CLAWS Newsletter

Cyber Index | Volume II | Issue 15

by Govind Nelika

@govindnelika
govind-nelika-4217969b

https://claws.co.in/category/newsletter/

* CLAWS Cyber Index Newsletter is a concise Bi-Monthly brief delivering Strategic Insights on Global Cyber Threats, Policy Shifts, and Geopolitical Technology Developments.

About us

The Centre for Land Warfare Studies (CLAWS) is an independent think tank based in New Delhi, India, dedicated to strategic studies and land warfare in the Indian context. Established in 2004 and registered under the Societies Registration Act, 1860, CLAWS operates as a membership-based organization governed by a Board of Governors and an Executive Council, under the Aegis of the Indian Army.

With a futuristic outlook and a policy-oriented approach, CLAWS focuses on national security issues, conventional military operations, and sub-conventional warfare. The Centre closely monitors regional conflicts and military developments within India's strategic frontiers, particularly in South Asia.

Committed to fostering strategic culture and informed policymaking, CLAWS disseminates its research to armed forces personnel, policymakers, members of the strategic community, and interested civilians. By facilitating in-depth studies and discussions, CLAWS contributes to shaping India's defense policies and military preparedness.

The CLAWS Newsletter is a fortnightly series under the leadership of Dr. Tara Kartha, Director Research & Academics. The newsletter features insightful content curated by CLAWS researchers, each specializing in their respective verticals. This initiative aims to provide in-depth analysis, strategic insights, and updates on key issues.

Contents

Internal

CERT-In Advisory on Emerging Threats Targeting Microsoft 365 (M365)

In a critical warning to enterprise defenders, the Indian Computer Emergency Response Team (CERT-In) has issued an advisory detailing a severe escalation in targeted cyber-attacks against Microsoft 365 (M365) environments, orchestrated through sophisticated identity compromise tactics. As organizations increasingly rely on cloud-based collaboration platforms, threat actors are aggressively exploiting legacy authentication protocols and identity workflows to bypass Multi-Factor Authentication (MFA) and establish persistent, unauthorized access to sensitive tenants. The observed campaigns employ a multi-pronged approach, heavily leveraging large-scale password spraying that abuses the Azure Resource Owner Password Credentials (ROPC) OAuth flow. Concurrently, attackers are utilizing subscription-based Phishing-as-a-Service (PhaaS) platforms to launch device-code phishing attacks. By tricking victims into authorizing attacker-controlled device codes via legitimate Microsoft sign-in pages, adversaries seamlessly capture OAuth access tokens and hijack authenticated sessions without directly stealing credentials.

Once initial access is achieved, the threat actors register rogue devices in Microsoft Entra ID and deploy hidden mailbox rules to maintain stealthy persistence. These compromised accounts often including trusted supplier or vendor identities are subsequently weaponized to conduct Business Email Compromise (BEC), facilitate lateral movement, and exfiltrate confidential data from Exchange Online, SharePoint, OneDrive, and Teams for financial extortion. To mitigate these advanced identity-centric threats, CERT-In urges organizations to immediately disable legacy authentication methods like ROPC, restrict device-code authentication strictly to emergency accounts, and transition toward phishing-resistant MFA frameworks such as FIDO2. For risk managers and cloud security architects, this development underscores a fundamental shift in the threat landscape: adversaries are no longer simply hacking networks; they are systematically logging in by weaponizing the very OAuth and API workflows designed to facilitate seamless cloud access, necessitating a rapid pivot to continuous access evaluation and stringent Conditional Access policies.

Read more: https://www.cert-in.org.in/s2cMainServlet?pageid=PUBVLNOTES02&VLCODE=CIAD-2026-0037

Bharat Electronics Limited, Esri India sign MoU to develop GIS, GeoAI solutions for India's defence sector

In a pivotal move to modernize military intelligence architectures, state-owned defense contractor Bharat Electronics Limited (BEL) has formalized a strategic partnership with mapping and spatial analytics firm Esri India to co-develop advanced Geographic Information System (GIS) and GeoAI solutions for the Indian armed forces. Amidst a geopolitical landscape where algorithmic warfare and multi-domain operational awareness dictate tactical superiority, the integration of artificial intelligence with spatial data has become a critical strategic requirement. The Memorandum of Understanding (MoU) establishes a framework to fuse BEL's extensive portfolio of defence electronics, unmanned systems, and C4I (Command, Control, Communications, Computers, and Intelligence) platforms with Esri's sophisticated location intelligence capabilities. Specifically, the initiative centres on deploying Esri's 'Indo ArcGIS' platform to build resilient, mission-ready geospatial capabilities tailored to urgent operational requirements.

As part of this capacity-building effort, Esri India will execute targeted training programs to establish in-house GeoAI proficiency within BEL. By replacing fragmented legacy mapping systems with AI-driven spatial intelligence, the collaboration aims to provide military commanders with accelerated threat analysis, enhanced situational awareness, and precision response coordination across contested environments. Crucially, the partnership aligns with the government's 'Make in India' and Atmanirbhar Bharat mandates, directly mitigating supply chain risks and reducing reliance on foreign intelligence infrastructure. For defence sector stakeholders and security architects, this convergence of sovereign defence hardware and advanced GeoAI signals a fundamental capability upgrade. It underscores a broader systemic shift wherein indigenous technological ecosystems are mobilized to secure national defence networks, ensuring that strategic geospatial

data remains under sovereign control while equipping operational forces to navigate complex, data-centric threat landscapes.

Read more: https://economictimes.indiatimes.com/news/defence/bharat-electronics-and-esri-india-join-forces-to-revolutionize-indias-defence-with-gis-and-geoai-solutions/articleshow/132836815.cms?from=mdr

Inside India's Rs 1 trillion plan to take defence prototypes towards production starting with YETI

In a strategic effort to bridge the "missing middle" of defence procurement, the Indian government has operationalized its ₹1-lakh-crore Research, Development and Innovation (RDI) Scheme, targeting the commercial scale-up of sovereign deep-tech and cyber-physical systems. Amidst escalating geopolitical tensions and a global pivot toward autonomous warfare, emerging defence contractors frequently face a "valley of death" where functional prototypes exhaust capital before securing dependable military acquisition. The RDI Scheme disrupts this cycle by providing milestone-linked debt financing with maximum tenures of up to 15 years for strategic technologies at Technology Readiness Level (TRL) 4 and above, covering up to 50% of assessed project costs.

The inaugural beneficiary of this capital injection is ideaForge's Project YETI, which secured a Letter of Intent for ₹151 crore in debt financing. Project YETI is an autonomous, fixed-wing vertical take-off and landing (VTOL) platform engineered for military logistics, boasting a developmental payload capacity of up to 200 kg and a mission range of 200 km. Beyond autonomous aviation, the six-year initiative explicitly prioritizes critical sunrise domains, including quantum technologies, artificial intelligence, robotics, and biotechnology. By absorbing the industrialization risk that traditional banks and venture investors typically avoid, the Ministry of Defence is actively decentralizing its defence-industrial base. For national security stakeholders and risk managers, this financing model represents a vital evolution in cyber-physical resilience. Transitioning from reliance on foreign original equipment manufacturers (OEMs) to a mature, indigenous ecosystem mitigates critical supply chain vulnerabilities and ensures that the advanced capabilities required for future multidomain operations can successfully survive the leap from laboratory demonstration to active theatre deployment.

