perimattic

AI-Generated Threats And AI-Powered Defense

The Cybersecurity Arms Race

AUGUST 2025

perimattic

Table Of Contents

1. Introduction
- Automation, Sophistication, Evasion
- The Other Side: Defense Powered by AI
2. ‘Smarter’ Cybersecurity Threats
- Why AI is Such an Enabler for Cybersecurity Threats
3. Why AI Threats Feel More ‘Real’ Than Human-Generated Ones
4. Industries Most Vulnerable to AI Cyber Threats
5. How Threat Actors Exploit Industry-Specific Vulnerabilities with AI
6. Different Impacts Across Sectors
7. Staying Ahead of Miscreants: The Rise for Proactive AI-Powered Defense
- Level 1: Private and Corporate Initiatives
- Level 2: Government-Level Efforts
- Level 3: Personal-Level Protection
8. Turning Innovation into Advantage

perimattic

Introduction

Every transformative technology carries both promise and peril. From the advent of computers and the World Wide Web to the development of faster arms, more precise ammunition, and advancements in automotive engineering, the same tools that enable progress can be repurposed for harm. Cars provided mobility, yet also became instruments of destruction; the web democratized information, yet opened channels for global cyber threats. This duality underscores a timeless truth: innovation often begets new vulnerabilities.

Artificial Intelligence (AI) is no exception. As it fuels innovation by using information and data to learn and perform tasks, it also arms threat actors with more potent methods. Cybercriminals use AI to automate attacks, craft highly targeted phishing campaigns, and generate evasive malware.

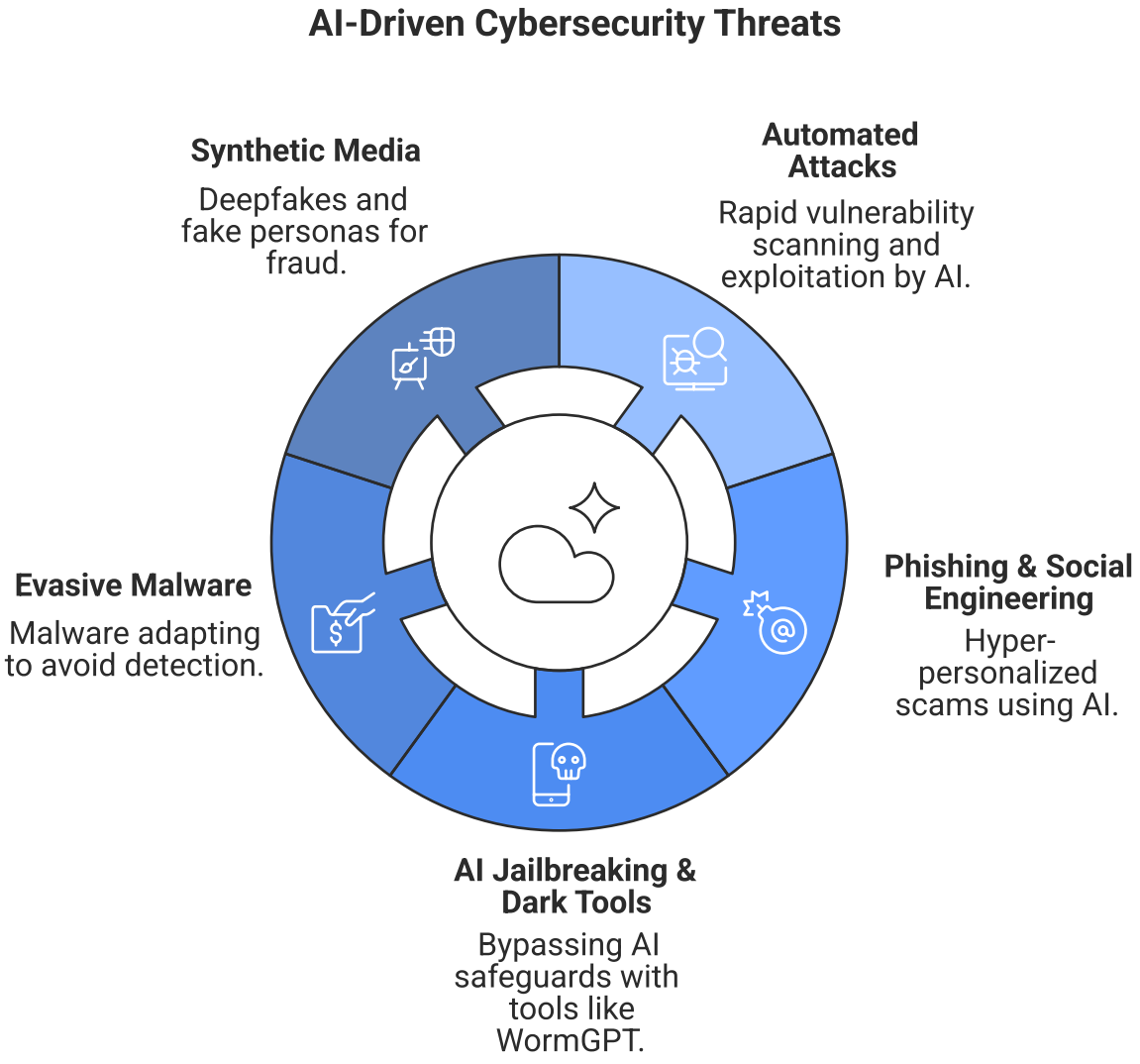
In KELA's 2025 AI Threat Report, researchers noted **200% surge** in mentions of malicious AI tools on cybercrime forums, signalling an explosive shift toward weaponizing AI in underground spaces.

