CYBERSECURITY AT MIT SLOAN

Working Paper Series CAMS25.1114

Rethinking the Cybersecurity Arms Race

When 80% of Ransomware Attacks are AI-Driven

Michael Siegel
Principal Research Scientist
Director, *Cybersecurity at MIT Sloan*
Massachusetts Institute of Technology
msiegel@mit.edu

Sander Zeijlemaker
Research Affiliate, *Cybersecurity at MIT Sloan*
Massachusetts Institute of Technology
szeijl@mit.edu

Vidit Baxi
Co-Founder & CISO
Safe Security

Sharavanan Raajah
Threat Researcher
Safe Security

April 10, 2025

Executive summary

Cyber risk and resilience management have always been a race, with technological advancements continuously reshaping the battlefield. Today, artificial intelligence (AI) is revolutionizing cyber threats, making them more scalable, adaptive, and autonomous. AI-enhanced phishing, deep-fake deception, and polymorphic malware are just the beginning. Our data-driven research of 2800 ransomware attacks reveals an even more alarming reality. Adversarial AI is now automating entire attack sequences, executing with minimal human intervention, and dynamically adapting to exploit weaknesses in real time.

Threat actors rapidly adopt AI-driven techniques to enhance their attacks' sophistication, automation, and efficiency in targeting organizations. Malicious actors leverage AI for sophisticated phishing campaigns, advanced social engineering, and the evasion of traditional security mechanisms. In contrast, the adoption of AI-driven defensive measures cannot keep up due to challenges accounting for resource limitations, regulatory constraints, and the complexity of implementation. This discrepancy favours the adversaries who can exploit AI's capabilities with fewer restrictions. Bridging this gap requires a strategic approach to AI adoption in cybersecurity, ensuring that defensive mechanisms evolve in parallel with emerging threats. In 2023 – 2024, AI drives 80% of the ransomware attacks.

To counter this new reality, organizations must move beyond reactive security measures. AI-powered cybersecurity tools alone are insufficient; a proactive, multi-layered approach—integrating human oversight, governance frameworks, AI-driven simulations, and real-time intelligence sharing—is essential. Executives must act decisively to close the widening gap between attackers and defenders. By embracing AI in cyber risk management, organizations can enhance resilience, reduce operational burdens, and maintain a proactive stance. A structured, AI-enabled security roadmap will be crucial to staying ahead and safeguarding critical assets in an increasingly complex threat landscape.

Table of Contents

Defenders must stay ahead of an AI-based cyber threat landscape

Cyber risk management and resilience has always been a race—one where technological innovation dictates the pace. For decades, advancements in computing have reshaped the battlefield, forcing organizations to evolve in response to ever-more sophisticated threats[1]. Today, artificial intelligence (AI) is not just shifting the balance—it is redefining the very nature of cyber threats.

Recent reports highlight AI-enhanced phishing, deep-fake-driven deception, and even in a lab setting polymorphic malware designed to outmaneuver traditional defenses. But recent data-driven research of 2800 ransomware attacks reveals an even more alarming reality: AI is not merely improving cyberattack techniques—it is transforming cyber threats into something far more scalable, adaptive, and autonomous.

Adversarial AI is now automating[2] entire attack sequences, enabling seamless execution across the kill chain with minimal human intervention. More concerning, AI-driven agentic systems are emerging—threat actors that not only react but autonomously act, adapting their strategies in real time to exploit specific weaknesses in an organization's defenses. These autonomous, aggressive agents[3] mark a paradigm shift: cyber threats that think, learn, and execute with unprecedented speed and precision[4].

To confront this new era, organizations must move beyond reactive security strategies[5]. AI-powered cybersecurity tools alone will not suffice. A proactive, multi-layered approach—integrating human oversight, governance frameworks, AI-driven threat simulations, and real-time intelligence sharing—is critical. The ability to anticipate and disrupt AI-driven attacks before they materialize will separate leaders from those left vulnerable in the wake of machine-speed threats.

This work challenges conventional assumptions about AI-driven cyber threats by grounding the conversation in hard data and real-world evidence. We present an unfiltered view of today's AI-powered attack landscape and offer strategic guidance to close the widening gap between attackers and defenders. The cyber arms race has entered a new phase—executives

[1] Davies, V. (2021, October 4). The history of cybersecurity. Cyber Magazine. https://cybermagazine.com/cyber-security/history-cybersecurity

[2] Zhang, C., Pal, R., Nicholson, C., & Siegel, M. (2024, December). (Gen) AI Versus (Gen) AI in Industrial Control Cybersecurity. In 2024 Winter Simulation Conference (WSC) (pp. 2739-2750). IEEE. https://dl.acm.org/doi/10.5555/3712729.3712956.

[3] The Grungq. (2017). The Triple A Threat: Aggressive autonomous agents. BlackHat. https://www.blackhat.com/docs/webcast/12142017-the-triple-a-threat.pdf

[4] Schröer, S.L., Apruzzese, G., Human, S., Laskov, P., Anderson, H. S, Bernroider, E. W. N., Fass, A., Nassi, B., Rimmer, V., Roli, F., Salam, S., Shen, A, Sunyaev, A, Wadhwa-Brown, T., Wagner, I., Wang, G. (2025) SoK: On the Offensive Potential of AI. IEEE Conference on Secure and Trustworthy Machine Learning.

[5] Zhang, C., Pal, R., Nicholson, C., & Siegel, M. (2024, December). (Gen) AI Versus (Gen) AI in Industrial Control Cybersecurity. In 2024 Winter Simulation Conference (WSC) (pp. 2739-2750). IEEE. https://dl.acm.org/doi/10.5555/3712729.3712956.

must act now to stay ahead. We close this work with recommendations on how defenders can close this gap.

How artificial Intelligence can leverage the adversary

The rapid evolution of artificial intelligence is redefining the cyber threat landscape, arming adversaries with unprecedented capabilities that challenge conventional security frameworks. AI is no longer merely an augmentation tool—it is now a core driver of automation, deception, and efficiency in cyberattacks. The implications for enterprises, governments, and cybersecurity professionals are profound: organizations must recognize and address the increasingly sophisticated adversarial tactics enabled by AI. The AI-Powered Adversary is characterized by three different roles of AI's in cyberattacks today.

First, AI-Generated Cyber Threats where AI automates the creation of sophisticated malware, phishing campaigns, and deep-fake-driven social engineering. Large language models (LLMs) such as ChatGPT and Claude are being leveraged to generate malicious code and tailor phishing content. Malware strains like SugarGh0st RAT, AsyncRAT, and Win32/Wkysol illustrate how AI enhances attack methodologies [6][7][8]. Generative Adversarial Networks (GANs) create malware that mimics legitimate software, bypassing traditional security measures[9]. AI-generated deepfake attacks, such as synthetic voice fraud, have been used in high-profile corporate breaches where attackers impersonate executives to authorize fraudulent transactions.

Second, AI-Enhanced Cyber Kill Chain – AI is optimizing each phase of the cyber kill chain, making attacks more precise and resilient:

- Reconnaissance: AI-driven machine learning models scan vast networks, identifying vulnerabilities at scale. Automated scanning tools such as Shodan and Censys leverage AI to detect weaknesses in real time[10].
- Weaponization: AI enables the development of polymorphic malware like BlackMamba[11], which continuously mutates its code to evade signature-based

[6] OpenAI. (2024, October 9). An update on disrupting deceptive uses of AI. OpenAI. https://openai.com/global-affairs/an-update-on-disrupting-deceptive-uses-of-ai/

[7] Iascu, I. (2024, April 10). Malicious PowerShell script pushing malware looks AI-written. Bleeding Computer. https://www.bleepingcomputer.com/news/security/malicious-powershell-script-pushing-malware-looks-ai-written/

[8] Toulas, B. (2024, September 24). Hackers deploy AI-written malware in targeted attacks. Bleeping Computer. https://www.bleepingcomputer.com/news/security/hackers-deploy-ai-written-malware-in-targeted-attacks/

[9] Rigaki, M., & Garcia, S. (2018). Bringing a GAN to a Knife-Fight: Adapting Malware Communication to Avoid Detection, IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 2018, pp. 70-75, doi: 10.1109/SPW.2018.00019.

[10] Alabdulatif, A., & Thilakarathne, N. N. (2025). Hacking Exposed: Leveraging Google Dorks, Shodan, and Censys for Cyber Attacks and the Defense Against Them. Computers, 14(1), 24. https://www.mdpi.com/2073-431X/14/1/24

[11] Sims, J. (2023, July 31). BlackMamba: Using AI to generate polymorphic malware. Hyas. https://www.hyas.com/blog/blackmamba-using-ai-to-generate-polymorphic-malware

detection. The DeepLocker malware[12] framework uses AI to trigger attacks only under specific conditions, avoiding premature exposure.