Read more: https://www.financialexpress.com/business/defence/inside-indias-rs-1-trillion-plan-to-take-defence-prototypes-towards-productionstarting-with-yeti/4318761/

External

Global Focus Brief

Visa beefs up cybersecurity offerings with $2.4 billion BioCatch deal

In a major consolidation of the financial cybersecurity landscape, global payments giant Visa has reached a definitive agreement to acquire Israeli behavioural biometrics pioneer BioCatch for $2.4 billion in cash. This strategic acquisition arrives as the global economy faces over $1 trillion in annual losses from account takeovers, first-party fraud, and scams, increasingly supercharged by artificial intelligence. With real-time payments compressing the window for fraud detection, traditional transaction monitoring is no longer sufficient, driving payment networks to prioritize upstream threat prevention before a transaction ever occurs. BioCatch specializes in behavioural-first, multi-signal fraud intelligence, shifting identity verification from static credentials to continuous behavioural analysis. Instead of relying on passwords, its artificial intelligence and machine-learning models process over 3,000 anonymized signals per millisecond including typing cadence, touchscreen gestures, mouse movements, device handling angles, and signs of hesitation or coercion. By analysing 19 billion monthly user sessions across 1.8 billion devices, the platform protects 760 million users spanning 350 global banking institutions.

Crucially, the technology has evolved to counter sophisticated AI-driven threats, utilizing recent innovations like DeviceIQ to detect agentic browsers, deepfake injections, robotic process automation, and jailbroken devices. By identifying anomalies at account registration or login, the system proactively disrupts money mules and social engineering scams in real time. For risk managers and enterprise defenders, Visa's acquisition signals a fundamental evolution in cyber resilience and identity architecture. By integrating behavioral biometrics directly into its core infrastructure, Visa is constructing a proprietary trust layer that addresses

the vulnerabilities of agentic commerce and AI-enabled financial crime. This move underscores a broader industry trend where the boundaries between payment processing and proactive cybersecurity are dissolving, equipping financial institutions with the continuous intelligence required to distinguish legitimate human intent from sophisticated, automated exploitation.

Read more: https://www.reuters.com/legal/transactional/visa-buy-fraud-intelligence-provider-biocatch-24-billion-2026-08-03/

Stealing Reasoning Traces from Proprietary LLM APIs

In a critical architectural failure affecting the foundation model ecosystem, researchers have uncovered a systemic vulnerability across major AI providers including OpenAI, Anthropic, and Google that allows attackers to decrypt and steal concealed chain-of-thought (CoT) reasoning traces. As AI vendors increasingly hide their models' step-by-step reasoning to protect intellectual property and prevent model distillation, they have relied on returning these traces to clients as opaque, encrypted text blocks for state management. However, this approach inadvertently introduced a severe flaw: a lack of cryptographic contextual binding. Because these encrypted blocks were fully interchangeable across different sessions, user accounts, and even distinct models within a single provider's infrastructure, threat actors could execute a scalable "decryption jailbreak." By injecting an encrypted reasoning trace generated by a highly capable, heavily safeguarded model into a weaker, more permissive sibling model, attackers forced the latter to decode and output the proprietary reasoning verbatim in plaintext.

This vulnerability enabled four distinct exploitation vectors. Beyond circumventing anti-distillation mechanisms to facilitate AI intellectual property theft, the flaw exposed significant privacy risks. By scraping and decoding 315,320 reasoning blocks inadvertently published by developers in public repositories, researchers recovered 367 personally identifiable information (PII) artifacts and 182 credentials. Furthermore, the exploit allowed adversaries to surface hazardous content hidden within the reasoning steps of otherwise safely refused queries, and to smuggle invisible prompt-injection payloads inside the encrypted blocks to poison public agentic rollouts. Following coordinated disclosure, vendors have deployed server-side patches to bind reasoning blocks to their originating contexts, effectively neutralizing the immediate exploit. Nevertheless, this development exposes a critical blind spot in client-side state management for AI APIs, underscoring the urgent necessity for robust cryptographic isolation and highlighting the residual data exposure risks for organizations that previously logged and shared these ostensibly secure artifacts.

Read more: https://arxiv.org/abs/2608.09867

U.S. Cyber Command plans Silicon Valley office to drive innovation

To accelerate the operationalization of advanced digital warfare capabilities, United States Cyber Command is establishing a strategic outpost in Silicon Valley, designated Cyber Command-West (CC-W), as a cornerstone of its "CYBERCOM 2.0" modernization initiative. Operating within a geopolitical landscape where state-sponsored threat actors increasingly leverage asymmetric and artificial intelligence-augmented tactics, the rapid integration of commercial technological innovation has become a critical imperative for national defense. This strategic deployment aims to bridge the persistent procurement gap often termed the "valley of death" between commercial technology demonstration and active military deployment. Co-located initially with the Pentagon's Defense Innovation Unit (DIU), the CC-W satellite office will directly support the newly formed Cyber Innovation Warfare Center (CIWC).

The operational framework centers on rapid, iterative prototyping by placing military cyber operators and private-sector software engineers into a shared, collaborative environment to test emerging concepts against realistic threat scenarios and real-world telemetry. By establishing a direct physical footprint on the West Coast, U.S. Cyber Command seeks to circumvent traditional bureaucratic bottlenecks, directly engaging established enterprise entities and specialized startups to cultivate specialized talent and acquire cutting-edge

offensive and defensive tooling. This structural evolution reflects a broader paradigm shift in national security doctrine, explicitly acknowledging that maintaining tactical superiority and systemic resilience in the modern threat landscape requires an agile, symbiotic public-private partnership. For risk management practitioners and strategic policy stakeholders, the activation of CC-W underscores an ongoing decentralization of military cyber capability development, signaling that future national cyber resilience will depend heavily on frictionless, continuous collaboration between military apparatuses and commercial technology innovators to effectively counteract highly adaptive global adversaries.

Read more: https://therecord.media/cyber-command-plans-silicon-valley-office-to-drive-innovation

AI-Assisted Code Can Alter Forensic DNA Scan Files Without Any Detectable Trace

In a critical development that fundamentally threatens the integrity of digital forensics and the criminal justice system, security researchers have demonstrated the ability to use AI-assisted code to undetectably alter computerized DNA scan files generated by standard crime-lab equipment. Emerging against a backdrop of increasingly accessible AI-driven exploit generation and escalating attacks on civic infrastructure, this vulnerability shatters the long-held assumption that digitized physical evidence remains tamper-proof. According to the research, threat actors can leverage AI-crafted malware to intercept and manipulate raw genomic data directly as it is processed by forensic scanning machines. By injecting malicious code capable of modifying or entirely swapping DNA profiles before the data is finalized, logged, or cryptographically hashed, attackers can fabricate or erase evidentiary ties without triggering standard integrity alerts or leaving detectable digital footprints.