perimattic

Automation, Sophistication, Evasion

AI is turbocharging cybercrime. Phishing and social engineering campaigns becoming more effective, and malware, especially ransomware and infostealers, is also evolving to evade detection, thanks to AI's speed and adaptability. Jailbreaking of public AI models is on the rise, with criminals exploiting prompt injection techniques to bypass safety protocols and co-opt powerful language models like ChatGPT, Gemini, and Claude.

Simultaneously, Europol and other agencies are sounding alarms: **organized crime is adopting AI to automate multilingual messaging, craft believable impersonations, and deploy synthetic media like deepfakes and voice cloning for fraud, recruitment, and exploitation.**



perimattic

Generative AI intensifies threats manifold. Deepfakes (face-swapped videos), voice-cloned calls, entirely synthetic personas, all of these undermine trust in media, enable fraud, and manipulate information. Scenarios like digital identity theft, ‘digital arrests’ via fabricated evidence, and misinformation campaigns underscore how generative AI elevates the stakes.

perimattic

The Other Side: Defense Powered by AI

While threat actors escalate their tactics, defenders are stepping up, also powered by AI. Cybersecurity teams are deploying machine learning to detect anomalies across networks, flag deepfake media, and triage threats at machine speed. Tech giants are coming forward to invest in research and education for responsible use of AI

Google's Responsible AI Progress Report notes that the organization has spent for $120 million for **AI education and training around the world**. Also notable is that 19,000 security professionals have taken their **SAIF Risk Self-Assessment to receive a personalized report of AI risks relevant to their organization.**

Industry experts emphasize AI's indispensable role in cyber resiliency, advocating for zero-trust models, continuous behavioural monitoring, and adversarial testing to stay one step ahead.

It is a cybersecurity arms race and this whitepaper considers both sides so that organizations can make informed decisions pertaining to the threats and security measures they need to consider. We will look at:

- **How cyber threats have become 'smarter'** and why traditional defenses are straining against rapid automation and sophistication.
- **Which industries are most vulnerable** to AI-powered attacks and how they can mitigate these emerging risks.
- **How defenders are adapting** by innovating AI-native solutions, refining strategies, and harnessing real-time intelligence in an accelerating arms race.

perimattic

‘Smarter’ Cybersecurity Threats

AI has not invented cybercrime, but it has radically augmented the efficiency, precision, and believability of attacks. What used to require days or weeks of manual planning can now be executed in minutes, at a scale human attackers could never match. Traditional defenses, designed to detect known patterns or respond to slower-moving threats, are struggling to keep up with the speed and sophistication that AI brings to the table.

Why AI is Such an Enabler for Cybersecurity Threats

Several technical factors make AI a force multiplier for malicious activity:

1. Automation at Scale

AI can autonomously scan millions of IP addresses and deploy exploits, in parallel. This removes the bottleneck of human labour and vastly increases the scope of an attack.