- Delivery & Exploitation: AI analyzes behavioral patterns to optimize attack timing and vector selection. AI-driven social engineering tools tailor phishing emails based on an individual's digital footprint, significantly increasing the likelihood of success.
- Command & Control (C2): AI-powered botnets use natural language processing (NLP) to evade detection by mimicking human communication[13]. For example, adversarial AI chatbots have been used to engage in multi-turn conversations with cybersecurity analysts, misleading them while maintaining control over compromised systems.
- Actions on Objectives: AI automates data exfiltration, blending malicious traffic into normal network behavior. AI-enhanced rootkits adapt to monitoring techniques, ensuring prolonged persistence within a compromised system.

Finally, AI-Driven Attack Automation – AI integrates and automates multiple attack stages, enabling more persistent and scalable threat. Examples are Federated Learning & Decentralized Botnets, Automated Aggressive Agents (AAA)[14], and Generative AI for New Attack Forms. Federated Learning & Decentralized Botnets where attackers leverage these learning mechanisms to train botnets like Mozi without centralized control, making them highly resistant to takedown efforts[15][16][17]. AAAs where cyber attacks have agency. Some past examples of malware are WannaCry and NotPetya exemplify AI-powered self-propagating attacks that require minimal human intervention. More recent examples include EyeSpy[18], which automates reconnaissance, payload delivery, and evasion techniques. Generative AI for New Attack Forms where GAN-based tools like attackGAN[19], ISDGAN[20], and MalwareGAN[21] continuously probe security systems to generate novel attack techniques, effectively outpacing traditional security solutions.

[12] Ph, M., Stoecklin, J., Jang, J., & Kirat, D. (2018, August 8). DeepLocker: How AI can power a stealthy new breed of malware. Security Intelligence. https://securityintelligence.com/deeplocker-how-ai-can-power-a-stealthy-new-breed-of-malware/

[13] Yang, K.-C., & Menczer, F. (2023, July 30). Anatomy of an AI-powered malicious social botnet. arXiv Preprint. https://arxiv.org/abs/2307.16336

[14] The Grungq. (2017). The Triple A Threat: Aggressive autonomous agents. BlackHat. https://www.blackhat.com/docs/webcast/12142017-the-triple-a-threat.pdf

[15] Digital, N. (2023, January 16). Benefits of federated learning explained. OctaiPipe. Retrieved September 16, 2024, from https://octaipipe.ai/benefits-of-federated-learning-explained/

[16] Belova, K. (2023, December 15). What is federated learning: Key benefits, applications, and working principles explained. PixelPlex. https://pixelplex.io/blog/federated-learning-guide/

[17] Nelson, N. (2023, November 3). Somebody Just Killed the Mozi Botnet., Darkreading, https://www.darkreading.com/ics-ot-security/somebody-just-killed-mozi-botnet

[18] Hays. (2024). EyeSpy: Proof-of-concept. Hays. https://www.hyas.com/read-the-eyespy-proof-of-concept

[19] Zhao, S., Li, J., Wang, J., Zhang, Z., Zhu, L., & Zhang, Y. (2021). attackGAN: Adversarial attack against black-box IDS using generative adversarial networks. Procedia Computer Science, 187, 128–133. https://doi.org/10.1016/j.procs.2021.04.118

[20] Lin, Z., Shi, Y., & Xue, Z. (2022). IDSGAN: Generative adversarial networks for attack generation against intrusion detection. In Lecture Notes in Computer Science (pp. 79–91). Springer International Publishing. http://dx.doi.org/10.1007/978-3-031-05981-0_7

[21] Huang, Y., Huang, L., & Zhu, Q. (2022). Reinforcement learning for feedback-enabled cyber resilience. Annual Reviews in Control, 53, 273–295. https://doi.org/10.1016/j.arcontrol.2022.01.001

AI is fundamentally reshaping the cyber threat landscape by enhancing the speed, precision, and adaptability of cyberattacks[22]. From automated reconnaissance to AI-powered deception techniques, adversaries are leveraging machine learning to bypass security measures with unprecedented efficiency. As AI-driven threats become more sophisticated, organizations must acknowledge the evolving nature of these attacks and rethink their cybersecurity strategies.

Our learnings from AI-powered attacks in 2024

Since 2022, the rise of artificial intelligence (AI) has transformed various industries, including cybersecurity. While AI has the potential to strengthen security systems, it has also become a tool for cybercriminals, leading to increasingly sophisticated and harder-to-detect threats.

One of the most concerning developments is how AI is accelerating the spread of malware. Hackers now use AI-powered tools to automate the creation of malicious software, making it easier to evade traditional security measures. Additionally, AI enables cybercriminals to customize attacks for specific individuals or organizations, making them more convincing and successful.

Another growing concern is the role of AI in social engineering scams, especially phishing. AI-generated emails can mimic real messages so effectively that even cautious users may struggle to spot fraud. The emergence of deepfake technology further complicates the issue. By generating realistic fake audio and video, deepfakes allow scammers to impersonate people convincingly, increasing the likelihood of financial fraud and data breaches.

Recent statistics reflect the severity of the problem. In 2024, phishing scams and deepfake-related fraud linked to AI saw a staggering 60% increase. Businesses and security professionals are using AI-driven defenses to combat these evolving threats. Machine learning algorithms can detect suspicious patterns and anomalies in real-time, helping organizations respond before significant damage is done.

However, technology alone isn't enough to tackle AI-powered cybercrime. Collaboration between governments, businesses, and security experts is essential. Cybersecurity professionals can stay ahead of cybercriminals and minimize risks by sharing threat intelligence, developing countermeasures, and increasing awareness.

[22] Schröer, S.L., Apruzzese, G., Human, S., Laskov, P., Anderson, H. S, Bernroider, E. W. N., Fass, A., Nassi, B., Rimmer, V., Roli, F., Salam, S., Shen, A, Sunyaev, A, Wadhwa-Brown, T., Wagner, I., Wang, G. (2025) SoK: On the Offensive Potential of AI. IEEE Conference on Secure and Trustworthy Machine Learning.

While AI has immense potential to improve cybersecurity, it presents new challenges. The key to staying protected lies in proactive security measures, AI-driven defense systems, and a united front against emerging threats. We can navigate this rapidly evolving digital landscape with greater security and confidence by working together and staying informed.

Some examples of AI-powered attack vectors seen in 2023-24 are,

1. AI-powered Phishing
2. AI-powered Social Engineering
3. Voice Cloning
4. Deepfakes
5. AI-powered Malware
6. AI-powered Password Cracking
7. AI-powered CAPTCHA Bypass
8. AI-powered Exploitation

What attacks were analyzed?

Our recent analysis of over 2800 ransomware incidents has revealed an alarming trend: *AI plays an increasingly significant role in these attacks*. In 2024, 80.83% of recorded ransomware events were attributed to threat actors utilizing AI.

This surge in AI-driven ransomware attacks highlights the evolving threat landscape and the growing sophistication of cybercriminals. By leveraging AI, attackers can automate and enhance various stages of the ransomware attack lifecycle, from target selection and vulnerability identification to payload delivery and extortion.

The implications of this trend are substantial. We will likely see even more sophisticated and widespread ransomware attacks as AI advances. This poses a significant challenge for organizations of all sizes, as they must adapt their cybersecurity strategies to defend against this evolving threat.

2024 Key Statistics

The 2023-2024 period saw a dramatic surge in ransomware attacks, with 2,811 recorded incidents. A staggering 80.83% of these attacks, equating to 2,272 events, were directly attributed to AI-enabled Threat Actors. This highlights the growing sophistication and danger of AI-powered cybercrime.

Among the numerous AI-enabled ransomware groups, several emerged as particularly prolific. LockBit led the pack with 815 attacks, followed by RansomHub with 548. Akira and ALPHV/BlackCat were responsible for 314 and 155 attacks, respectively. Other prominent groups included BlackBasta (189 attacks), Cactus (93 attacks), DragonForce (84 attacks), and Funksec (74 attacks).