The attack targets the vulnerable operational technology (OT) and legacy data pipelines within forensic laboratories, seamlessly compromising the digital chain of custody at the point of conversion from a physical sample to a digital file. This discovery underscores a severe systemic risk for law enforcement, judicial bodies, and enterprise security leaders relying on biometric or forensic data. The implications extend far beyond isolated data breaches, directly threatening legal outcomes and the foundational trust in forensic science. As AI lowers the barrier for executing highly sophisticated data-manipulation attacks, mitigating this threat will require an urgent overhaul of crime-lab cybersecurity postures, necessitating the implementation of zero-trust architectures, strictly air-gapped processing environments, and hardware-level cryptographic signing of forensic data at the exact moment of generation to ensure true non-repudiation.

Read more: https://www.techtimes.com/articles/322771/20260803/ai-assisted-code-can-alter-forensic-dna-scan-files-without-any-detectable-trace.htm

Department of the Navy Establishes Direct Reporting Portfolio Manager for Robotic and Autonomous Systems

In a strategic overhaul of its technology acquisition pipeline, the U.S. Department of the Navy (DON) has established the Direct Reporting Portfolio Manager for Robotic and Autonomous Systems (DRPM RAS) to centralize and accelerate the deployment of unmanned maritime technologies. Set against a backdrop of intensifying geopolitical competition and a paradigm shift toward AI-driven naval warfare, this organizational pivot addresses a critical vulnerability in defense tech integration. For years, fragmented acquisition practices and legacy bureaucratic processes have bottlenecked the deployment of advanced autonomous capabilities, diluting accountability and complicating engagements with the defense industrial base. As nation-states increasingly leverage unmanned systems for asymmetric maritime operations and intelligence gathering, rapidly transitioning these emerging technologies from development to the operational theater has become a strict imperative for national security decision-makers.

Operationally, the DON is adapting a proven, centralized management framework originally engineered for the nuclear submarine fleet to its robotics portfolio. Directed by Acting Secretary of the Navy Hung Cao, the initiative places Christopher Miller as the acting DRPM RAS and Portfolio Acquisition Executive, reporting directly to the Under Secretary level. This unified command structure will synchronize with broader Department

of War frameworks for unmanned offensive and defensive systems (UxS). To further tighten the feedback loop between engineers and end-users, the DON is actively evaluating the relocation of DRPM RAS personnel away from Washington, D.C., aiming to embed them directly with Fleet warfighters to ensure capability development is driven by immediate tactical realities. This structural realignment signals a broader trend in global defense risk management: treating robotic and autonomous systems as core strategic platforms requiring streamlined governance. By dismantling bureaucratic friction and accelerating industry partnerships, the DON is fortifying its operational resilience, reinforcing maritime deterrence, and ensuring sustained technological overmatch in an increasingly automated threat landscape.

Read more: https://www.navy.mil/Press-Office/Press-Releases/display-pressreleases/Article/4565150/department-of-the-navy-establishes-direct-reporting-portfolio-manager-for-robot/

United States of America (USA)

US Air Force jet engine manufacturing plagued by 'significant challenges'

In a critical development for defense supply chain resilience and the broader technological risk landscape, the U.S. Air Force has initiated a sweeping overhaul of its fighter jet engine procurement strategy in response to severe manufacturing bottlenecks, quality control failures, and critical component obsolescence across its current contractor ecosystem. Amid escalating geopolitical tensions most notably China's restrictions on the export of rare earth elements the fragility of the defense industrial base has emerged as a systemic vulnerability threatening national security and operational readiness. Addressing these structural weaknesses, the Air Force has issued a formal Request for Information (deadline August 28) demanding a paradigm shift toward industry-led technological evolution, supply chain hardening, and lifecycle accountability. This strategic policy move is driven by compounding operational setbacks in high-profile platforms: chronic contractor deficiencies have persistently derailed F-35 deliveries and upgrades, while B-52 modernization efforts have suffered a $3 billion cost overrun and a 15-month delay in initial operational capability directly tied to engine complications.

To mitigate these hardware and logistical risks, the Air Force plans to scale up procurement to over 180 engines annually by 2034, specifically targeting the F-15EX and F-16 fleets. The initiative enforces stringent producibility requirements, compelling prospective partners to map out raw material dependencies such as specialized titanium and nickel alloys and resolve advanced casting or forging constraints. Furthermore, vendors must demonstrate how minimum production volumes will incentivize private capital investments in automated tooling and facility expansion to structurally lower long-term unit costs. Ultimately, this initiative underscores a vital evolution in national security risk management: treating the defense manufacturing supply chain with the same urgency as critical infrastructure protection, where establishing resilient, scalable, and fully independent production capabilities is now a prerequisite for maintaining air superiority and global stability in an increasingly contested environment.

Read more: https://www.militarytimes.com/industry/techwatch/2026/08/03/us-air-force-jet-engine-manufacturing-plagued-by-significant-challenges/

US to boost Patriot and THAAD interceptor production as missile stockpiles come under pressure

In a critical move to harden defense supply chains against escalating global kinetic threats, the U.S. Department of Defense (DoD) has finalized framework agreements exceeding $3 billion with major aerospace contractors Lockheed Martin and Northrop Grumman to rapidly scale the production of Patriot PAC-3 and Terminal High Altitude Area Defense (THAAD) interceptors. This strategic policy shift emerges within a fraught geopolitical landscape where prolonged, high-volume engagements particularly the conflicts in Ukraine and Iran involving extensive ballistic missile and drone barrages have severely depleted U.S. precision munition stockpiles. Recent strategic assessments indicate the U.S. military inventory has dwindled to fewer than 1,000 Patriot and 250 THAAD interceptors, exposing a significant vulnerability in the resilience of Western defensive architectures.

To rectify this operational deficit and secure the defense industrial base, the DoD is structurally altering its acquisition strategy by engaging directly with component-level munition suppliers rather than relying exclusively on prime contractors. Operationally, the agreements aim to triple the production capacity for Patriot interceptors and quadruple the output for THAAD systems, ultimately supporting a long-term procurement objective of nearly 14,000 interceptors. Together, these systems comprise the backbone of the U.S. layered missile defense network, with THAAD providing first-line defense against high-altitude, long-range ballistic threats, while the PAC-3 system serves as a lower-tier shield against cruise missiles and penetrating aerial threats. For defense stakeholders and risk management professionals, this development highlights a necessary evolution in supply chain resilience. By actively diversifying manufacturing networks down to the component level, the DoD is not only closing an immediate strategic capability gap but also preemptively adapting its infrastructure to withstand the sustained, high-intensity logistical attrition that increasingly characterizes the modern threat landscape.

Read more: https://timesofindia.indiatimes.com/defence/international/us-to-boost-patriot-and-thaad-

interceptor-production-as-missile-stockpiles-come-under-pressure/articleshow/132893075.cms

U.S. Army Expands Hydra-70 Rocket Supply Chain Through Innovative OTA Agreements

In a strategic move to fortify critical supply chains against single points of failure, the U.S. Department of the Army's Aviation Rocket and Small Guided Munitions (ARSGM) Product Office has executed a series of Other Transaction Authority (OTA) agreements to diversify production for the Hydra-70 rocket system. Set against a backdrop of escalating global tensions and an urgent mandate to revitalize the defense industrial base, this development highlights a shift toward agile procurement designed to eliminate vendor lock-in and mitigate systemic logistical risks. Leveraging the One Nation Innovation (ONI) marketplace, the Army awarded contracts to four defense sector vendors: Firehawk, iRocket, Nammo Perry, and Albers. Operationally, each contractor is tasked with delivering twenty-four prototype Hydra-70 M151 High Explosive rockets, coupled with a modernized Technical Data Package (TDP), within an 18-month timeframe. This initiative is engineered to yield immediate strategic dividends by establishing multiple viable production channels, thereby driving competitive pricing for annual procurements typically ranging from 100,000 to 200,000 units.