2. Pattern Recognition & Exploit Discovery

Machine learning models can sift through massive datasets, from leaked credentials to code repositories, and identify exploitable patterns faster than traditional scanning tools.

3. Adaptive Evasion

AI-driven malware can alter its code, communication methods, and behaviour on the fly to bypass intrusion detection systems and spam filters.

4. Hyper-Personalization

Large Language Models (LLMs) can craft highly tailored phishing messages, deepfake content, or fraudulent scripts that match a target’s language, tone, and even cultural references, making detection far harder.

Sophisticated cyberattacks used to require deep technical expertise and expensive infrastructure. Today, malicious AI tools are widely accessible, often sold or shared on dark web marketplaces, encrypted chat groups, and even public GitHub repositories. 'Plug-and-play' AI models and jailbreak guides are making it possible for low-skill actors to deploy highly advanced attacks with minimal effort or cost. This democratization of attack capability is one of the most alarming shifts in modern cybersecurity.

Recently, cybercriminals released a fake AI-generated video of YouTube CEO Neal Mohan announcing changes to the platform's payout policy.

This incident highlights a critical vulnerability: YouTube is not just a video platform, it is a central hub for **creators, influencers, brands, and advertisers**. An AI-driven scam here has far-reaching consequences:

- **Creators** could be tricked into sharing login credentials or banking information.
- **Influencers** could inadvertently spread fraudulent announcements to millions of followers.
- **Brands** could suffer reputational damage if their ad campaigns appear alongside scam content.
- **Advertisers** risk financial loss if ad spend is redirected through fraudulent channels.

The credibility of a platform amplifies the damage when that trust is exploited. AI deepfakes erode that credibility faster than any human-generated scam could.

perimattic

Why AI Threats Feel More ‘Real’ Than Human-Generated Ones

A study on Arvix found phishing emails created by large language models had a 54% click-through rate, far higher than the 12% rate for likely human-written messages. This is because AI-generated content:

- Mimics natural language patterns with minimal grammatical errors.
- Adjusts tone, style, and cultural context based on the recipient’s profile.
- Embeds persuasive, contextually relevant details that a generic human-written email would likely miss.

For the recipient, the message doesn’t feel like a scam, it feels like communication from a known, credible source.