- Total recorded ransomware events: 2,811
- Ransomware events linked to AI-enabled Threat Actors: 2,272 (80.83%)
- Prominent AI-enabled Ransomware Groups:

#	Threat Actor Name	Events
1	LockBit[23]	815
2	RansomHub[24]	548
3	Akira[25]	314
4	ALPHV/BlackCat[26]	189
5	BlackBasta[27]	155
6	Cactus[28]	93
7	DragonForce[29]	84
8	Funksec[30]	74

[23] Cybersecurity and Infrastructure Security Agency. (n.d.). LockBit overview. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-165a
[24] Cybersecurity and Infrastructure Security Agency. (n.d.). RansomHub overview. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-242a
[25] AkiraDecryptor. (2025). Akira ransomware & AI-driven attacks: Poised to unleash AI-driven attacks in 2025, experts warn. https://akiradecryptor.org/akira-ransomware-poised-to-unleash-ai-driven-attacks-in-2025-experts-warn/
[26] Cybersecurity and Infrastructure Security Agency. (n.d.). ALPHV Blackcat overview. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-353a
[27] Cybersecurity and Infrastructure Security Agency. (n.d.). BlackBasta overview. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-131a
[28] Darktrace (2024, April 24). How Cactus Ransomware was detected and stopped. https://www.darktrace.com/blog/a-thorn-in-attackers-sides-how-darktrace-uncovered-a-cactus-ransomware-infection
[29] Halcyon. (n.d.). DragonForce ransomware attack threatens EvoEvents Limited. Halcyon. https://www.halcyon.ai/attacks/dragonforce-ransomware-attack-threatens-evoevents-limited: Incident Data used for analysis is up to February 2025
[30] Check Point Research. (2025). AI-powered ransomware: FunkSec—Alleged top ransomware group powered by AI. Check Point. https://research.checkpoint.com/2025/funksec-alleged-top-ransomware-group-powered-by-ai

AI-powered Ransomware Attack

Ransomware Trends

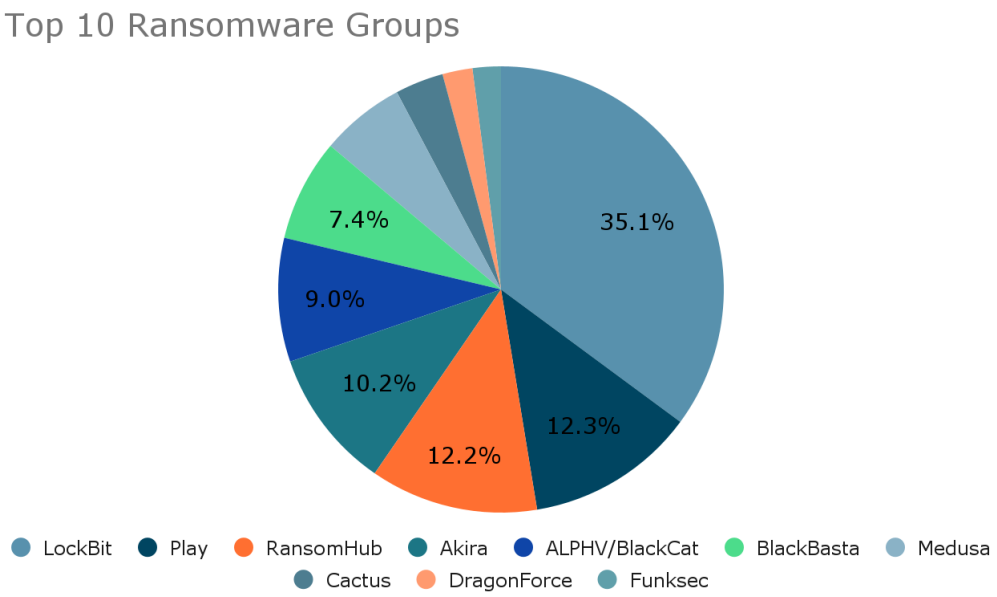

Image 1.1: Cybercriminal Groups Using AI in Ransomware Attacks 2023-2024: Top 10 and Their Distribution

The above chart from our analysis shows the distribution of ransomware attacks in 2023-24 by the top 10 cybercriminal groups known to be leveraging AI and spreading Ransomware malware, which accounts for almost 80% of the overall Ransomware attacks.

AI has made ransomware attacks faster, more efficient, and harder to detect. Specifically, conducting more convincing phishing attacks or developing/modifying a malware tool that can evade detection has been more efficient. Additionally, AI-generated deepfake voices or videos that impersonate executives help trick employees into installing malware. Furthermore, the rise of Ransomware-as-a-Service (RaaS) facilitated by AI has also made it easier for new threat actors to begin their journey. These benefits of AI have enabled adversaries to operate at scale, increasing the frequency and severity of Ransomware attacks.

As adversaries continue weaponizing AI, cybersecurity professionals must adopt equally advanced countermeasures, integrating AI-driven threat detection, predictive analytics, and automated incident response to stay ahead of the ever-evolving ransomware landscape.

Common Capabilities of AI-powered Ransomware

With its increasing sophistication and the integration of AI, ransomware employs advanced machine learning and data-driven techniques to enhance its capabilities, making it more efficient, adaptive, and challenging to counter.

Some of the core capabilities of AI-powered ransomware are as follows.

1. Targeted File Selection
    - AI analyzes the victim's system to identify and prioritize high-value files, such as sensitive data, intellectual property, or critical business documents, ensuring maximum impact.
    - It avoids encrypting unnecessary or decoy files, saving time and resources.
    - Example:
        - *The Clop Ransomware*[31] is a strong example of AI-powered ransomware that exhibits the Targeted File Selection feature. It is known for targeting high-value files and critical business documents, particularly in industries like finance and healthcare. It avoids encrypting non-essential files to ensure faster and more efficient attacks. Clop often avoids unnecessary files, such as system files or decoys, reducing its footprint while causing maximum disruption. The Clop Ransomware[32] exemplifies AI-driven ransomware with its Targeted File Selection capability. Focusing on high-value and critical business documents, especially finance and healthcare, Clop prioritizes efficiency and speed by avoiding non-essential files. This targeted approach minimizes Clop's footprint while maximizing disruption, ensuring that only the most impactful files are encrypted.

2. Evasion Tactics
    - AI-powered ransomware employs advanced evasion techniques, such as:
        - Polymorphism: Regularly altering its code structure to evade signature-based detection.
        - Behavioral Adaptation: Learning the behavior of endpoint detection systems and adjusting their actions to bypass them.
    - Identify and neutralize security software or sandboxes.
    - Example:

[31] BleepingComputer. (n.d.). *Clop overview*. https://www.bleepingcomputer.com/tag/clop-ransomware/
[32] Palo Alto Networks. (n.d.). *Clop technical analysis*. https://unit42.paloaltonetworks.com/clop-ransomware/

- *The Maze Ransomware*[33] and *LockBit 3.0*[34] are excellent examples of ransomware exhibiting AI-powered evasion tactics.

- Maze Ransomware[35] regularly modifies its code to avoid detection by traditional signature-based antivirus systems.
- In addition to modifying the code, LockBit 3.0 actively searches for and turns off antivirus, backup services, and endpoint security solutions.

3. Dynamic Ransom Demands
    - To set ransom amounts, AI analyzes victim data, such as financial status, industry, and past responses to cyber incidents.
    - Personalize ransom messages to pressure victims into compliance.
    - Example:
        - *REvil (Sodinokibi)*[36] is a prime example of AI-powered ransomware demonstrating Dynamic Ransom Demands. It tailors ransom demands based on the victim's financial data, industry, and organization size. It uses intelligence gathered from compromised systems to assess what the victim can afford and sets ransom demands accordingly. REvil[37] adjusts ransom amounts based on prior responses to ransomware attacks in the same industry.

4. Faster Spread and Infection
    - AI optimizes attack vectors by identifying the most vulnerable systems and efficiently propagating through networks via:
        - Exploiting weak credentials and misconfigurations.
        - Using predictive modeling to locate additional systems susceptible to compromise.
    - Example:
        - *Emotet*[38] and *Ryuk*[39] ransomware families are examples of ransomware that exhibit Faster Spread and Infection using AI-like features to optimize attack vectors. Emotet acts as a loader, delivering ransomware payloads like Ryuk or Conti[40] with advanced propagation mechanisms. It uses brute-force techniques to exploit weak

[33] BleepingComputer. (n.d.). *Maze ransomware analysis*. https://www.bleepingcomputer.com/news/security/maze-ransomware/
[34] SentinelOne. (n.d.). *LockBit 3.0 overview*. https://www.sentinelone.com/anthology/lockbit-3-0-lockbit-black/
[35] Unit 42. (n.d.). *Technical analysis*. Palo Alto Networks. https://unit42.paloaltonetworks.com/maze-ransomware-campaign/
[36] Cybersecurity and Infrastructure Security Agency. (n.d.). *REvil overview*. CISA. https://www.cisa.gov/stopransomware/revil-sodinokibi
[37] Wired. (n.d.). *Kaseya attack – Detailed report*. https://www.wired.com/story/kaseya-ransomware-attack-revil/
[38] United States Computer Emergency Readiness Team. (n.d.). *Emotet technical overview*. US-CERT. https://us-cert.cisa.gov/ncas/alerts/aa20-280a
[39] Symantec. (n.d.). *Emotet and Ryuk collaboration*. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/emotet-ryuk-conti-ransomware
[40] Symantec. (n.d.). *Emotet and Ryuk collaboration*. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/emotet-ryuk-conti-ransomware

credentials in Active Directory environments. It also scans for misconfigured systems and lateral movement opportunities within the network to spread faster.