Furthermore, updating the government-owned product definition directly enhances manufacturability, reducing friction in future production cycles. While the focus remains on kinetic weaponry, the overarching policy move mirrors broader operational resilience strategies seen across critical infrastructure and cyber risk landscapes: expanding vendor ecosystems to secure surge capacity and ensure operational continuity during crises. By aggressively restructuring its munition supply lines, the U.S. Army is directly addressing vulnerabilities inherent in legacy, consolidated defense manufacturing. Ultimately, this framework not only ensures rapid response capabilities for emergent mission requirements but also serves as a blueprint for national security stakeholders aiming to balance competitive innovation with robust risk management, thereby enhancing long-term readiness and deterrence in an increasingly volatile strategic environment.

Read more: https://www.army.mil/article/294415/army_expands_hydra_70_rocket_supply_chain_through_innovative_ota_agreements

The Commonwealth of Australia

Australia advances AUKUS capability with dry dock design appointment

The Australian Department of Defence has advanced its strategic maritime capabilities under the AUKUS framework by appointing defense contractor Leidos Gibbs and Cox Australia to lead the preliminary design of a contingency floating dry dock in Western Australia. Set against the backdrop of escalating geopolitical tensions and a highly contested strategic environment in the Indo-Pacific region, this infrastructure development represents a critical risk mitigation measure for the allied defense industrial base. As nation-state actors expand their naval footprints and asymmetric capabilities, securing sovereign maintenance infrastructure for next-generation technological platforms has become an imperative for national security decision-makers. The primary operational development centers on establishing this domestic dry dock to support Australia's incoming conventionally-armed, nuclear-powered submarine fleet. Strategically, the facility serves as a mandatory pre-condition for the Royal Australian Navy to receive its first Virginia-class submarines in the early 2030s, and it establishes the necessary framework for complex depot-level maintenance scheduled for the late 2030s.

Integrated into a broader initial $12 billion sovereign investment across HMAS Stirling and the Henderson Defence Precinct, the project will operate under stringent international regulatory protocols governing nuclear safety, security, and safeguards. While situated in the physical defense sector, this policy move mirrors broader trends in supply chain resilience and critical infrastructure protection, emphasizing the need to eliminate single points of failure in strategic operational networks. By cementing domestic sustainment capabilities, allied nations reduce their exposure to logistical disruptions and systemic vulnerabilities that could be exploited by adversaries. Ultimately, this development substantially bolsters the resilience of the AUKUS alliance, ensuring that critical maritime assets can be sustained without over-reliance on external logistical hubs, thereby reinforcing collective deterrence, minimizing systemic risk, and enhancing long-term geopolitical stability across a volatile theatre.

Read more: https://www.minister.defence.gov.au/media-releases/2026-07-29/australia-advances-aukus-capability-dry-dock-design-appointment-0

European Union (EU)

The EU AI Act

The European Union has activated its landmark Artificial Intelligence Act, establishing a comprehensive, risk-based regulatory framework that holds AI developers, deployers, and general-purpose AI (GPAI) providers strictly accountable for the security, transparency, and ethical deployment of their systems. Operating in a threat landscape defined by rapidly escalating model capabilities and corresponding systemic risks, the European Commission and the newly empowered EU AI Office seek to enforce structural guardrails that balance technological innovation with the protection of critical infrastructure and fundamental rights. The legislation categorizes AI into distinct risk tiers, aggressively targeting malicious or highly invasive use cases. Systems posing an "unacceptable risk" including social scoring, untargeted biometric scraping, and harmful behavioral manipulation are strictly prohibited under rules that largely took effect in early 2025. Conversely, "high-risk" applications deployed in sensitive environments such as critical infrastructure, law enforcement, and employment screening must meet stringent compliance mandates by December 2027.

These obligations mandate robust pre-market risk assessments, continuous cybersecurity testing, rigorous human oversight, and comprehensive activity logging. Recent policy maneuvers, notably the July 2026 "AI Omnibus" regulation, have centralized the oversight of frontier GPAI models under the AI Office and explicitly banned non-consensual AI-generated explicit content. Concurrently, the Commission and ENISA have launched a targeted Action Plan to secure advanced AI systems across critical sectors like energy and finance. Furthermore, transparency rules enforceable as of August 2026 dictate that AI-generated content such as deepfakes or automated chatbot outputs must be clearly labeled, while GPAI providers are compelled to publish detailed summaries of their training data. Ultimately, this regulatory posture dictates a profound transformation in global cyber risk management. Security leaders and compliance analysts must now embed continuous post-market monitoring, exhaustive data traceability, and rigid third-party auditing into their AI development pipelines to ensure enterprise resilience and avoid crippling penalties across the European market.

Read more: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

EU launches AI Gigafactories call to boost Europe's computing capacity and unlock more than €30 billion in investment

The European Union has initiated a pivotal effort to assert technological sovereignty and mitigate systemic digital supply chain vulnerabilities by launching a call for tenders to establish up to seven advanced "AI Gigafactories" across the continent. Against a backdrop of intensifying geopolitical friction over artificial intelligence dominance and growing concerns regarding the security implications of relying on non-domestic computing architectures, this policy move signals a strategic shift aimed at safeguarding European data and algorithmic integrity. Fueled by up to €10 billion in joint EU and national public funding, the initiative is designed to unlock at least €20 billion in private sector investment to massively scale the bloc's localized compute capacity. Operationally, these Gigafactories will deploy next-generation AI processors, integrated software and cloud technology stacks, ultra-high-speed connectivity protocols, and highly energy-efficient data centers.

Augmenting the EU's existing operational network of 19 AI Factories, this expanded infrastructure will furnish startups, enterprise industry, academia, and public authorities with robust, localized environments to securely train, fine-tune, and execute inference on frontier AI models without exporting sensitive telemetry to offshore providers. From a risk management and compliance perspective, this localized architecture guarantees that state-of-the-art AI development will remain strictly tethered to European digital infrastructure, enforcing native compliance with the bloc's rigorous data protection, safety, and operational security standards. For defenders and national security stakeholders, this €30 billion mobilization represents a crucial structural fortification of the broader cyber risk landscape. By intentionally reducing reliance on foreign technological monopolies, the EU is actively attempting to shrink its strategic attack surface, insulate its critical emerging technology pipelines

from international volatility, and anchor the future of foundational AI firmly within the bounds of European regulatory control and cyber resilience.