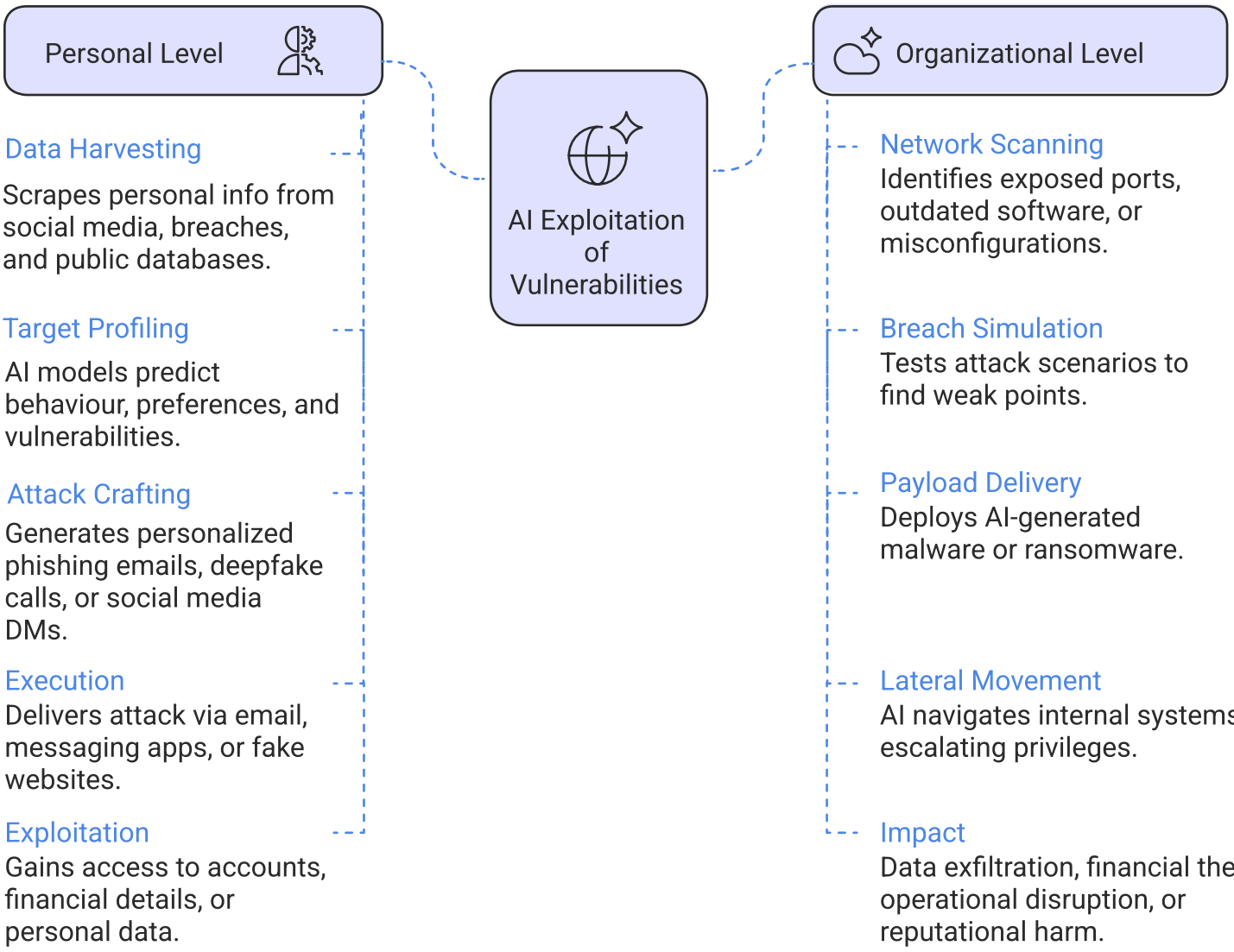
Phishing, malware, ransomware, and impersonation scams existed long before machine learning models. But AI fundamentally changes the scale and quality of these attacks.

1. **Speed**: AI can write and launch thousands of phishing campaigns or malware variants in seconds, compressing what used to be weeks of work into moments.
2. **Volume**: Automation allows for relentless, simultaneous targeting of thousands (even millions) of individuals or organizations without additional human effort.
3. **Sophistication**: AI refines attack methods through continuous feedback, making each subsequent attempt more convincing and harder to detect. For example, failed phishing attempts can be automatically analysed, with future messages optimized for better success rates.

Traditional defenses, often reactive and rule-based, were never designed for this level of constant, adaptive assault.

