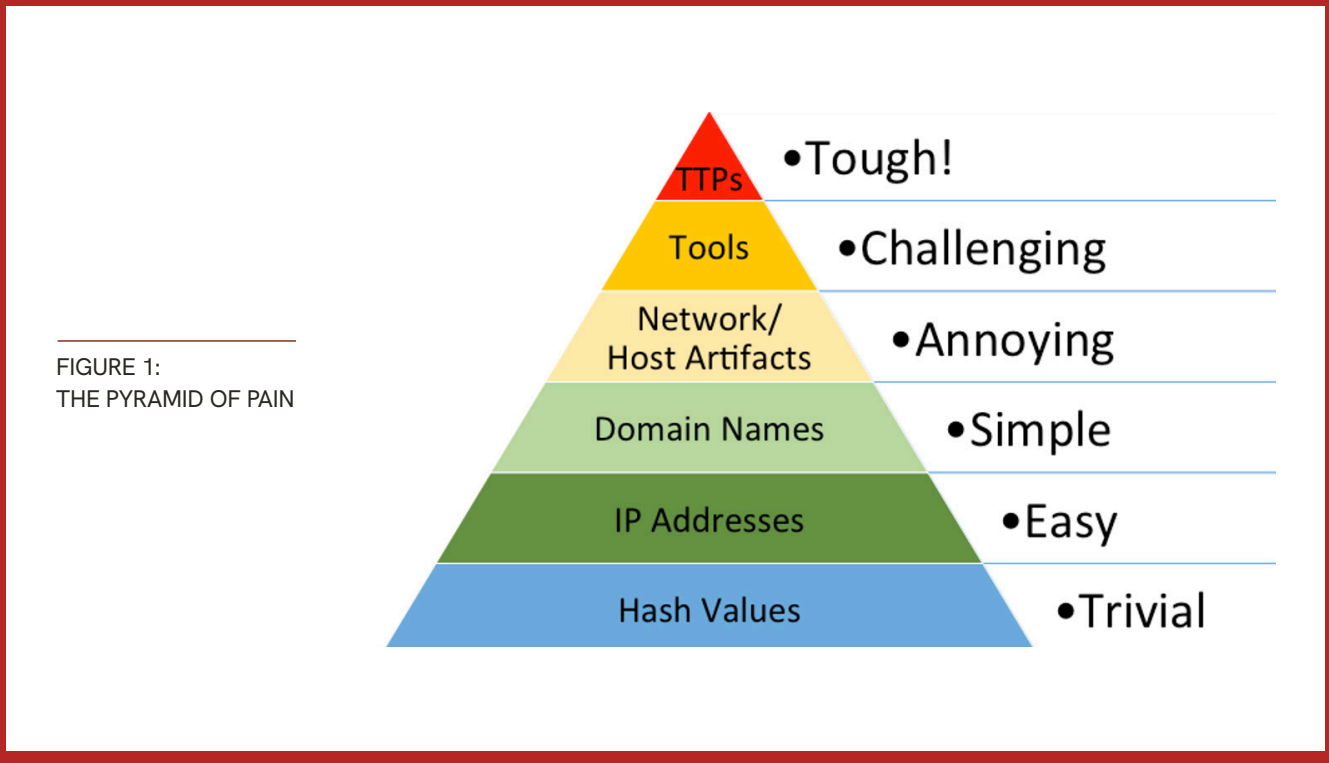


FIGURE 1:
THE PYRAMID OF PAIN

A Pebble in the Ocean: Maximizing Log Fidelity in Container Environments

by Zachary Salva

READ THE RESEARCH

Log fidelity is crucial for Incident Response Teams to investigate and contain cyber incidents but can be difficult to optimize in containerized environments. To improve log fidelity, organizations must invest in understanding their operating environment, and then select and implement a security control framework like NIST 800-53 that helps meet business objectives. A holistic approach to security, including proper logging configurations and proactive monitoring supported by security control frameworks, is necessary to maximize log fidelity. By applying logging-related controls, simulating an attack on a Kubernetes environment, and performing an incident response log collection, we demonstrate the challenges related to logging and security controls in a containerized environment.