Read more: https://digital-strategy.ec.europa.eu/en/news/eu-launches-ai-gigafactories-call-boost-europes-computing-capacity-and-unlock-more-eu30-billion

People's Republic of China (PRC) | China

Storm-1175 actor deploys new StormEncryptor ransomware after N-central vulnerability exploitation

The emergence of a new bespoke ransomware variant dubbed StormEncryptor signals a strategic evolution by the China-nexus, financially motivated threat cluster tracked as Storm-1175. Historically operating as an affiliate within the Medusa ransomware cartel, Storm-1175's pivot to proprietary malware underscores a broader trend wherein sophisticated extortion groups are abandoning established ransomware-as-a-service (RaaS) models in favor of custom tooling to maximize operational autonomy and evade signature-based detection. This resurgence—the group's first documented activity since April 2026—capitalizes on the systemic vulnerabilities inherent in centralized IT administration, specifically targeting CVE-2026-18577, a critical authentication-bypass vulnerability within N-able's N-central remote monitoring and management (RMM) platform. By exploiting this flaw for initial access, the attackers effectively weaponize the very infrastructure designed to manage enterprise environments. Upon breaching the perimeter, Storm-1175 demonstrates alarming operational velocity, rapidly transitioning from initial compromise to data exfiltration and encryption. Their post-exploitation playbook relies heavily on living-off-the-land techniques and dual-use software; the actors utilize AnyDesk and SimpleHelp for persistent remote access, alongside Advanced IP Scanner and Mimikatz for network reconnaissance and credential harvesting.

The final payload, StormEncryptor, is a C++ compiled ransomware that appends the ".encrypted" extension to compromised files and executes a double-extortion scheme, issuing a stringent three-day deadline to prevent data leakage. Microsoft Threat Intelligence emphasizes the highly compressed dwell time characterizing these intrusions, leaving defenders with a minimal window for disruption. For security practitioners and risk managers, this campaign reinforces the critical threat of RMM solutions acting as high-value choke points for supply chain attacks. The immediate imperative for organizations requires aggressively auditing internet-facing RMM exposure, enforcing strict multi-factor authentication controls, and swiftly applying N-able's hotfix to mitigate CVE-2026-18577 to defend against this accelerated time-to-encryption threat model.

Read more: https://www.scworld.com/brief/storm-1175-actor-deploys-new-stormencryptor-ransomware-after-n-central-vulnerability-exploitation

Jewelbug: APT Group Runs Espionage and Crypto Fraud Operations Side by Side

In a striking example of the increasingly blurred lines between state-sponsored espionage and financially motivated cybercrime, a newly identified China-based hackers-for-hire group tracked as Jewelbug has been observed conducting sweeping intelligence-gathering operations alongside a parallel cryptocurrency fraud enterprise. Operating within a threat landscape where mercenary syndicates frequently moonlight for profit, Jewelbug represents a highly sophisticated convergence of objectives. The group actively targets government, military, and telecom infrastructure across the Middle East, Southeast Asia, and South Asia, while simultaneously running an industrial-scale search-engine-optimization (SEO) poisoning pipeline aimed at Chinese-speaking cryptocurrency users. Both missions are orchestrated from a centralized command-and-control framework dubbed XG-Web, a browser-centric remote-access and information-stealing platform built on React and Node.js.

The operators rely on a versatile, multi-platform toolkit that includes the Antino Windows backdoor, a Linux and router implant known as ClientKing, and a highly invasive "PDF Viewer" malicious extension for Chrome and Firefox that harvests credentials, intercepts web traffic, and bypasses browser sandboxing. In their most significant espionage campaign to date, Jewelbug operators compromised a shared web-hosting platform utilized by a Middle Eastern state telecommunications provider, deploying a single malicious script that successfully established a watering hole across more than 15 government webmail tenants simultaneously. This operation, combined with their broader

targeting, yielded over one million implant check-ins and 580,000 stolen browser cookies in under three months. For defenders and national security stakeholders, Jewelbug's dual-mandate operations underscore a critical shift in adversary behavior, where the operational infrastructure and advanced tradecraft traditionally reserved for geopolitical intelligence collection are actively dual-purposed for illicit financial gain. Mitigating this hybridized threat requires organizations to adopt stringent identity verification, robust browser security policies, and continuous monitoring of third-party hosting environments to prevent sprawling supply-chain compromises from facilitating both catastrophic data exfiltration and targeted commodity fraud.

Read more: https://www.security.com/threat-intelligence/jewelbug-crypto-fraud-espionage

Chinese authority penalizes netizens for fabricating, spreading flood-related rumors

Chinese cybersecurity authorities operating under the Ministry of Public Security have initiated a targeted crackdown on individuals utilizing artificial intelligence to fabricate and disseminate disinformation during natural disasters. Amidst a global surge in generative AI-driven information operations, the manipulation of public perception during crises has emerged as a critical vector for societal disruption, forcing state entities to aggressively police synthetic media to preserve public trust and maintain operational continuity. Authorities recently disclosed 15 enforcement cases detailing how domestic netizens leveraged AI tools to engineer sophisticated rumors regarding recent severe flooding and typhoon impacts. Exploitation activities included doctoring official government safety advisories, generating synthetic notices of school and business closures distributed via WeChat, and circulating deceptively edited or AI-generated video content on platforms like Douyin to maximize virality.

Other incidents involved the wholesale invention of mass-casualty flash flood events, falsely claiming over 100 fatalities and failed rescue efforts to artificially inflate online engagement. Operating primarily within regions such as Guangdong Province, these operations actively polluted the localized information environment, which authorities reported directly impacted real-world disaster relief and crisis management efforts. All identified perpetrators faced administrative penalties under Chinese law for disrupting social order. This enforcement action underscores the escalating threat that democratized AI tools pose to crisis communications and national resilience. For risk managers and policy stakeholders, the rapid weaponization of synthetic media during emergencies highlights the urgent need for robust authentication protocols for official communications, as well as agile regulatory frameworks capable of neutralizing AI-enabled disinformation campaigns before they compromise localized operational security and public safety.

Read more: https://www.globaltimes.cn/page/202608/1367867.shtml

DeepSeek Plans 1GW AI Data Center in Inner Mongolia

In a move that underscores the intensifying global arms race for artificial intelligence compute sovereignty, Chinese AI developer DeepSeek is reportedly planning to construct a massive one-gigawatt (GW) data center in Ulanqab, Inner Mongolia, targeting partial operational capability by late 2027 or early 2028. This development arrives amidst a highly fractured geopolitical landscape where the United States and China are aggressively competing to secure the infrastructure required to train and deploy frontier large language models (LLMs). DeepSeek, which gained significant prominence in early 2025 by releasing highly efficient, open-source models that rivaled leading Western counterparts with significantly fewer computing resources, is driving this expansion as part of a broader push to scale its capabilities. Crucially, the company is also reportedly developing its own proprietary AI chips to power future model generations, directly attempting to bypass Western semiconductor monopolies and export controls.

Operationally, the 1GW facility aligns with Beijing's strategic "Eastern Data, Western Compute" initiative, a national blueprint designed to balance domestic computing power by leveraging the abundant land and energy resources of western regions like Inner Mongolia. By vertically integrating chip design with localized, hyperscale data center infrastructure, DeepSeek is working to achieve end-to-end technology self-sufficiency. For defenders and national security strategists, this massive mobilization of infrastructure signals that export controls on advanced compute hardware may only

temporarily delay, rather than permanently stall, the proliferation of state-of-the-art AI capabilities in rival nations. The strategic implications are profound: as Chinese entities establish independent, massive-scale compute architectures resilient to Western supply chain chokepoints, the global cybersecurity landscape must prepare for a future where advanced, domestically fueled AI engines accelerate both offensive capabilities and information operations from autonomous technological ecosystems.