5. Intelligent Reconnaissance
    - AI-powered ransomware conducts thorough reconnaissance, identifying backup systems, cloud services, and disaster recovery plans to disable them before launching encryption.
    - Example:
        - *Conti*[41] ransomware is a notable example of Intelligent Reconnaissance features. It performs extensive network scanning to locate backup systems, cloud services, and disaster recovery plans. It also disables or deletes shadow copies, backups, and recovery services to prevent victims from restoring their systems. It integrates tools like *Mimikatz* for gathering credentials and identifying privileged accounts to access critical systems.

6. Autonomous Decision-Making
    - AI models enable ransomware to act autonomously, making decisions such as:
        - When to deploy encryption for maximum disruption.
        - How to respond to security countermeasures or restore attempts.
    - Example:
        - *LockBit* 3.0[42] (LockBit Black) is a prime example of ransomware demonstrating Autonomous decision-making features. It uses AI models to determine the optimal time to launch encryption, often deploying during low-activity periods (nights, weekends, or holidays) for maximum disruption. It can autonomously respond to security tools by altering its behavior to bypass detection or block manual restore attempts. The ransomware turns off endpoint protection services and backups without direct human intervention. LockBit 3.0's[43] AI-driven logic identifies and prioritizes targets within a network, adjusting its encryption tactics based on system responses.

[41] Sophos. (2021, February 16). *Conti ransomware technical analysis*. https://news.sophos.com/en-us/2021/02/16/conti-ransomware-evasive-by-nature/

[42] Trend Micro. (n.d.). *LockBit analysis*. https://www.trendmicro.com/vinfo/us/security/news/ransomware-spotlight/ransomware-spotlight-lockbit

[43] Cybersecurity and Infrastructure Security Agency. (n.d.). *Ransomware-as-a-service*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-075a

Threat Actors Leveraging AI in Their Attack Cycle

As artificial intelligence (AI) continues to reshape the digital landscape, its capabilities are being exploited for innovation and cybercriminal activity. Integrating AI into cyberattacks has given rise to more sophisticated, evasive, and damaging threats, with various advanced persistent threat (APT) groups and cybercriminal organizations leveraging AI to enhance their attack strategies.

The table below outlines some of the most active AI-powered threat actors[44], detailing their tactics, techniques, and procedures (TTPs) in modern cyber operations. These groups employ AI-driven automation[45], deep learning, and real-time adaptation to bypass security measures, personalize phishing attacks, optimize malware delivery, and exfiltrate critical data with greater efficiency. The mentioned list of threat actors are not the only ones who embrace the usage of AI but are the ones for whom resources represent

#	Threat Actor Name	Attack Trend
1	Emerald Sleet (*THALLIUM / Kimusky/ APT 43/ Lazarus / Velvet Chollima /* TA406)	Phishing & Social Engineering, Malware
2	Charcoal Typhoon (*Aquatic Panda / Earth Lusca*)	Data Exfiltration, Malware
3	Forest Blizzard (APT 28 / *Fancy Bear / Strontium*)	Phishing & Social Engineering, Malware
4	Crimson Sandstrom (*Imperial Kitten*)	Phishing & Social Engineering, Data Exfiltration, Malware
5	Salmon Typhoon (*Maverick Panda / Sodium*)	Phishing & Social Engineering, Malware
6	CyberAv3ngers	Phishing & Social Engineering, Malware
7	SweetSpectre (*UNK_SweetSpecter*)	Phishing & Social Engineering, Malware

[44] Microsoft. (2024, February 14). *Staying ahead of threat actors in the age of* AI. Microsoft Security Blog. https://www.microsoft.com/en-us/security/blog/2024/02/14/staying-ahead-of-threat-actors-in-the-age-of-ai/

[45] CyberShujin. (n.d.). *Threat actors using* AI. GitHub. https://github.com/cybershujin/Threat-Actors-use-of-Artifical-Intelligence?tab=readme-ov-file

#	Threat Actor Name	Attack Trend
8	STORM-0817	Phishing & Social Engineering, Malware
9	Salt Typhoon (*GhostEmperor*, *FamousSparrow*, *Earth Estries*, UNC2286)	Phishing & Social Engineering, Data Exfiltration, Malware
10	Bitter APT (APT-C-08, *Aramanberry*)	Phishing & Social Engineering, Data Exfiltration
11	TA547 (*Scully Spider*)	Data Exfiltration, Malware
12	APT 44 (*Sandworm Team*)	Phishing & Social Engineering, Data Exfiltration
13	Baller423 (*goober2*)	Malware
14	Star23 (*baller13*)	Malware
15	GXC Team	Phishing & Social Engineering, Malware
16	Scattered Spider	Phishing & Social Engineering, Ransomware, Malware
17	Indrik Spider (*Evil Corp*)	Phishing & Social Engineering, Malware
18	CanadianKingpin12	Malware
19	DragonBridge	Phishing & Social Engineering
20	Lockbit	Ransomware, Malware
21	Ransomhub	Ransomware
22	BlackBasta	Phishing & Social Engineering, Ransomware, Data Exfiltration

#	Threat Actor Name	Attack Trend
23	ALPHV (*BlackCat*)	Ransomware, Data Exfiltration, Malware
24	Cactus Ransomware Group	Ransomware, Data Exfiltration
25	FunkSec	Phishing & Social Engineering, Ransomware, Malware
26	SneakyChef	Phishing & Social Engineering, Malware
27	TA499 (*Vovanand Lexus*)	Malware
28	Akira	Ransomware, Data Exfiltration
29	DragonForce	Ransomware, Data Exfiltration

Top 10 AI Associated Ransomware Threat Actors and Recorded Incidents

The table below highlights the top 10 AI-associated ransomware threat actors and their recorded incidents over the past three years (2023–2025[46]). These groups have evolved their tactics by integrating artificial intelligence (AI) into various stages of their attack cycles, making them some of the world's most dangerous and resilient ransomware operators.

By leveraging AI, these ransomware groups have been able to automate attack execution, improve target selection, and bypass security defenses more efficiently. AI-driven reconnaissance allows them to scan networks in real-time, identify critical assets, and exploit vulnerabilities precisely, ensuring that their attacks are fast and highly disruptive. Additionally, AI-powered malware enables adaptive evasion techniques, allowing ransomware to alter its code dynamically and remain undetected by traditional security measures.

As these groups continue to refine their methods, AI-driven ransomware attacks are expected to grow in scale, complexity, and effectiveness. Defending against these evolving

[46] Note: Incident Data used for analysis is up to February 2025

threats requires a multi-layered cybersecurity approach, including AI-powered threat detection, proactive monitoring, and stronger collaboration between organizations and security researchers to stay ahead of these increasingly sophisticated adversaries.

Threat Actor	**2023**	**2024**	**2025**[47]	**Total Events**
LockBit	1151	815	16	1982
Play	344	340	10	694
RansomHub	117	548	23	688
Akira	250	314	10	574
ALPHV (BlackCat)	346	155	5	506
BlackBasta	223	189	6	418
Medusa	145	199	2	346
Cactus	102	93	3	198
DragonForce	39	84	0	123
Funksec	34	74	8	116

Attack Trends Beyond Ransomware

Where various open-source threat intelligence sources indicate emerging events of AI usage by the adversary our analysis gives a more concerning view. Our analysis identifies structural attack trends and impacts where the advisory structurally usages artificial intelligence in their mode of operandi that goes beyond the ransomware threat[48]. While AI-driven

[47] Note: Incident Data used for analysis is up to February 2025
[48] Note: Incident Data used for analysis is up to February 2025

ransomware remains a major concern threat actors increasingly leverage artificial intelligence (AI) for a wide range of cyberattacks beyond just encryption-based extortion. The table above highlights emerging AI-powered attack trends that exploit machine learning models, automation, and advanced data analysis to enhance the effectiveness and efficiency of cyber threats.