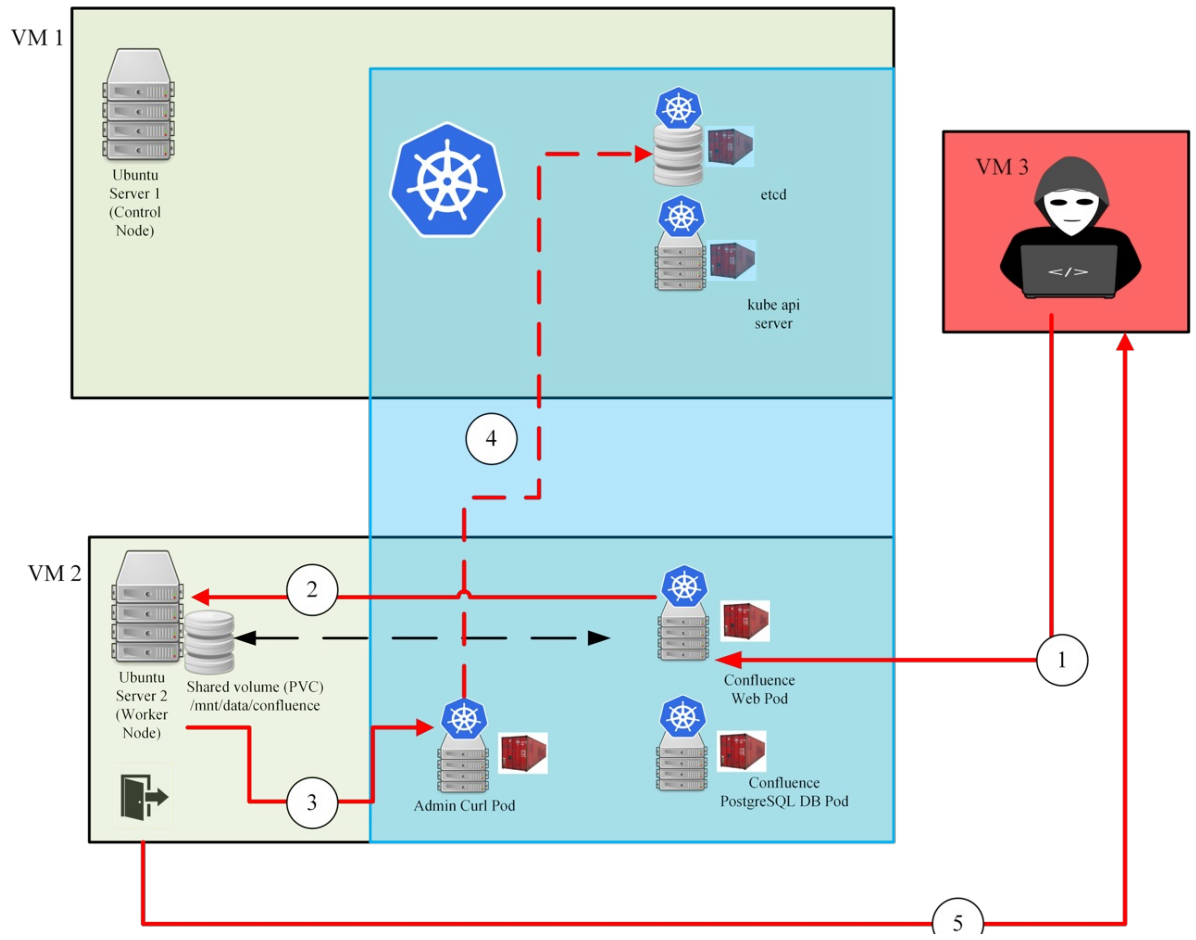


FIGURE 1: VIRTUAL MACHINE SET UP

Catching the Hand in the Cookie Jar: Canary Session Cookies

by Caleb Patten

READ THE RESEARCH

Multifactor Authentication (MFA) has advanced information security beyond the dark days of a simple username and password. While this additional layer of protection is essential, the foundation of internet authentication still largely rests on an antiquated (and inherently insecure) technology: the browser cookie. No matter how many authentication factors used, many web applications still ultimately grant or deny access based on the contents of cookies. The cookie is a bearer token, meaning anyone with possession of the authentication cookie is granted access to the resource – no questions (or passwords) required. While MFA added a mechanism to *authenticate* users, there have been few advancements in securing *the actual token* derived from that authentication process. The value in these cookies and other browser-stored information (autofill data like passwords and credit card numbers) is well known to hackers, as the information stealer business has grown over the last five years. According to one estimate, in 2024 alone, over 450 million people had their cookies and other sensitive data pilfered by just one infostealer (Flashpoint, 2024). This project demonstrates how even applications secured with MFA are still vulnerable to hijacked session cookies. Given the persistent threats posed to organizations by stolen authentication cookies, this research proposes implementing Canary session cookies to detect the theft and malicious use of credentials.