Read more: https://datacenters.economictimes.indiatimes.com/news/construction-site-development/deepseek-plans-1gw-ai-data-center-in-inner-mongolia/132817758

Republic of China (ROC) | Taiwan

Taiwan says it was targeted in AI-driven hacking campaign in July

In a significant shift in U.S. cyber strategy, the White House has authorized a new National Security Presidential Memorandum (NSPM) that integrates private-sector companies into federal cyber operations against foreign transnational criminal organizations (TCOs). Amidst an escalating threat landscape defined by rampant ransomware attacks and sophisticated financial fraud originating from foreign jurisdictions, this policy aims to weaponize commercial cyber capabilities to defend American interests. Under the framework overseen by the Department of Homeland Security (DHS) and the Department of Justice vetted private entities are now empowered to enter intelligence-sharing agreements and propose direct cyber operations against designated criminal targets. Crucially, the directive authorizes both "cyber surveillance operations" and "cyber effects operations". The latter explicitly permits the manipulation, disruption, denial, degradation, or destruction of information systems, networks, and virtual or physical infrastructure controlled by these global crime syndicates. To participate in these government-directed operations, private companies are required to maintain a bond or escrow of at least $1 million, ensuring a baseline of accountability.

This unprecedented fusion of federal law enforcement authority and private-sector innovation represents a drastic escalation in how the U.S. combats transnational cybercrime. While leveraging industry agility may accelerate the disruption of ransomware cartels and financial fraud rings, it fundamentally alters the national security risk calculus. Authorizing commercial entities to execute destructive operations under federal oversight introduces profound implications for international stability, raising acute concerns over collateral damage, unintended escalation, and the legal complexities of state-sponsored operations in foreign jurisdictions. For security leaders and policy stakeholders, this development signals a new era of hybrid cyber warfare where the boundary between corporate cyber defence and offensive state action is increasingly dissolved.

Read more: https://www.straitstimes.com/asia/east-asia/taiwan-says-it-was-targeted-last-month-in-ai-driven-hacking-campaign

The Republic of Poland | Rzeczpospolita Polska

Follow-Up Report of the December 2025 Energy Sector Incident

In a critical escalation of operational technology (OT) threats, Poland's national cybersecurity team (CERT Polska) has disclosed new findings detailing a purely destructive cyberattack against a Polish combined heat and power (CHP) plant, executed via a novel private Access Point Name (APN) exploitation vector. Set against the backdrop of heightened geopolitical tensions and an increasingly contested threat landscape where sophisticated actors actively probe critical infrastructure, this disclosure highlights severe vulnerabilities in how industrial environments integrate modern cellular communications. The incident, which occurred in parallel with previously reported attacks on 30 renewable energy installations and a larger facility on December 29, 2025, marks the first observed instance of a purely destructive cyber campaign targeting Poland's energy sector. During this coordinated strike on a smaller CHP plant serving 50,000 residents, threat actors successfully shut down a steam turbine and a water treatment system, abruptly interrupting the facility's cogeneration of electricity and heat.

While rapid intervention by plant operators prevented actual heat supply outages for consumers, subsequent forensic investigations revealed that attackers breached the OT network by exploiting a widespread private APN misconfiguration. This architectural flaw permitted arbitrary, lateral communication between devices within the private cellular network, granting the attackers a direct, previously undocumented pathway to manipulate industrial control systems. CERT Polska warned that this inherently insecure configuration is highly prevalent globally among organizations utilizing similar APN solutions. Ultimately, this development exposes a severe systemic risk in modern industrial network design,

demonstrating that adversaries are aggressively pioneering novel access vectors to bypass traditional IT security perimeters. For defenders, risk managers, and critical infrastructure operators worldwide, this incident underscores the urgent necessity of auditing private APN architectures, validating cellular-to-OT boundaries, and enforcing strict network segmentation to safeguard vital public utilities against destructive sabotage.

Read more: https://cert.pl/en/posts/2026/08/incident-follow-up-report-energy-sector-2025/

Russian Federation | Российская Федерация, Rossiyskaya Federatsiya

CaptiveCrunch: Midnight Blizzard targets travelers worldwide for malware delivery and credential theft

A newly identified cyberespionage campaign dubbed "CaptiveCrunch" has exposed a sophisticated operation by Storm-2945 a sub-cluster of the Russian SVR-linked Midnight Blizzard (APT29) targeting global corporate travelers through compromised hospitality and captive portal networks. Operating since May 2026, this campaign highlights a critical vulnerability in shared venue infrastructure, demonstrating how state-sponsored threat actors are increasingly weaponizing public Wi-Fi networks at hotels and conference centers to circumvent enterprise identity perimeters and conduct targeted intelligence gathering. The attackers execute adversary-in-the-middle (AiTM) operations by intercepting automated browser connectivity checks via DNS and HTTP traffic manipulation. This redirects victims to AI-augmented "ClickFix" social engineering pages that masquerade as manual driver repairs, Google verification failures, or urgent operating system updates. Victims who execute the prompts are infected with "CornFlake," a fully featured, Golang-based remote access trojan (RAT) that establishes persistent command-and-control communications using an encrypted custom JSON protocol secured by an Elliptic Curve Diffie-Hellman (ECDH) P-256 key exchange.

Masquerading as a legitimate svchost32 process, CornFlake deploys a comprehensive surveillance toolkit capable of WASAPI-based audio recording, webcam capture, extensive host enumeration, and advanced credential theft notably utilizing ChromeKatz-derived modules to bypass Chrome's App-Bound Encryption (ABE). Concurrently, the threat group leverages "ChocoShell," an in-memory PowerShell infostealer engineered to rapidly harvest Microsoft 365 Single Sign-On (SSO) tokens and session cookies. The attackers are also actively exploiting Entra ID device code authentication flows to quietly register malicious devices. For security decision-makers, the CaptiveCrunch campaign underscores the urgent need to harden traveler security protocols and enforce strict, phishing-resistant authentication frameworks. As elite state actors increasingly exploit the intersection of physical mobility and insecure peripheral networks, defending against these highly evasive, token-stealing operations has become a paramount priority for global corporate cyber resilience.

Read more: https://www.microsoft.com/en-us/security/blog/2026/07/31/captivecrunch-midnight-blizzard-targets-travelers-worldwide-for-malware-delivery-and-credential-theft/

Middle East | West Asia

UAE thwarts cyber-attacks on aviation, energy and education sectors

In a testament to the escalating cyber threat landscape in the Middle East, the United Arab Emirates (UAE) Cybersecurity Council recently announced the successful disruption of a coordinated, multi-vector campaign targeting the nation's aviation, energy, and education sectors. Amidst a notable regional spike in sophisticated threat activity including AI-driven malware operations against the UAE financial sector in July and disruptive ransomware deployments earlier in 2026 this incident highlights the persistent targeting of critical infrastructure by advanced threat actors. The thwarted intrusion attempts employed a combination of tactics, heavily relying on targeted phishing campaigns designed to exploit end-users as initial access vectors into sensitive environments. Following this attempted access, the attackers engaged in systematic efforts to compromise digital infrastructure, harvest operational accounts, and infiltrate critical data environments.