perimattic

How AI Exploits Vulnerabilities



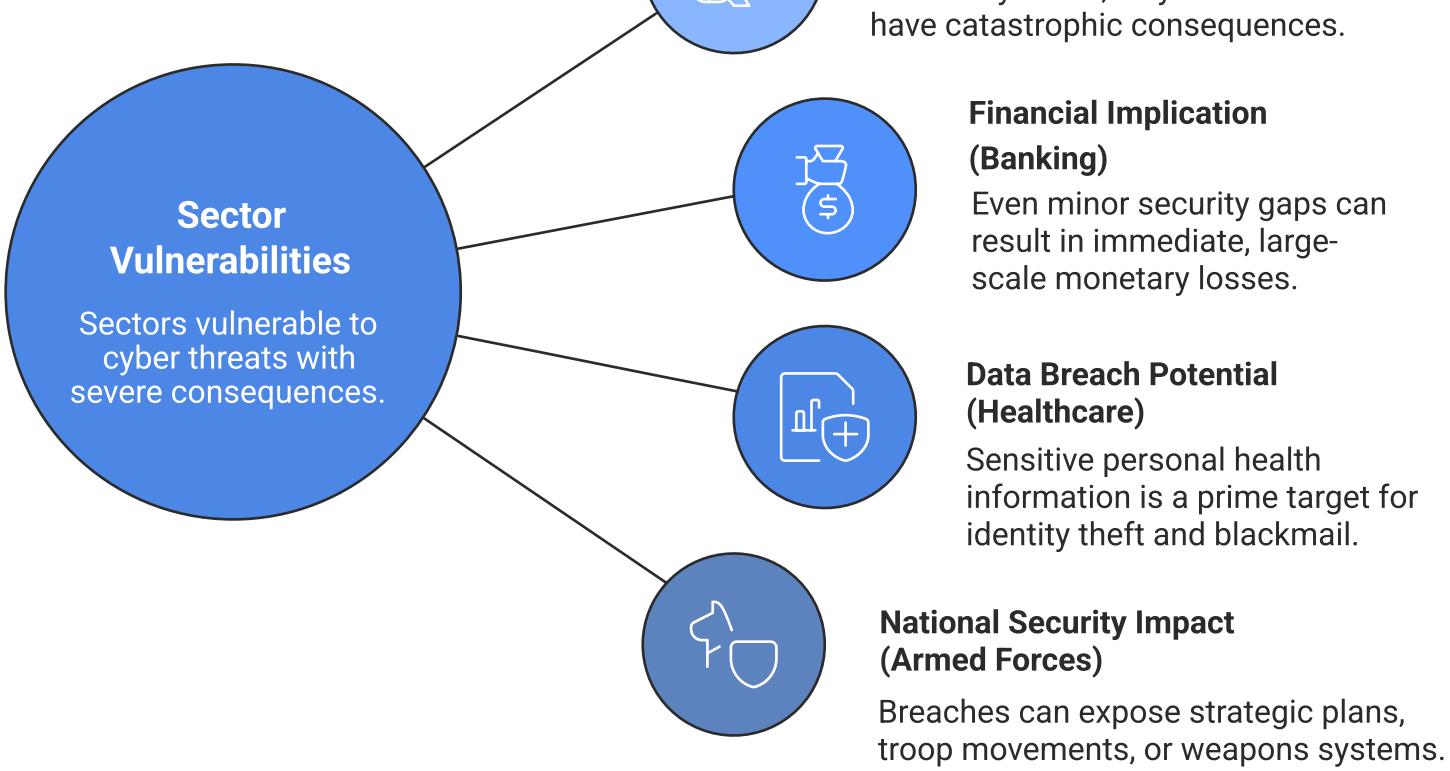
perimattic

Industries Most Vulnerable to AI Cyber Threats

No sector is immune to cyber risk, and the integration of AI into daily operations has amplified exposure across the board. AI is now accessible at the fingertips and it is capable of producing convincing phishing messages, generating deepfake videos, and automating attacks with minimal technical skill required.

According to the Cybercrime Trends 2025 report by SoSafe, **98% of security professionals believe the resilience gap is widening**, meaning the ability of organizations to defend themselves is not keeping pace with the speed at which threats are evolving. This imbalance leaves critical sectors such as essential infrastructure and small enterprises, increasingly exposed.

Factors That Make Certain Sectors More Vulnerable



perimattic

How Threat Actors Exploit Industry-Specific Vulnerabilities with AI

Some industries carry more inherent risk because of the value, sensitivity, or criticality of the data and systems they handle. AI compounds these vulnerabilities by enabling faster reconnaissance, more convincing deception, and more adaptive attack strategies.

Below are seven key industries where AI-enhanced cyber threats are particularly concerning.

Technology Sector

The technology sector is both an innovator and a prime target. In July 2025, Cisco disclosed a breach in which a cybercriminal used a voice phishing (vishing) scheme to convince a company representative to grant access to a third-party database. This led to the theft of basic profile information for Cisco.com users. The incident demonstrates that even sophisticated tech companies, with layered security architectures, can fall victim to AI-assisted social engineering. AI can generate hyper-realistic voices that mimic known employees or executives, eroding one of the last lines of human-based verification: voice recognition.

For technology companies, the stakes are high. They are custodians of not just their own data but also the tools and infrastructure used by countless other industries. An AI-assisted breach can lead to compromised products, malicious code insertion into widely deployed software, and theft of proprietary research. This can lead to consequences that ripple far beyond the breached organization.

Healthcare

Healthcare's vulnerability lies in the depth and sensitivity of the data it manages. The Tricare breach underscored this reality: backup tapes containing electronic health records were stolen from a vehicle during transportation. Security numbers, personal health data, clinical notes, lab results, and prescriptions-related information was breached.