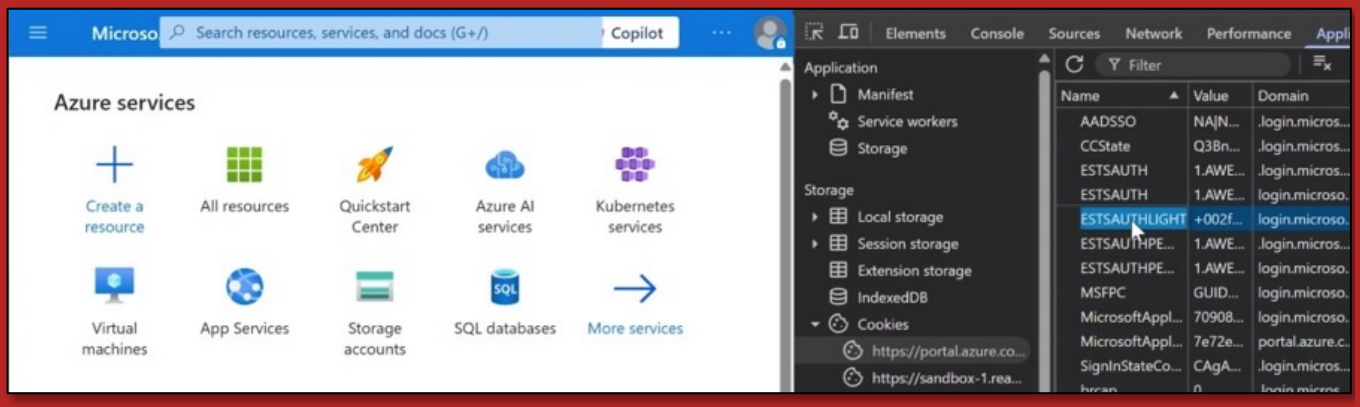


FIGURE 4: SUCCESSFULLY HIJACKED SESSION (ATTACKER)

Forensic Investigation of Bluetooth-Based Credit Card Skimmers

by John Passaro READ THE RESEARCH

Hidden Bluetooth Low Energy (BLE) credit skimmers are a growing threat to credit card fraud. Criminals can set up practical and inexpensive systems built on top of modules, such as the HM-19, to collect and transmit stolen data covertly across wireless channels. Criminals are utilizing modern technology to complicate traditional forensic processes by reducing the device's footprint, encrypting onboard storage, and creating potentially unpredictable behavior. New forensic processes must be generated to account for the increasing changes in technology. Customized Python-based scripts can be generated to assist with capturing live Bluetooth Low-Energy (BLE) data based on known patterns within historical devices and interfaces.

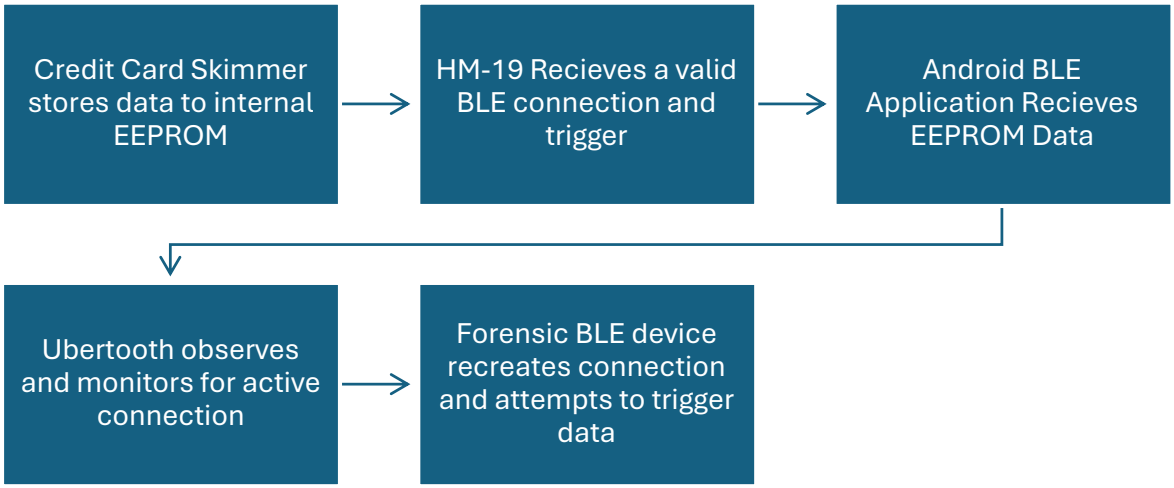


FIGURE 3: DATA FLOW DIAGRAM

STUDENT **HIGHLIGHT**

Webs of Deception: Using the SANS ICS Kill Chain to Flip the Advantage to the Defender

by Oren Niskin

READ THE RESEARCH

Defending a small Industrial Control System (ICS) against sophisticated threats can seem futile. The resource disparity between small ICS defenders and sophisticated attackers poses a significant security challenge. Communities rely on small ICS organizations to provide critical services like electricity and clean water. IT and ICS teams are managed separately and have distinct cultures within many small ICS organizations. Traditional ICS defense strategies primarily focus on monitoring the ICS network for threats. However, once the attacker is inside the ICS network, defenders' opportunities to prevent an incident become more limited. By looking for malicious activity across the wider attack chain, the SANS ICS Cyber Kill Chain provides the defender more opportunities to block and detect threats earlier in the attack chain.