Other AI-powered Attacks Trend

#	AI-powered Attack	AI Usage	Description
1	AI-Generated Phishing[49]	LLM-enhanced email crafting, automated spear-phishing, real-time adaptation	Cybercriminals use AI to generate highly personalized phishing emails that mimic legitimate communication. AI analyzes target behavior, past email history, and online activity to create compelling lures. AI-generated phishing emails have increased by over 1,000% since 2022, primarily targeting credential theft and financial fraud.
2	AI-Augmented Social Engineering[50]	AI-driven chatbots, deepfake-enhanced deception, NLP-based persuasion	AI is transforming social engineering attacks, making fake customer support calls, executive impersonation, and fraudulent communications more convincing. Hackers deploy AI chatbots and voice synthesis tools to engage in real-time conversations that manipulate victims into sharing sensitive information.

[49] SlashNext (2023, October 30). State of Phishing Report 2023. https://slashnext.com/press-release/slashnexts-2023-state-of-phishing-report-reveals-a-1265-increase-in-phishing-emails-since-the-launch-of-chatgpt-in-november-2022-signaling-a-new-era-of-cybercrime-fueled-by-generative-ai/

[50] Darktrace (2024, September 30). Business Email Compromise (BEC) in the Age of AI. https://www.darktrace.com/blog/business-email-compromise-bec-in-the-age-of-ai

#	AI-powered Attack	AI Usage	Description
3	Voice Cloning & AI-Impersonation[51]	Deep learning speech synthesis, biometric spoofing, voice morphing	AI-driven voice cloning technology enables cybercriminals to replicate a person's voice with high accuracy, bypassing voice authentication systems. Attackers use cloned voices to conduct scams, authorize financial transactions, and impersonate executives in corporate fraud schemes.
4	Deepfake Manipulation[52]	AI-generated video, synthetic identity creation, facial reenactment	AI is used to create realistic deepfake videos, allowing hackers to impersonate individuals in video calls, security verifications, and social engineering scams. This technology has been exploited for fraud, political misinformation, and extortion.
5	AI-Powered Malware[53]	AI-powered polymorphic code, self-modifying scripts, adaptive attack evasion	AI enables malware to learn from its environment, adjusting its code structure and behavior to avoid detection. AI-powered malware can alter signatures in real-time, bypassing antivirus tools and endpoint security. Some strains can autonomously detect sandbox environments and remain dormant until they reach a live target.

[51] McAfee (2023, May 15). Artificial Imposters—Cybercriminals Turn to AI Voice Cloning for a New Breed of Scam. https://www.mcafee.com/blogs/privacy-identity-protection/artificial-imposters-cybercriminals-turn-to-ai-voice-cloning-for-a-new-breed-of-scam/

[52] Deloitte (2025, January). The rise of deepfakes: What digital platforms and technology organizations should know. https://www.deloitte.com/uk/en/Industries/tmt/analysis/the-rise-of-deepfakes-what-digital-platforms-and-technology-organizations-should-know.html

[53] Impact Networking. (2024, December 27). AI-Generated Malware and How It's Changing Cybersecurity. https://www.impactmybiz.com/blog/how-ai-generated-malware-is-changing-cybersecurity/

#	AI-powered Attack	AI Usage	Description
6	AI-Accelerated Password Cracking[54]	Neural network-based brute force, predictive password analysis, AI-enhanced key discovery	AI dramatically improves password-cracking techniques by analyzing leaked credentials, predicting password patterns, and optimizing brute-force attacks. With machine learning, hackers reduce the time needed to crack passwords from weeks to hours, making traditional security measures more vulnerable.
7	AI-Powered CAPTCHA Bypass[55]	Adversarial AI image/text recognition, human behavior emulation	AI-driven bots are trained to solve CAPTCHAs at human-like accuracy, allowing cybercriminals to automate credential stuffing and mass account takeovers. These AI-powered CAPTCHA bypass tools mimic human interaction patterns to evade bot detection mechanisms.
8	ML Models Exploitation[56]	AI-assisted vulnerability research, adversarial input manipulation, model poisoning	Hackers are manipulating AI systems by feeding them deceptive data to cause incorrect outputs. This includes tricking fraud detection systems, bypassing AI-driven cybersecurity defenses, and disrupting facial recognition tools through adversarial attacks.

[54] Okta. (2025, February 12). How cracking passwords can be easier in the age of AI/ML. https://www.okta.com/blog/2025/02/how-cracking-passwords-can-be-easier-in-the-age-of-aiml/

[55] Scrapingapi. (2024, December 12). CAPTCHA Wars: Latest Statistics on Anti-Scraping Measures and Success Rates. https://scrapingapi.ai/blog/captcha-wars-latest-statistics-on-anti-scraping-measures-and-success-rates

[56] Radanliev, P & Santos, O. (2023). Adversarial Attacks Can Deceive AI Systems, Leading to Misclassification or Incorrect Decisions. https://www.preprints.org/manuscript/202309.2064

#	AI-powered Attack	AI Usage	Description
9	AI-Powered Data Scraping and Exploitation[57]	Automated web crawling, AI-assisted pattern recognition, targeted data mining	AI-powered scraping tools extract and analyze massive datasets from public and private sources faster than ever before. Cybercriminals use AI to correlate stolen data, identify high-value targets, and refine attack strategies. AI also helps bypass security filters meant to prevent automated data extraction.

Impact Caused By AI-powered Attacks

AI-driven cyber threats are reshaping the digital attack landscape, making traditional security measures increasingly ineffective. These advanced threats leverage artificial intelligence to accelerate attacks, enhance stealth, and optimize impact, often with minimal human intervention. Below are some of the most significant AI-powered attack outcomes and their far-reaching consequences.

#	Attack Outcomes	Description
1	AI-Enhanced Data Exfiltration	AI-driven attacks automate and optimize data theft by bypassing detection systems, compressing stolen data efficiently, mimicking normal traffic patterns, and using advanced obfuscation techniques to exfiltrate information stealthily. This allows cybercriminals to extract large amounts of sensitive data without triggering traditional security alerts. Example: ● Data compression efficiency, detection evasion success rate, stealth exfiltration accuracy, and encryption bypass effectiveness.

[57] Scrapingapi. (2024, December 4). The Rise of AI in Web Scraping: 2024 Stats That Will Surprise You. https://scrapingapi.ai/blog/the-rise-of-ai-in-web-scraping

#	Attack Outcomes	Description
2	AI-Augmented Ransomware Attacks	AI enhances ransomware attacks by identifying and prioritizing high-value targets, adapting encryption techniques to avoid detection, automating ransom negotiations, and dynamically generating polymorphic malware payloads. These enhancements make ransomware attacks faster, more targeted, and harder to mitigate. Example: ● Target prioritization accuracy, encryption speed, payload mutation rate, and negotiation response efficiency.
3	AI-Controlled Botnets & DDoS Attacks	AI-powered botnets leverage adaptive control mechanisms, real-time traffic analysis, and self-healing techniques to orchestrate massive, unpredictable, and highly resilient Distributed Denial-of-Service (DDoS) attacks. AI helps botnets dynamically adjust attack patterns to bypass mitigation tools, making them more effective against traditional defense mechanisms. Example: ● Attack adaptation speed, botnet resilience, evasion rate, and traffic modulation precision.
4	AI-Manipulated Data Corruption	Cybercriminals use AI to alter, poison, or selectively corrupt critical datasets, leading to misleading analytics, flawed decision-making, and business disruptions. AI-driven data poisoning is also a major threat to machine learning models, where attackers manipulate training data to degrade AI performance or introduce hidden vulnerabilities. Example: ● Data integrity degradation rate, anomaly evasion success rate, stealth injection precision, and AI model corruption effectiveness.

AI-Powered Ransomware Threats Not Going Away

We have highlighted some key attack scenarios, demonstrated adversarial functionalities observed in controlled environments, and explored methodologies inspired by past advancements in cyber threats. Our data-driven research confirms that AI-powered threats are not just emerging—they have already taken hold, and their impact has been unfolding for some time.

This research involves 2,800 recorded ransomware incidents of which more than 80% involves AI used by the most prominent ransomware groups. Our work identified close to 30 threat actors that represent resources of AI-driven automation, deep learning, and real-time adaptation to bypass security measures, personalize phishing attacks, optimize malware delivery, and exfiltrate critical data with greater efficiency.

We observe that AI powered ransomware attacks are widely adopted by different ransomware groups including criminal organizations, ransomware-as-a -service providers, and alleged state sponsored adversaries. This evolution makes ransomware attacks faster, more efficient, and harder to detect. These observed functionalities are de facto the standard of common adversarial ransomware capabilities as from today and it reveals the application of adversarial agentic agents.