However, proactive national cyber defense systems successfully detected the anomalies early in the attack lifecycle. Incident response teams tracked the adversaries' network paths, isolated associated indicators of compromise (IoCs), and implemented rapid technical countermeasures before the threat

actors could execute their primary objectives or disrupt vital public services. While the council has not formally attributed the campaign to a specific nation-state or transnational criminal syndicate, the simultaneous targeting of energy and aviation infrastructure aligns with intelligence-gathering or pre-positioning objectives typical of advanced persistent threats (APTs). This incident underscores a broader strategic shift toward fortified national cyber resilience, heavily supported by the UAE's recent launch of its AI-powered "Cyber Factory" initiative aimed at accelerating proactive defense capabilities. For enterprise defenders and national security stakeholders, these thwarted attacks reinforce the critical necessity of deploying continuous, intelligence-driven monitoring, robust identity management, and defense-in-depth architectures to preempt complex intrusion attempts before they can precipitate systemic operational failure or threaten international stability.

Read more: https://thearabweekly.com/uae-thwarts-cyber-attacks-aviation-energy-and-education-sectors

Iran-Linked Hackers Disrupt U.S. Water Systems via PLCs

In a stark escalation of state-aligned cyber aggression targeting critical infrastructure, Iran-linked threat actors have successfully disrupted physical operations at more than 30 U.S. water utilities across at least seven states by exploiting internet-exposed programmable logic controllers (PLCs). This campaign, primarily attributed to the Islamic Revolutionary Guard Corps (IRGC)-linked persona CyberAv3ngers alongside claims by the MOIS-associated hacktivist brand Handala, signals a dangerous shift from nuisance defacement to the direct manipulation of safety-instrumented systems. Taking advantage of the decentralized and often under-resourced nature of U.S. municipal water management, the attackers capitalized on widespread structural vulnerabilities, leveraging direct internet exposure and default credentials to gain initial access. The intrusion set targeted a broader array of equipment than previously observed, extending beyond Unitronics to include Rockwell Automation MicroLogix and ControlLogix systems, actively exploiting protocol exposures across EtherNet/IP, Modbus, S7comm, and SSH, as well as a known authentication-bypass vulnerability (CVE-2021-22681). Once inside, the operators utilized legitimate engineering software to alter PLC logic, disable safety shutdown behaviours, spoof HMI reporting to blind operators, and lock out legitimate staff by changing IP addresses and passwords; they further established persistence by deploying Dropbear SSH on connected cellular modems. This OT-level interference yielded tangible physical consequences, including localized pressure loss, boil-water advisories, and flooding. For operational technology defenders and national security policymakers, this incident emphatically underscores the precarious state of domestic cyber-physical resilience. It highlights the urgent imperative to eliminate direct internet exposure of control systems, enforce robust IT/OT segmentation, and mandate hardware-level protections such as switching PLCs to "RUN" mode to block remote alterations lest adversarial actors continue to transform trivial IT exposures into systemic kinetic threats against critical public utilities.

Read more: https://www.secureblink.com/threat-feeds/iran-linked-hackers-disrupt-u-s-water-systems-via-pl-cs

The French Republic | République française

France probes unprecedented cyberattack after tax data of 678,000 users stolen

In a severe breach of national data security, France's Directorate-General for Public Finances (DGFiP) has suffered a sophisticated cyberattack, resulting in the exfiltration of highly sensitive tax and personal data belonging to approximately 678,000 citizens and businesses. Occurring against a backdrop of escalating public sector IT compromises and a documented surge in physical "wrench attacks" violent robberies specifically targeting high-net-worth individuals and cryptocurrency holders across France this incident underscores the critical, real-world risks associated with centralized government data repositories. The intrusion, claimed by a threat actor operating under the alias "ZeroBytes," took place between June and July 2026. According to the French Finance Ministry, the attackers gained unauthorized access to internal DGFiP systems by leveraging stolen login credentials belonging to an employee and an authorized third party, subsequently executing a multi-factor authentication (MFA) bypass. Once inside, the threat actors reportedly abused an internal search tool to systematically extract comprehensive taxpayer profiles.

The compromised dataset, now reportedly advertised on dark web marketplaces, includes names, home addresses, dates of birth, phone numbers, email addresses, tax identifiers, company SIREN numbers, and precise financial metrics such as reference tax incomes and property details. While DGFiP confirmed that direct account passwords for the public-facing tax portal (impots.gouv.fr) were not compromised and that the illicit access was terminated upon detection, the exposed information provides a lucrative blueprint for financially motivated threat actors. The granularity of the data which precisely identifies tens of thousands of top-tier income earners drastically elevates the threat of highly tailored spear-phishing, systemic identity fraud, extortion, and targeted home invasions. For cybersecurity leaders and policymakers, this breach exposes the fragility of traditional identity perimeters reliant on standard MFA and highlights the urgent necessity for zero-trust architectures, rigorous behavioral anomaly detection, and robust data loss prevention (DLP) controls to secure critical state infrastructure against credential-based data extraction.

Read more: https://www.rfi.fr/en/france/20260815-france-probes-unprecedented-cyberattack-after-tax-data-of-678-000-users-stolen

Malware & Vulnerabilities

Multi-Functional Linux Botnet "Evooo1Bot"

FortiGuard Labs has identified a sophisticated, multi-functional Linux botnet dubbed "Evooo1Bot," expanding upon the notorious Mirai codebase to transform compromised edge devices into persistent proxies and attack nodes. Emerging within a threat landscape where poorly secured internet-facing infrastructure is increasingly weaponized for distributed operations, this discovery highlights the evolving complexity of router and IoT-centric malware. Active since at least July 2026, Evooo1Bot is compromising targets globally by exploiting a diverse array of known remote code execution and command injection vulnerabilities, including CVE-2024-29269 (Telesquare), CVE-2025-10123 (D-Link), and older flaws affecting Atlassian, Hikvision, and TP-Link devices. Once downloaded via a shell script loader (wget.sh), the heavily obfuscated, multi-layered binary runs extensive evasion checks scanning for analysis tools, sandboxes, and virtualized environments before establishing encrypted command-and-control (C2) communications over TCP port 443. Beyond incorporating 16 distinct DDoS attack vectors (such as UDP, TCP SYN/ACK, and OVH-bypass floods), Evooo1Bot features highly advanced operational modules. It deploys a built-in SOCKS5 proxy relay, allowing threat actors to route malicious traffic or build anonymous proxy networks through infected hosts.

It also utilizes a custom SSH brute-force scanner with a 150+ credential dictionary targeting enterprise service accounts (e.g., jenkins, postgres), uniquely integrating dual-stage honeypot detection to evade systems like Cowrie or Kippo. The botnet establishes aggressive persistence across the host using systemd, SysV init, cron jobs, and rc.local, while manipulating the oom_score_adj and /dev/watchdog to prevent process termination. For security practitioners, Evooo1Bot underscores the critical necessity of rapid patch management for edge and IoT devices. The botnet's blend of automated exploitation, sophisticated honeypot evasion, and traffic relay capabilities signals a dangerous shift toward highly resilient, multi-purpose edge infections that can seamlessly pivot from disruptive DDoS campaigns to stealthy enterprise network intrusions.