Using the SANS ICS Cyber Kill Chain, the research below implemented a representative ICS network to evaluate the effectiveness of security controls for use by small ICS defenders. Complementing typical ICS security controls like firewalls and secure remote access, the research identified three high-leverage deception tactics that are simple to implement and add high-confidence opportunities to detect threats of any sophistication. This research seeks to reinforce that "*Defense is doable*."

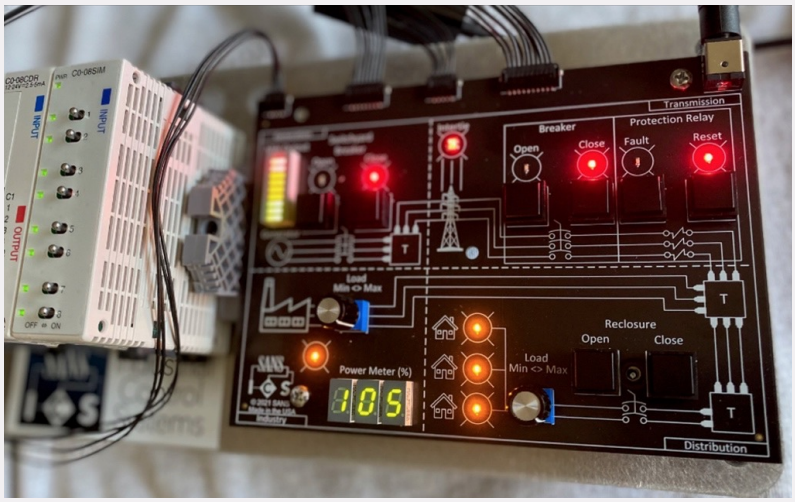


FIGURE 2:
SANS ICS515 COURSE KIT
(LEE, 2021).

"Oren extends common cyber deception techniques into the ICS realm to provide an early warning system to detect malicious insiders and threat actors probing to learn about and access OT systems. Development of low-cost and easy-to-deploy artifacts provides the opportunity for high-fidelity alerts within the business network, giving defenders an edge against attackers wishing to interrupt control system processes."

– DAVID FLETCHER, FACULTY RESEARCH ADVISOR

Code Modularity as a Heuristic for Malware Design

by Joseph Edwards

READ THE RESEARCH

Malware targeting industrial control systems (ICS) and critical infrastructure often exhibits a modular architecture, using a central loader to execute interchangeable payload modules. Well-known ICS malware such as Industroyer/CrashOverride and Havex were designed to support plugins for scanning or exploitation. This research investigates whether modularity and plugin-based architecture can serve as a static heuristic to detect new or unknown malware. The experimental method tested a set of YARA hunting signatures that match code patterns characteristic of modular loaders. The methodology involved reverse-engineering known modular malware samples to identify unique code patterns that facilitate loading plugins. Four versions of the signature were tested: a Normal rule with wildcarded addresses, a Loose rule wildcarding common stack constants, an experimental "Stack Constants" rule prioritizing mainly constants and structures on the stack, and a refined version of the Stack Constants rule. Each rule was tested via retro-hunting over the past year on two major malware repositories. By the end of the fourth phase, the total matches had increased by 28% from the baseline, with a False Positive rate of 0%. These results confirm that modular code can be an indicator of advanced malware. However, the relatively low number of overall matches reinforces the need for more flexible signatures, a dynamic in-memory scanning approach and the importance of unpacking technologies.

FIGURE 2:
HAVEX LOADER CALLING PLUGIN EXPORTS

```
int32_t __fastcall loadModuleRunExport(struct moduleStruct* arg1)

1002902c        int32_t* var_1c = &var_30
1002902f        void* const var_34_2 = &data_1001124d
10029035        int32_t var_8_4 = 0xfffffffe
1002903c        int32_t var_c = var_8_3
10029042        fsbase->NtTib.ExceptionList = &ExceptionList
1001124f        struct moduleStruct* var_20 = arg1
10011256        WCHAR* modulePath
10011256
10011256        if (arg1->field_30 u< 8)
1001125d            modulePath = &arg1->modulePath
10011256        else
10011258            modulePath = arg1->modulePath
10011258
10011261        HMODULE hModule = LoadLibraryW(lpLibFileName: modulePath)
10011267        arg1->hmodule = hModule
10011267
1001126c        if (hModule != 0)
1001127a            int32_t eax = GetProcAddress(hModule, lpProcName: "runDll")
1001127a
1001127e            if (eax != 0)
1001128e            label_1001128e:
1001128e                arg1->exportFound = true
10011292                int32_t var_8_1 = 0
10011296                eax()
10011298                int32_t var_8_2 = 0xfffffffe
1001129f                sub_100112a9(arg1)
1001127e            else
10011288                eax = GetProcAddress(hModule: arg1->hmodule, lpProcName: "RunDllEntry")
10011288
1001128c                if (eax != 0)
1001128c                    goto label_1001128e
1001128c
100112c3                if (sub_1000e55b() == 0)
100112c6                    deleteFile(&arg1->modulePath2)
100112c6
```