This rapid advances in artificial intelligence presents both a transformative opportunity and a formidable challenge in cyber defense. Adversaries are increasingly leveraging AI to develop sophisticated threats, escalating the complexity and scale of cyberattacks, particularly ransomware. As AI-driven threats evolve, organizations must proactively adapt their cybersecurity strategies to counteract these emerging risks.

Defenders' Reaction: AI-Powered Cyber Risk and Resilience Management

As cyber threats grow in sophistication and scale, defenders must leverage AI to enhance their cyber risk management and resilience strategies. AI offers transformative capabilities key areas include: Automated Security Hygiene, Autonomous and Deceptive Defense Systems, and Augmented Oversight and Reporting. These AI-driven solutions improve proactive defenses, streamline security operations, and strengthen overall organizational resilience against emerging cyber threats.

Automated Security Hygiene strengthens the foundational cyber defenses. A significant portion of cyber attacks exploit fundamental vulnerabilities like unpatched software, misconfigurations, and compromised credentials. AI-driven automation can address these issues by solutions like self-healing software code[58], self-patching systems[59], continuous

[58] ABN. (2021, March 19). ABN AMRO first buyer of innovative self-healing cybersecurity software. ABN AMRO. https://www.abnamro.com/en/news/abn-amro-first-buyer-of-innovative-self-healing-cybersecurity-software

[59] Sibanda, I. (2024, November 29). Automated patch management: A proactive way to stay ahead of threats. ComputerWeekly. https://www.computerweekly.com/feature/Automated-patch-management-A-proactive-way-to-stay-ahead-of-threats

attack surface management[60], zero trust based architecture[61], and self-driving trustworthy networks[62]. By integrating AI into security hygiene, organizations reduce manual workload while enhancing resilience against attacks that exploit foundational weaknesses.

Autonomous and Deceptive Defense Systems provides proactive and adaptive cybersecurity. AI enables defenders to shift from reactive to proactive defense strategies. Autonomous cyber defense systems, such as Extended detection and response (XDR)[63] and Security orchestration, automation, and response (SOAR)[64] systems, use real-time analytics and machine learning to identify, contain, and mitigate threats autonomously while continuously learning from telemetry data to anticipate and counteract threats. Known examples are Simultaneously automated moving target defense[65] automatically and frequently alters system configurations that increase complexity for attackers and reduces windows of opportunities. And finally deceptive tactics and misinformation[66] can mislead adversaries, slowing their progress, reducing attack success rates while the defender as a longer window of opportunity to detect the adversary. These capabilities shorten threat detection times, reduce response efforts, and enhance overall cyber resilience.

Augmented Oversight and Reporting will enhance decision-making with AI. These AI-powered analytics provide executives with a deeper understanding of cyber risks and their impact on business operations and provide near real-time data driven insights to strengthen decision making. For instance automated AI driven risk quantification helps to assess and evaluate emerging vulnerabilities and predict future risk impacts[67]. In addition simulation technology provide interactive and exploratory environment to simulate cyber threats and test long term effectiveness of security strategies before implementation[6869]. An advanced supply chain security analytics map interconnected risks and improve third-party risk

[60] Vindhya, L., Mahima, B. Gowda, Gowramma Gaari Sindhu, & Keerthan, V. (2023). International Journal of Advanced Research in Science, Communication and Technology (IJARSCT, 3(8), April 2023. https://ijarsct.co.in/Paper9533.pdf

[61] K. Chokkanathan, S. M. Karpagavalli, G. Priyanka, K. Vanitha, K. Anitha and P. Shenbagavalli, "AI-Driven Zero Trust Architecture: Enhancing Cyber-Security Resilience," 2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 2024, pp. 1-6, doi: 10.1109/CSITSS64042.2024.10816746.

[62] Hireche, O., Benzaïd, C., & Taleb, T. (2022). Deep data plane programming and AI for zero-trust self-driven networking in beyond 5G. Computer Networks, 203, 108668. https://doi.org/10.1016/j.comnet.2021.108668

[63] Joshi, J. (2024, May 13). The impact of AI on endpoint detection and response. Proficio. https://www.proficio.com/blog/ai-endpoint-detection-and-response-edr/

[64] Cyware. (2023, December 1). SOAR and AI in cybersecurity: Reshaping your security operations. Cyware. https://cyware.com/security-guides/security-orchestration-automation-and-response/from-insight-to-action-how-ai-and-soar-are-reshaping-security-operations-13d9

[65] Jajodia, S., Cybenko, G., Liu, P., Wang, C., & Wellman, M. (Eds.). (2019). Adversarial and uncertain reasoning for adaptive cyber defense: Control-and game-theoretic approaches to cybersecurity (Vol. 11830). Springer Nature.

[66] Duy, P. T., Hoang, H. D., Khoa, N. H., Hien, D. T., & Pham, V. (2022). Fool your enemies: Enable cyber deception and moving target defense for intrusion detection in SDN. 2022 21st International Symposium on Communications and Information Technologies (ISCIT), 27–32. https://ieeexplore.ieee.org/abstract/document/9931208

[67] Sanna, N. (2025, February 25). How SAFE Is Transforming Cyber Risk Management for a Secure Digital Future. Save Security. https://safe.security/resources/blog/safe-transforming-cyber-risk-management-secure-digital-future/

[68] Zeijlemaker, S. and Siegel, M., "Capturing the Dynamic Nature of Cyber Risk: Evidence from an Explorative Case Study" (2023). Hawaii International Conference on System Sciences 2023 (HICSS-56). 5. https://aisel.aisnet.org/hicss-56/os/cybersecurity/5

[69] Zeijlemaker, S.; Pal, R.; and Siegel, Ml, "Strengthening Managerial Foresight to Defeat Cyber Threats" (2024). AMCIS 2024 Proceedings. 18. https://aisel.aisnet.org/amcis2024/security/security/18

management[70]. By leveraging AI for cyber risk governance and oversight, organizations gain predictive insights and strategic clarity, strengthening cyber risk and resilience management.

AI-driven cybersecurity solutions address all steps of the adversarial attack process. Automated Security Hygiene ensures a strong foundation for security, while Autonomous and Deceptive Defense Systems provide real-time, adaptive responses to cyber threats. AI-powered oversight enhances decision-making and governance. While AI-driven cybersecurity solutions offer significant advantages, they also introduce new challenges. Attackers will seek to exploit AI systems, necessitating continuous monitoring and oversight of AI-powered defenses. Organizations must integrate AI strategically, ensuring that security teams are equipped to manage and govern these technologies effectively.

Unfortunately, this is easier said than done as many defenders already cope with resource limitations[71], face regulatory or ethical constraints in using AI[72], and deal with the complexity of such implementations[73]. This discrepancy favors the adversaries who can exploit AI's capabilities with fewer restrictions.
By timely embracing AI in cyber risk management, defenders can enhance resilience, reduce operational burdens, and maintain a proactive stance against evolving cyber threats. A structured, AI-enabled security roadmap ensures that organizations remain ahead of adversaries, safeguarding business operations and critical assets in an increasingly complex cyber landscape.

[70] Wasi, A. T., Islam, M. D., Akib, A. R., & Bappy, M. M. (2024). Graph Neural Networks in Supply Chain Analytics and Optimization: Concepts, Perspectives, Dataset and Benchmarks. arXiv preprint arXiv:2411.08550.

[71] Morgan, S. (2023, April 14). Cybersecurity Jobs Report: 3.5 Million Unfilled Positions In 2025, Cybercrime Magazine. https://cybersecurityventures.com/jobs/

[72] Helkala, K., Cook, J., Lucas, G., Pasquale, F., Reichberg, G., & Syse, H. (2022). AI in cyber operations: ethical and legal considerations for end-users. In Artificial Intelligence and Cybersecurity: Theory and Applications (pp. 185-206). Cham: Springer International Publishing.