AI magnifies these risks by making stolen healthcare data more actionable. Machine learning algorithms can quickly sort, match, and cross-reference information with other breached datasets to build complete identity profiles. In some cases, AI can predict future medical needs, enabling targeted fraud or blackmail. The healthcare sector's complex web of providers, insurers, and technology vendors offers multiple entry points for AI-assisted intrusion.

Energy and Utilities

AI-enhanced cyberattacks on the energy grids, oil rigs, nuclear or gas plants, could be aimed at shutting down critical infrastructure, manipulating control systems, or causing cascading failures across dependent industries.

AI's ability to simulate operational behaviour means attackers can develop precision exploits without triggering alerts during the reconnaissance phase. In addition, utilities often operate with legacy systems that are difficult to update, giving AI-assisted attackers a wider attack surface.

Banking and Financial Services

The banking sector is an attractive target for obvious reasons: direct access to capital and financial assets. NatWest, for example, faces an onslaught of 100 million cyberattacks every month. The sheer volume, driven by AI-powered automation, makes it increasingly challenging to distinguish genuine customer interactions from fraudulent ones.

AI enables fraudsters to execute highly personalized scams, impersonating bank officials, generating convincing transaction alerts, or producing fake account statements. The consequences extend beyond immediate monetary loss; a successful breach undermines trust in the institution, potentially leading to client attrition and regulatory penalties. Financial organizations must now contend not just with criminals but with algorithms designed to learn from every failed attempt and adapt accordingly.

Retail and E-Commerce

Retailers process vast amounts of payment information and consumer data. AI can be used to create convincing fake e-commerce websites, complete with cloned branding, product listings, and payment gateways. It can also generate fraudulent product reviews to mislead consumers or harm competitors.

Large retail platforms are also prime targets for credential stuffing attacks, where AI automates login attempts using stolen username-password pairs. The speed and scale of such attacks can overwhelm even advanced fraud detection systems, leading to widespread account takeovers and financial loss.

Air Travel and Aviation

The air travel sector operates on interconnected systems that manage everything from flight bookings to aircraft navigation. AI can be weaponized to manipulate scheduling systems, interfere with cargo and baggage logistics, or disrupt communication channels between pilots and ground control. Safety is the paramount concern; even a minor system compromise can endanger lives.

Threat actors can also use AI for targeted phishing campaigns against aviation staff, generating emails that precisely mimic official communications. The potential for ransomware attacks that shut down airline operations (stranding passengers worldwide) is high, with impacts felt immediately in public safety and customer trust.

Education and Research Institutions

Universities and research facilities are repositories of intellectual property such as groundbreaking medical research, defense-related technology, and more. AI can accelerate cyber espionage by scanning for unprotected databases, identifying valuable datasets, and exfiltrating them without detection.

AI-powered phishing campaigns targeting students, faculty, and administrative staff can also yield login credentials that provide attackers with deep access to institutional networks. The loss of research data can derail years of work and result in significant financial and reputational damage.

perimattic

Different Impacts Across Sectors

While all industries face similar categories of threat (data breaches, fraud, operational disruption) the impacts vary in severity and nature.

Financial Impact:

Banking and retail face direct monetary losses. Healthcare and education may incur significant regulatory fines and legal settlements.

Social Impact:

Breaches in healthcare, education, and public utilities can erode public trust and cause lasting harm to communities.

Brand Value:

Technology companies, airlines, and retailers risk long-term reputational damage that affects market position.

Trust:

Once a platform like YouTube or a bank suffers an AI-driven scam, user confidence can be difficult to rebuild.

Security:

In defense and critical infrastructure, breaches can have national security consequences, potentially endangering lives and destabilizing economies.

perimattic

Staying Ahead Of Miscreants: The Rise For Proactive AI-Powered Defense

The rapid acceleration of AI-enhanced cyber threats has made one truth unavoidable: **reactive defense is no longer enough**. By the time an attack is detected using traditional methods, it may already have inflicted significant damage. The only way forward is to use AI not just as a defensive shield but as an active, predictive force capable of anticipating and neutralizing threats before they manifest.

Cybersecurity today is an innovation race. The same AI capabilities that enable criminals to scale attacks can also be used by defenders to predict attack patterns, identify vulnerabilities in real time, and deploy countermeasures instantly. Industry leaders and governments alike are starting to embrace this mindset, investing in next-generation AI-driven security solutions.