Defensible IEC 61850 Substation Network Security Monitoring with Zeek

by Elliot Lee

READ THE RESEARCH

Modern substations built on the IEC 61850 protocol family (MMS, GOOSE, SMV) expand operational flexibility but also widen the cyberattack surface. It leaves critical protection and control functions vulnerable to identity spoofing, false data injection, denial-of-service attacks, and reconnaissance. While commercial monitoring platforms exist, they rely heavily on deep packet inspection and proprietary decoders, which can be costly and inflexible for utilities. Open-source solutions remain limited, leaving a gap for affordable, protocol-aware defenses. This study introduces a Zeek-based monitoring framework that leverages transport layer and layer two invariants, such as MAC and VLAN integrity, multicast group membership, traffic rates, and MMS connection behavior, to detect the most consequential precursors to substation misoperation. Using reproducible lab PCAPs, the framework validates lightweight detectors for baseline discovery, false data injection precursors, denial-of-service, spoofing, and eavesdropping/exposure. The framework also demonstrates that an allowlist-plus-threshold approach can deliver high-signal, low-overhead alerts. The result is a practical, low-cost monitoring posture that utilities can deploy or extend within change-controlled environments to strengthen the defensibility of IEC 61850 substation networks.

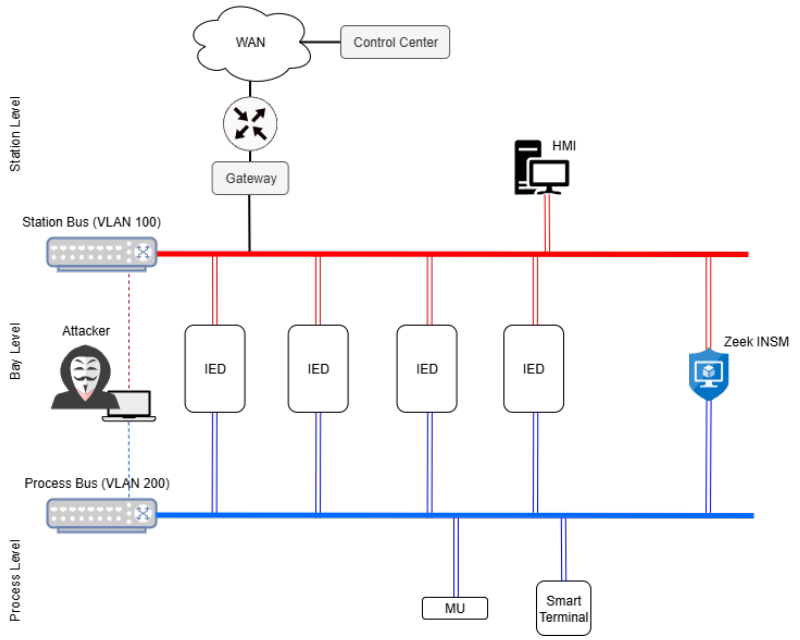


FIGURE 2.1:
IEC 61850 TEST LAB DIAGRAM

STUDENT **HIGHLIGHT**

Structural Vulnerability: Autodesk Revit Server WAN Exposure Versus Cost of Autodesk Construction Cloud

by Joshua Hall

READ THE RESEARCH

Autodesk Revit Server, a critical collaboration tool in the architecture, engineering, and construction (AEC) industry, was designed to operate within trusted networks. However, cost pressures and convenience create incentives to adopt an anti-pattern of forwarding port 808 through firewalls, exposing the service directly to the public internet. The result is open, unauthenticated access to business-critical design models: any external actor with foreknowledge of Revit Server's operation and a target IP can connect, acquire, and modify elements without challenge. Packet captures confirm that Revit transmits Simple Object Access Protocol (SOAP) operations and project metadata in clear text.

This paper argues that security leaders must move beyond reflexive prohibitions and vendor defaults, offering cost-conscious and workable alternatives. VPNs with multi-factor authentication, a Security Service Edge (SSE), and layered compensating controls offer viable ways to close unrestricted access while respecting the financial realities of project-driven organizations.