[73] AL-Dosari, K., Fetais, N., & Kucukvar, M. (2024). Artificial intelligence and cyber defense system for banking industry: A qualitative study of AI applications and challenges. Cybernetics and systems, 55(2), 302-330. https://www.tandfonline.com/doi/pdf/10.1080/01969722.2022.2112539

References

1. ABN. (2021, March 19). ABN AMRO first buyer of innovative self-healing cybersecurity software. ABN AMRO. https://www.abnamro.com/en/news/abn-amro-first-buyer-of-innovative-self-healing-cybersecurity-software
2. AkiraDecryptor. (2025). *Akira ransomware & AI-driven attacks: Poised to unleash AI-driven attacks in 2025, experts warn*. https://akiradecryptor.org/akira-ransomware-poised-to-unleash-ai-driven-attacks-in-2025-experts-warn/
3. Alabdulatif, A., & Thilakarathne, N. N. (2025). Hacking Exposed: Leveraging Google Dorks, Shodan, and Censys for Cyber Attacks and the Defense Against Them. Computers, 14(1), 24. https://www.mdpi.com/2073-431X/14/1/24
4. AL-Dosari, K., Fetais, N., & Kucukvar, M. (2024). Artificial intelligence and cyber defense system for banking industry: A qualitative study of AI applications and challenges. Cybernetics and systems, 55(2), 302-330. https://www.tandfonline.com/doi/pdf/10.1080/01969722.2022.2112539
5. Belova, K. (2023, December 15). What is federated learning: Key benefits, applications, and working principles explained. *PixelPlex*. https://pixelplex.io/blog/federated-learning-guide/
6. BleepingComputer. (n.d.). *Maze ransomware analysis*. https://www.bleepingcomputer.com/news/security/maze-ransomware/
7. BleepingComputer. (n.d.). *Clop overview*. https://www.bleepingcomputer.com/tag/clop-ransomware/
8. Check Point Research. (2022). *Predictive modeling in malware attacks*. https://research.checkpoint.com/2022/predictive-modeling-in-malware-attacks/
9. Check Point Research. (2025). *AI-powered ransomware: FunkSec—Alleged top ransomware group powered by AI*. Check Point. https://research.checkpoint.com/2025/funksec-alleged-top-ransomware-group-powered-by-ai
10. Chokkanathan, K., Karpagavalli, S. M., Priyanka, G., Vanitha, K., Anitha, K., & Shenbagavalli, P. (2024). AI-driven zero trust architecture: Enhancing cyber-security resilience. 2024 *8th International Conference on Computational System and Information Technology for Sustainable Solutions* (CSITSS), Bengaluru, India, 1–6. https://doi.org/10.1109/CSITSS64042.2024.10816746
11. CrowdStrike. (n.d.). *Ryuk ransomware report*. https://www.crowdstrike.com/blog/ryuk-ransomware-technical-analysis/
12. CrowdStrike. (2021). *Ransomware propagation techniques in network environments*. https://www.crowdstrike.com/blog/ransomware-propagation-techniques/
13. Cybersecurity and Infrastructure Security Agency. (n.d.). *Ransomware-as-a-service*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-075a
14. Cybersecurity and Infrastructure Security Agency. (n.d.). *ALPHV Blackcat overview*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-353a
15. Cybersecurity and Infrastructure Security Agency. (n.d.). *BlackBasta overview*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-131a
16. Cybersecurity and Infrastructure Security Agency. (n.d.). *LockBit overview*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-165a

17. Cybersecurity and Infrastructure Security Agency. (n.d.). *RansomHub overview*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-242a
18. Cybersecurity and Infrastructure Security Agency. (n.d.). *Conti leak analysis*. CISA. https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-265a
19. Cybersecurity and Infrastructure Security Agency. (n.d.). *REvil overview*. CISA. https://www.cisa.gov/stopransomware/revil-sodinokibi
20. CyberShujin. (n.d.). *Threat actors using AI*. GitHub. https://github.com/cybershujin/Threat-Actors-use-of-Artifical-Intelligence?tab=readme-ov-file
21. Cyware. (2023, December 1). SOAR and AI in cybersecurity: Reshaping your security operations. *Cyware*. https://cyware.com/security-guides/security-orchestration-automation-and-response/from-insight-to-action-how-ai-and-soar-are-reshaping-security-operations-13d9
22. Darktrace (2024, April 24). How Cactus Ransomware was detected and stopped. https://www.darktrace.com/blog/a-thorn-in-attackers-sides-how-darktrace-uncovered-a-cactus-ransomware-infection
23. Darktrace (2024, September 30). Business Email Compromise (BEC) in the Age of AI. https://www.darktrace.com/blog/business-email-compromise-bec-in-the-age-of-ai
24. Davies, V. (2021, October 4). The history of cybersecurity. *Cyber Magazine*. https://cybermagazine.com/cyber-security/history-cybersecurity
25. Deloitte (2025, January). The rise of deepfakes: What digital platforms and technology organizations should know. https://www.deloitte.com/uk/en/Industries/tmt/analysis/the-rise-of-deepfakes-what-digital-platforms-and-technology-organizations-should-know.html
26. Deloitte. (2024). *AI threat report*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-design-ai-threat-report-v2.pdf
27. Digital, N. (2023, January 16). Benefits of federated learning explained. *OctaiPipe*. Retrieved September 16, 2024, from https://octaipipe.ai/benefits-of-federated-learning-explained/
28. Duy, P. T., Hoang, H. D., Khoa, N. H., Hien, D. T., & Pham, V. (2022). Fool your enemies: Enable cyber deception and moving target defense for intrusion detection in SDN. *2022 21st International Symposium on Communications and Information Technologies (ISCIT)*, 27–32. https://ieeexplore.ieee.org/abstract/document/9931208
29. Halcyon. (n.d.). *DragonForce ransomware attack threatens EvoEvents Limited*. Halcyon. https://www.halcyon.ai/attacks/dragonforce-ransomware-attack-threatens-evoevents-limited
30. Hays. (2024). EyeSpy: Proof-of-concept. *Hays*. https://www.hyas.com/read-the-eyespy-proof-of-concept
31. Helkala, K., Cook, J., Lucas, G., Pasquale, F., Reichberg, G., & Syse, H. (2022). AI in cyber operations: ethical and legal considerations for end-users. In Artificial Intelligence and Cybersecurity: Theory and Applications (pp. 185-206). Cham: Springer International Publishing.
32. Hireche, O., Benzaïd, C., & Taleb, T. (2022). Deep data plane programming and AI for zero-trust self-driven networking in beyond 5G. *Computer Networks*, 203, 108668. https://doi.org/10.1016/j.comnet.2021.108668
33. Huang, Y., Huang, L., & Zhu, Q. (2022). Reinforcement learning for feedback-enabled cyber resilience. *Annual Reviews in Control*, 53, 273–295. https://doi.org/10.1016/j.arcontrol.2022.01.001

34. IBM Security. (2022). *Reconnaissance in ransomware attacks: Disabling backups and cloud services*. https://www.ibm.com/security/intelligence/ransomware-reconnaissance
35. Ilascu, I. (2024, April 10). Malicious PowerShell script pushing malware looks AI-written. *Bleeping Computer*. https://www.bleepingcomputer.com/news/security/malicious-powershell-script-pushing-malware-looks-ai-written/
36. Impact Networking. (2024, December 27). AI-Generated Malware and How It's Changing Cybersecurity. https://www.impactmybiz.com/blog/how-ai-generated-malware-is-changing-cybersecurity/
37. Jajodia, S., Cybenko, G., Liu, P., Wang, C., & Wellman, M. (Eds.). (2019). *Adversarial and uncertain reasoning for adaptive cyber defense: Control-and game-theoretic approaches to cybersecurity* (Vol. 11830). Springer Nature.
38. Joshi, J. (2024, May 13). The impact of AI on endpoint detection and response. *Proficio*. https://www.proficio.com/blog/ai-endpoint-detection-and-response-edr/
39. Kharraz, A., & Robertson, W. (2018). *Ransoms: The emerging threat of ransomware*. ACM Digital Library. https://dl.acm.org/doi/10.1145/3180445.3180449
40. K. Chokkanathan, S. M. Karpagavalli, G. Priyanka, K. Vanitha, K. Anitha, & P. Shenbagavalli. (2024). AI-Driven Zero Trust Architecture: Enhancing Cyber-Security Resilience. *2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, Bengaluru, India, 1–6. https://doi.org/10.1109/CSITSS64042.2024.10816746
41. Lin, Z., Shi, Y., & Xue, Z. (2022). IDSGAN: Generative adversarial networks for attack generation against intrusion detection. *Lecture Notes in Computer Science*, 79–91. Springer International Publishing. https://doi.org/10.1007/978-3-031-05981-0_7
42. McAfee (2023, May 15). Artificial Imposters—Cybercriminals Turn to AI Voice Cloning for a New Breed of Scam. https://www.mcafee.com/blogs/privacy-identity-protection/artificial-imposters-cybercriminals-turn-to-ai-voice-cloning-for-a-new-breed-of-scam/
43. Microsoft Security. (2021, February 11). *Rapid propagation techniques*. https://www.microsoft.com/security/blog/2021/02/11/ryuk-ransomware-analysis/
44. Microsoft Security Intelligence. (2021, May 27). *Ransomware's reconnaissance techniques*. https://www.microsoft.com/security/blog/2021/05/27/ransomware-reconnaissance-before-the-heist/
45. Microsoft. (2024, February 14). *Staying ahead of threat actors in the age of AI*. Microsoft Security Blog. https://www.microsoft.com/en-us/security/blog/2024/02/14/staying-ahead-of-threat-actors-in-the-age-of-ai/
46. MITRE ATT&CK Framework. (2023). *Defense evasion techniques in modern malware*. MITRE. https://attack.mitre.org/tactics/TA0005/
47. Morgan, S. (2023, April 14). Cybersecurity Jobs Report: 3.5 Million Unfilled Positions In 2025, Cybercrime Magazine. https://cybersecurityventures.com/jobs/
48. Nelson, N. (2023, November 3). Somebody just killed the Mozi Botnet. *DarkReading*. https://www.darkreading.com/ics-ot-security/somebody-just-killed-mozi-botnet
49. OpenAI. (2024, October 9). An update on disrupting deceptive uses of AI. *OpenAI*. https://openai.com/global-affairs/an-update-on-disrupting-deceptive-uses-of-ai/
50. Okta. (2025, February 12). How cracking passwords can be easier in the age of AI/ML. https://www.okta.com/blog/2025/02/how-cracking-passwords-can-be-easier-in-the-age-of-aiml/