Level 1: Private and Corporate Initiatives

Private organizations are often the first line of defense against cyberattacks, and many are developing proprietary AI solutions to protect their operations, customers, and partners.

Fujitsu's Multi-AI Agent Security Technology

Fujitsu has developed the world's first multi-AI agent security system, which brings together multiple AI agents, each specialized in a different aspect of cybersecurity. These agents combine deep knowledge of attack patterns with protective strategies, working collaboratively to identify vulnerabilities and counter new threats. Instead of relying on a single AI model, this approach creates a layered, adaptive defense that evolves alongside the threat landscape.

Accenture–Microsoft Partnership on GenAI Cybersecurity Tools

In another significant development, Accenture and Microsoft have joined forces to build generative AI-powered cybersecurity tools. The goal is to detect and block attacks and to help organizations tackle complex, evolving threats while optimizing their technology stack and lowering operational costs. This type of partnership reflects a growing recognition that advanced cyber defense requires the combined expertise of security specialists, cloud infrastructure providers, and AI innovators.

In both cases, the focus is on agility and on building systems that can learn, collaborate, and adapt at the speed of attack.

Level 2: Government-Level Efforts

Governments play a unique role in cybersecurity because their remit spans critical infrastructure, national defense, law enforcement, and public awareness. AI is now an essential component of national security strategies worldwide.

India's Cyber Crime Cells

In India, each state has established Cyber Crime Cells: specialist units staffed by trained investigators with expertise in cyber forensics and digital law enforcement. Working alongside local police, these cells investigate a wide range of cybercrimes, from AI-powered scams to attacks on critical systems. Their work is crucial in bridging the gap between technological threats and legal enforcement.

Japan's AI Integration in National Security Strategy

Japan's National Security Strategy, updated in December 2022, explicitly integrates AI into its cybersecurity and information warfare capabilities. The approach is multifaceted: AI is used for early threat detection, intelligence analysis, and defensive cyber operations. By embedding AI directly into its security framework, Japan aims to maintain technological superiority over potential adversaries, both in cyberspace and in broader defense contexts.

Government-level initiatives are particularly important because they establish the policy and legal frameworks within which both public and private sectors can coordinate AI-powered defense.

Level 3: Personal-Level Protection

Cybersecurity isn't just the responsibility of corporations and governments, individuals are targets too, and AI-driven tools can help protect them.

Truecaller's AI Call Scanner

The Swedish communication app Truecaller has rolled out an AI Call Scanner to combat the growing threat of AI-generated voice scams. This feature analyzes incoming calls in real time, distinguishing between genuine human voices and synthetic, AI-generated ones. Combined with caller ID, spam detection, and call blocking, it empowers users to protect themselves from fraud without having deep technical knowledge.

On a personal level, AI tools like these can help individuals secure their devices, identify suspicious content, and make better-informed decisions about online interactions. While personal defenses may not stop large-scale attacks, they can significantly reduce the success rate of scams and identity theft.

perimattic

Turning Innovation Into Advantage

The AI-driven cybersecurity arms race is not slowing down. Threat actors are innovating at speed, but so too are defenders. The key is to ensure that innovation on the defense side outpaces the evolution of attack methods. This requires:

- **Enhanced Vigilance:** Constant monitoring and readiness to adapt to new attack techniques.
- **Investment in Research and Development:** Building more advanced AI models for security, capable of predictive analysis and autonomous response.
- **Cross-Sector Collaboration:** Sharing intelligence and technology between private companies, governments, and individual users.
- **A Balanced View of AI:** Recognizing AI not just as a risk but as a powerful ally in securing digital ecosystems.

The future of cybersecurity will be defined by the side that can innovate faster. By harnessing AI as both shield and sentinel, defenders can not only respond to threats but anticipate and neutralize them before they cause harm.

perimattic

Custom, Secure And Scalable Digital Solutions For Your Business

Perimattic builds digital solutions that change how your business works from automation to cybersecurity and everything in between.

Contact Us Now

US +1 (442) 319-7124

UK +44 (20) 3769 0051

IND +91 92142 66896

sales@perimattic.com

www.perimattic.com

Follow Us