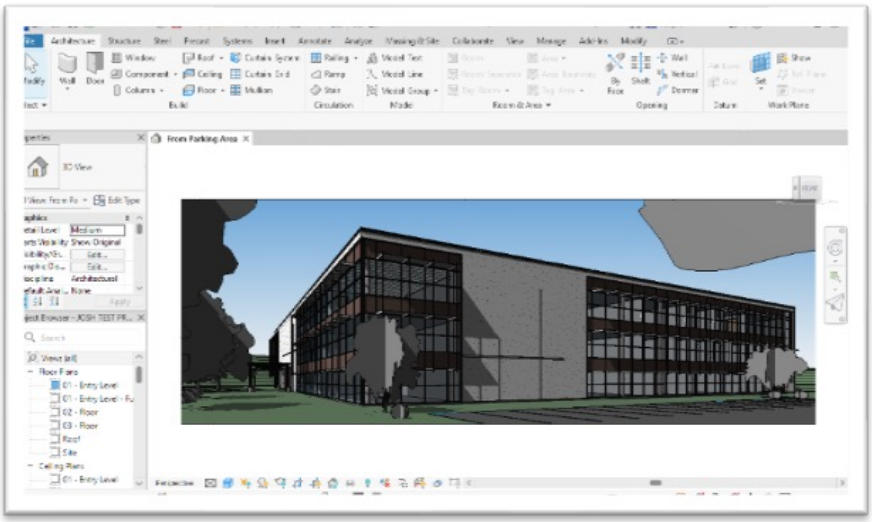


FIGURE 1:
SAMPLE REVIT MODEL

"Josh Hall tackled a problem that many security practitioners have concerns about, but that is rarely systematically analyzed. He conducted a comparative analysis of Autodesk Revit Server WAN exposure versus cloud alternatives, examining the security implications and business costs when organizations choose convenience over secure architecture. Josh didn't just identify potential risks, but provided a framework for evaluating the trade-offs and practical alternatives that work within operational constraints. His analysis equips security teams with a documented assessment and business case to discuss these architectural trade-offs with management. Such research moves the industry forward by analyzing real-world security concerns and providing actionable recommendations."

– **LENNY ZELTSER, FACULTY RESEARCH ADVISOR**

Privacy Protections: Are Stronger Laws Changing What We Reveal?

by Katie Christensen READ THE RESEARCH

As U.S. states enact privacy laws aimed at giving consumers more control over their personal data, little is known about whether privacy legislation influences individuals' willingness to disclose their identity on public platforms. This study addresses that gap by analyzing 22,000 Google Maps reviews across all 50 U.S. states, classifying reviewer names to determine whether users disclosed their real identities or used pseudonyms. Previous research has primarily focused on privacy attitudes rather than actual disclosure behaviors, and prior studies often overlook the intersection of demographic and cultural factors with legal protections. Reviewers in states with more mature privacy laws tended to disclose their real names less often. In contrast, states with recent, partial, or no privacy laws exhibited higher rates of real-name disclosure. These results suggest an association between stronger privacy legislation and reduced disclosure behavior, but only when laws are both comprehensive and have been in effect long enough to influence public awareness or enforcement practices. Although group differences were not statistically significant in the chi-square test ($p = .0921$), regression and correlation analyses revealed a moderate, statistically significant inverse relationship between law maturity and disclosure behavior ($r = -0.442$, $p = .0017$). This suggests that behavioral shifts may emerge more clearly when privacy law maturity is analyzed as a continuous variable rather than a categorical one.

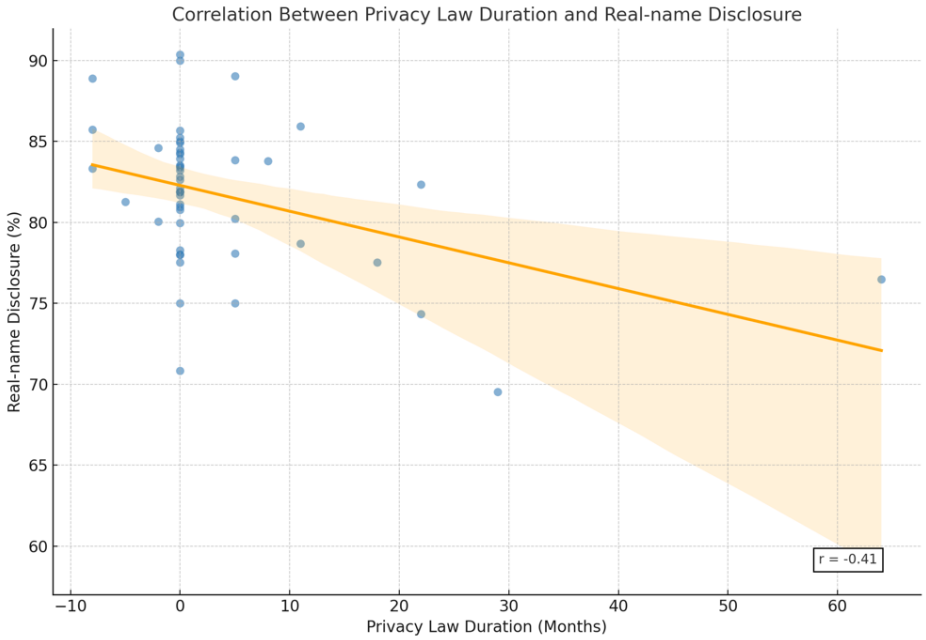


FIGURE 10: REAL-NAME DISCLOSURE VS. RURAL POPULATION

STUDENT **HIGHLIGHT**

The Mimic Octopus: Weaponizing File Corruption and Recoverability to Bypass Antivirus and Email Filtering