Read more: https://www.fortinet.com/blog/threat-research/multi-functional-linux-botnet-evooo1bot

Amazon identifies North Korean hacker group behind open-source supply chain attacks

Amazon Threat Intelligence has uncovered new evidence linking a sophisticated North Korean state-sponsored threat actor tracked interchangeably as SAPPHIRE SLEET, BlueNoroff, and STARDUST CHOLLIMA to a series of high-impact software supply chain attacks compromising major open-source Node Package Manager (NPM) libraries. This revelation arrives as the cybersecurity community continues to grapple with the fallout from the XZ Utils backdoor, underscoring a systemic vulnerability where patient, state-linked actors exploit the trust and limited resources of volunteer open-source maintainers to infiltrate critical global infrastructure. By injecting malicious code into foundational software building blocks, attackers can achieve downstream compromise across countless organizations simultaneously. Amazon's analysis connects several previously disparate incidents, revealing that this single Democratic People's Republic of Korea (DPRK) nexus actor is responsible

for compromising the typo-crypto package in March 2025, the debug and chalk packages in September 2025, and most recently, the highly ubiquitous axios library which averages over 100 million weekly downloads in March 2026.

Operationally, the threat actor consistently relied on social engineering tactics targeting trusted maintainers to gain access before pushing updates laced with malicious payloads. Technically, the campaigns exhibited shared tactics, techniques, and procedures (TTPs), specifically the deployment of trojanized NPM packages, the malicious utilization of post-install hooks to trigger automatic execution upon installation, and identifiable code reuse alongside shared command-and-control (C2) infrastructure. Furthermore, Amazon noted early signs of threat actors leveraging generative AI to alter the obfuscation and structure of malicious packages, as well as probing AI-assisted coding systems. For risk managers and defenders, these findings highlight an escalating, industrialized approach to supply chain compromise by North Korea, emphasizing the urgent need for stringent dependency auditing, enhanced maintainer security protocols, and advanced detection mechanisms capable of identifying increasingly AI-mutated malware within the global open-source ecosystem.

Read more: https://aws.amazon.com/blogs/security/amazon-identifies-north-korean-hacker-group-behind-open-source-supply-chain-attacks/

I'll Just Call You: Agent-to-Agent Privilege Boundary Failures in CI/CD on Google's ADK Repository

In a groundbreaking demonstration of vulnerability within multi-agent systems, Pillar Security researchers have identified the first practical instance of agent-to-agent privilege escalation in a production CI/CD environment, discovering a critical flaw in Google's adk-python repository (the Agent Development Kit for Python). As organizations increasingly integrate autonomous agents into software supply chains, the failure to establish robust privilege boundaries between AI agents introduces novel, previously un-modeled attack vectors. The vulnerability stemmed from the interplay between two classes of automated agents: a low-privileged, public-facing triage agent embedded in GitHub workflows (activated by pull requests) and high-privileged agents (gemini-invoke and gemini-review) reserved exclusively for repository maintainers. Researchers discovered that the triage agent, operating via a highly privileged "Collaborator" user account (rather than a restricted bot account), could be manipulated through prompt injection within a pull request to output specific trigger commands, such as @gemini-cli. This action successfully bypassed access controls, tricking the downstream dispatcher workflow into executing the high-privileged agent.

By reframing the malicious payload as standard compliance checks using the Pillar Security CFS Framework, the researchers forced the agent to leak its Model Context Protocol (MCP) tool capabilities and expose a Remote Code Execution (RCE) pathway within the GitHub runner. This access allowed extraction of the GITHUB_TOKEN, granting the attacker the ability to edit existing issue comments, impersonate maintainers, and manipulate the pull request approval lifecycle including dismissing reviews or falsely approving changes. While this specific exploit did not immediately yield direct code insertion, it provided all necessary conditions for a sophisticated supply chain compromise via high-impact social engineering. This incident underscores an urgent imperative for security practitioners: traditional threat models must rapidly evolve to account for inter-agent interactions, strict capability scoping, and the blast radius of chained agentic workflows in CI/CD pipelines.

Read more: https://www.pillar.security/blog/ill-just-call-you-agent-to-agent-privilege-boundary-failures-in-ci-cd-on-googles-adk-repository

State Sponsored Hackers Use Fake Job Offers to Deliver New Zero-Day Exploit

State-sponsored hackers affiliated with North Korea's Lazarus Group have weaponized fake job offers to deliver a novel Windows zero-day vulnerability (CVE-2026-68820), targeting the global defence, aerospace, and aviation sectors. Against a backdrop of intensifying geopolitical competition and the relentless pursuit of defence intelligence, this campaign a resurgence of the long-running "Operation Dream Job" demonstrates how advanced persistent threats (APTs) continually refine social engineering to bypass modern perimeter defences. The attack vector relies on malicious recruiters approaching professionals on platforms like LinkedIn with convincing PDF job descriptions. Researchers identified two primary infection chains: one utilizing DLL sideloading to execute the MISTPEN downloader, and another employing a trojanized PDF viewer ("SecurityPDF") to drop the previously undocumented "Troy" backdoor directly

into memory. Once initial access is achieved, the attackers deploy the zero-day exploit, which targets a core Windows component responsible for network connections (AFD.sys), enabling privilege escalation to system-level control.

This elevated access allows the deployment of a highly capable rootkit (v3.1) designed to evade Endpoint Detection and Response (EDR) solutions by disabling logging systems and tampering with Windows trust verification features. Notably, the threat actors eschewed dedicated command-and-control servers, opting instead to hijack vulnerable, internet-facing Roundcube webmail installations and content management systems using a new PHP webshell called RelayShell. This technique, combined with the use of commercial VPNs, effectively obfuscates malicious communications within legitimate traffic. Following coordinated disclosure, Microsoft patched CVE-2026-68820 on August 11, 2026. This development underscores a critical imperative for enterprise risk management: defenders must prioritize aggressive endpoint behavioural monitoring, rapid patch deployment, and rigorous posture validation of public-facing infrastructure to counter state-aligned adversaries who increasingly blend zero-day exploitation with highly personalized, human-centric attack vectors.

Read more: https://blog.checkpoint.com/research/state-sponsored-hackers-use-fake-job-offers-to-deliver-new-zero-day-exploit/

About the Author

Govind Nelika is a Researcher, Web Manager, and Outreach Coordinator at the Centre for Land Warfare Studies (CLAWS), working on national security issues at the intersection of technology, cybersecurity, and geopolitics. His research focuses on hybrid warfare, digital influence operations, semiconductor geopolitics, AI-enabled conflict, and cyber governance, with publications covering topics such as U.S.–China tech rivalry, the Quad's cyber dynamics, and emerging risks in AI and supply chains. He previously worked at Pondicherry University under the UGC-SAP (DRS II) programme in the Department of Politics & International Studies, progressing from Project Fellow to Project Associate. He holds a degree in Political Science and a Data Science certification from IBM. Earlier in his career, he gained research and digital management experience with the Regional Centre of Expertise, Trivandrum (affiliated with the United Nations University), and the Bureau of Police Research & Development (BPRD), Ministry of Home Affairs where he conducted research on cybercrime trends in India. He was awarded the Chief of Army Staff (COAS) Commendation Card on Army Day 2025 for his contributions to CLAWS