51. Palo Alto Networks. (n.d.). *Clop technical analysis*. https://unit42.paloaltonetworks.com/clop-ransomware/
52. Palo Alto Networks. (2021). *Ransomware families and their evasion techniques*. Unit 42. https://unit42.paloaltonetworks.com
53. Ph, M., Stoecklin, J., Jang, J., & Kirat, D. (2018, August 8). DeepLocker: How AI can power a stealthy new breed of malware. Security Intelligence. Yang, K.-C., & Menczer, F. (2023, July 30). Anatomy of an AI-powered malicious social botnet. arXiv Preprint. https://arxiv.org/abs/2307.16336
54. Radanliev, P & Santos, O. (2023). Adversarial Attacks Can Deceive AI Systems, Leading to Misclassification or Incorrect Decisions. https://www.preprints.org/manuscript/202309.2064
55. Rigaki, M., & Garcia, S. (2018). Bringing a GAN to a knife-fight: Adapting malware communication to avoid detection. *IEEE Security and Privacy Workshops (SPW)*, 70–75. https://doi.org/10.1109/SPW.2018.00019
56. Sanna, N. (2025, February 25). How SAFE is transforming cyber risk management for a secure digital future. *SAFE Security*. https://safe.security/resources/blog/safe-transforming-cyber-risk-management-secure-digital-future/
57. Schröer, S. L., Apruzzese, G., Human, S., Laskov, P., Anderson, H. S., Bernroider, E. W. N., Fass, A., Nassi, B., Rimmer, V., Roli, F., Salam, S., Shen, A., Sunyaev, A., Wadhwa-Brown, T., Wagner, I., & Wang, G. (2025). SoK: On the offensive potential of AI. *IEEE Conference on Secure and Trustworthy Machine Learning*.
58. SentinelOne. (n.d.). *LockBit 3.0 overview*. https://www.sentinelone.com/anthology/lockbit-3-0-lockbit-black/
59. SentinelOne. (2024, February). *Cybercrime & AI-driven threats*. SentinelOne Blog. https://www.sentinelone.com/blog/february-2024-cybercrime-update-commercial-spyware-ai-driven-apts-and-flawed-rmms/
60. Scrapingapi. (2024, December 12). CAPTCHA Wars: Latest Statistics on Anti-Scraping Measures and Success Rates. https://scrapingapi.ai/blog/captcha-wars-latest-statistics-on-anti-scraping-measures-and-success-rates
61. Scrapingapi. (2024, December 4). The Rise of AI in Web Scraping: 2024 Stats That Will Surprise You. https://scrapingapi.ai/blog/the-rise-of-ai-in-web-scraping
62. SlashNext (2023, October 30). State of Phishing Report 2023. https://slashnext.com/press-release/slashnexts-2023-state-of-phishing-report-reveals-a-1265-increase-in-phishing-emails-since-the-launch-of-chatgpt-in-november-2022-signaling-a-new-era-of-cybercrime-fueled-by-generative-ai/
63. Sibanda, I. (2024, November 29). Automated patch management: A proactive way to stay ahead of threats. *ComputerWeekly*. https://www.computerweekly.com/feature/Automated-patch-management-A-proactive-way-to-stay-ahead-of-threats
64. Sims, J. (2023, July 31). BlackMamba: Using AI to generate polymorphic malware. Hyas. https://www.hyas.com/blog/blackmamba-using-ai-to-generate-polymorphic-malware
65. Sophos. (2021, February 16). *Conti ransomware technical analysis*. https://news.sophos.com/en-us/2021/02/16/conti-ransomware-evasive-by-nature/
66. Sophos. (2021, July 7). *Ransomware intelligence*. https://news.sophos.com/en-us/2021/07/07/revil-ransomware-attack-analysis/

67. Sophos. (2021). *The state of ransomware 2021*. https://www.sophos.com/en-us/medialibrary/pdfs/whitepaper/sophos-state-of-ransomware-2021-wp.pdf
68. Symantec. (n.d.). *Emotet and Ryuk collaboration*. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/emotet-ryuk-conti-ransomware
69. Symantec Threat Intelligence. (2022). *Polymorphic malware in advanced threat campaigns*. Symantec. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/polymorphic-malware-advanced-threats
70. The Grungq. (2017). The Triple A Threat: Aggressive autonomous agents. BlackHat. https://www.blackhat.com/docs/webcast/12142017-the-triple-a-threat.pdf
71. Toulas, B. (2024, September 24). Hackers deploy AI-written malware in targeted attacks. Bleeping Computer. https://www.bleepingcomputer.com/news/security/hackers-deploy-ai-written-malware-in-targeted-attacks/
72. Trend Micro. (n.d.). *LockBit analysis*. https://www.trendmicro.com/vinfo/us/security/news/ransomware-spotlight/ransomware-spotlight-lockbit
73. Trend Micro. (2022). *AI in ransomware: Personalized attacks and adaptive demands*. https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/ai-in-ransomware-personalized-attacks
74. Unit 42. (n.d.). *Advanced reconnaissance*. Palo Alto Networks. https://unit42.paloaltonetworks.com/conti-ransomware/
75. Unit 42. (n.d.). *Technical analysis*. Palo Alto Networks. https://unit42.paloaltonetworks.com/maze-ransomware-campaign/
76. United States Computer Emergency Readiness Team. (n.d.). *Emotet technical overview*. US-CERT. https://us-cert.cisa.gov/ncas/alerts/aa20-280a
77. Vindhya, L., Mahima, B. G., Gowda, G. G. S., & Keerthan, V. (2023). [Title of the article]. *International Journal of Advanced Research in Science, Communication and Technology* (IJARSCT), *3*(8). https://ijarsct.co.in/Paper9533.pdf
78. Wasi, A. T., Islam, M. D., Akib, A. R., & Bappy, M. M. (2024). Graph neural networks in supply chain analytics and optimization: Concepts, perspectives, dataset and benchmarks. *arXiv Preprint*. https://arxiv.org/abs/2411.08550
79. Wired. (n.d.). *Kaseya attack – Detailed report*. https://www.wired.com/story/kaseya-ransomware-attack-revil/
80. Yang, K.-C., & Menczer, F. (2023, July 30). Anatomy of an AI-powered malicious social botnet. *arXiv Preprint*. https://arxiv.org/abs/2307.16336
81. Zeijlemaker, S., & Siegel, M. (2023). Capturing the dynamic nature of cyber risk: Evidence from an explorative case study. *Hawaii International Conference on System Sciences 2023* (HICSS-56). https://aisel.aisnet.org/hicss-56/os/cybersecurity/5
82. Zeijlemaker, S., Pal, R., & Siegel, M. (2024). Strengthening managerial foresight to defeat cyber threats. *AMCIS 2024 Proceedings*, (18). https://aisel.aisnet.org/amcis2024/security/security/18
83. Zhang, C., Pal, R., Nicholson, C., & Siegel, M. (2024, December). (Gen) AI versus (Gen) AI in industrial control cybersecurity. *2024 Winter Simulation Conference* (WSC), 2739–2750. IEEE. https://doi.org/10.5555/3712729.3712956

84. Zhao, S., Li, J., Wang, J., Zhang, Z., Zhu, L., & Zhang, Y. (2021). attackGAN: Adversarial attack against black-box IDS using generative adversarial networks. Procedia Computer Science, 187, 128–133. https://doi.org/10.1016/j.procs.2021.04.118