by Justin Gazick READ THE RESEARCH

This paper investigates a novel tactic in phishing operations where threat actors intentionally corrupt document and archive files, such as DOCX, DOCM, PDF, and ZIP, to evade antivirus (AV) and email filtering systems. These files, though malformed, are recoverable by native tools like Microsoft Word, Adobe Reader, and WinRAR. As a result, malicious payloads can still execute after delivery. Building on prior findings by Any.Run (Any.Run, 2024), this study expands the corruption methodology to include multiple structural modifications and evaluates their impact on AV detection via VirusTotal and behavior in the Any.Run sandbox. A custom corruption suite and detection tool were developed to automate corruption detection and analyze results across formats.

Findings reveal that minor corruption can suppress static detection without disrupting file recoverability or execution, exposing gaps in modern AV pipelines. The results underscore the need for advanced structural integrity checks and sandbox automation to simulate user-driven recovery. This research contributes actionable recommendations for enhancing threat detection and lays the groundwork for future defenses against corruption-based evasion tactics.

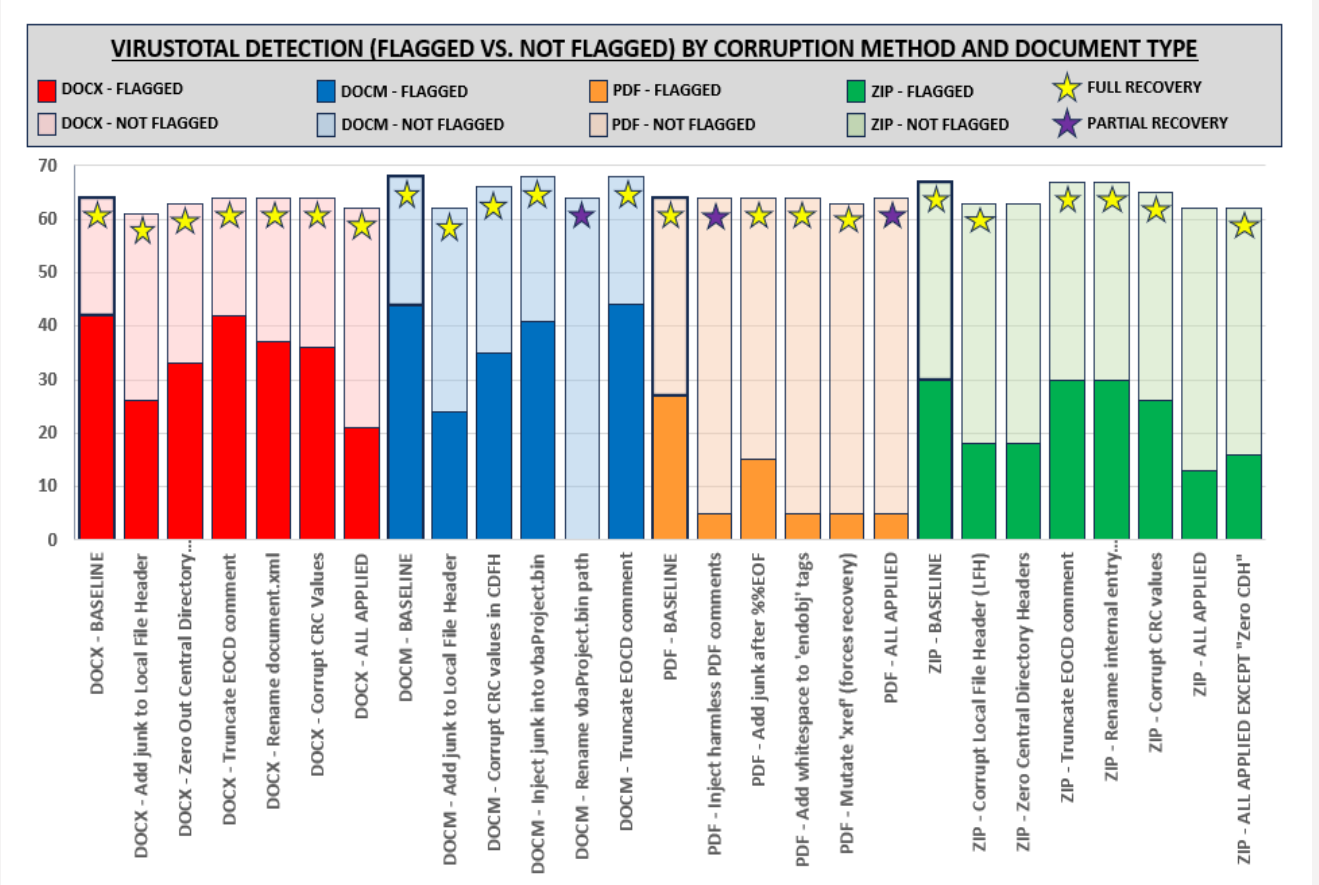


FIGURE 3:
VIRUSTOTAL DETECTIONS (FLAGGED VS. NOT FLAGGED) BY CORRUPTION METHOD AND DOCUMENT TYPE, WITH RECOVERY OUTCOMES

The Mimic Octopus: Weaponizing File Corruption and Recoverability to Bypass Antivirus and Email Filtering CONT.

"Justin's research highlights a creative evasion technique in which attackers intentionally corrupt document and archive files to bypass antivirus and email filtering systems while remaining recoverable by common user applications. Highlighting this in DOCX, DOCM, PDF, and ZIP formats, the paper demonstrates how minor structural corruption can suppress detection while still enabling malicious payload delivery. As phishing continues to be a dominant initial access vector, this work provides valuable insight for defenders responsible for email security, malware detection, and endpoint protection."

– JONATHAN RISTO, FACULTY RESEARCH ADVISOR

Interrogators: Attack Surface Mapping in an Agentic World

by Michael Samson READ THE RESEARCH

This research introduces the concept of AI agent interrogators and the open-source project Agent Interrogator, an opaque box interrogation framework designed to map the attack surface of agentic systems. As the adoption of AI agents rapidly expands, there is a growing need to develop the ability to map the attack surface behind their natural language interface, which traditional security tooling cannot accomplish. Through Agent Interrogator, a two-stage AI-assisted interrogation process is employed. Identifying the agent's high-level capabilities in the initial interrogation stage and then enumerating the supporting invokable tools for each capability. This research validates the approach against test targets utilizing LangChain and Model Context Protocol (MCP) to deliver agentic capabilities. The product of the interrogation is a structured profile mapping the agent's attack surface, enabling security practitioners to identify vulnerabilities such as excessive agency and conduct targeted fuzzing. This work provides a critical foundation for securing the next generation of AI systems and the development of automated attack surface mapping in complex, multi-agent ecosystems.

	DVLA	DVLA w/ MCP	DVLA Without Tools
Tool Detection	100%	100%	N/A
Hallucinations	0	0	0
Detection Frequency Range	95%	45% - 100%	N/A

FIGURE 10: OVERALL FINDINGS SUMMARY

From Crash to Compromise: Unlocking the Potential of Windows Crash Dumps in Offensive Security

by Jason Mull READ THE RESEARCH

Windows crash dump files, frequently overlooked in offensive contexts despite their forensic value, contain several sensitive elements that threat actors can exploit for privilege escalation, credential harvesting, lateral movement, and data exfiltration. Furthermore, the operating system already generates these crash dumps, so a threat actor does not need to risk discovery by dumping sensitive processes using noisier, more traditional methods. Offensive security education often emphasizes exploit development, network pivoting, and credential theft but largely overlooks the value of memory forensics in post-exploitation scenarios. This research explores how offensive security practitioners can incorporate crash dump analysis into their workflows to extract sensitive data such as plaintext credentials, encryption keys, and files from memory. It also investigates methods to detect the creation of crash dump files, allowing defensive practitioners to identify and respond to their presence within an organization.

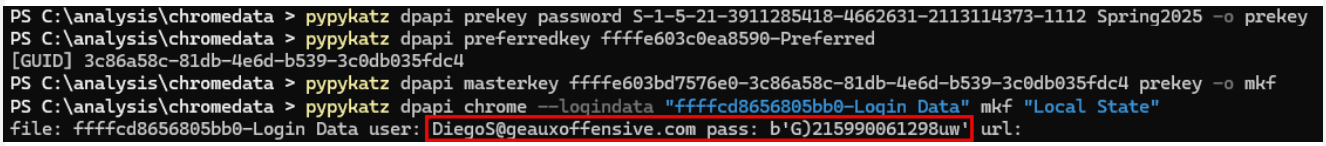


FIGURE 11: COMPLETE MEMORY DUMP - CHROME PASSWORD ACQUISITION

SANS Technology Institute offers undergraduate and graduate programs on the cutting edge of information security.

SANS TECHNOLOGY INSTITUTE

sans.edu/research