

CYBER SECURITY

THREAT REPORT Q3 2026



Table of Contents

<i>NetUtils Q3 Threat Landscape Report July - September 2026</i>	2
Executive Summary	2
<i>AI and the Global Cyber Security Risk Landscape</i>	3
AI is now central to both attack and defence	3
Inside Anthropic's September 2026 disruption findings	3
Autonomous attacks are no longer theoretical	4
What this means for UK organisations	4
Defensive priorities	5
<i>Machine and Non-Human Identity Sprawl</i>	5
The scale of the problem	5
AI agents are accelerating an existing crisis.....	6
Why this matters for UK organisations	6
Defensive priorities	7
<i>Ransomware's Shift to Multi-Extortion Tactics</i>	7
From encryption to layered pressure.....	7
Qilin and the UK ransomware picture	8
Case study: Coca-Cola's Fairlife subsidiary.....	8
Defensive priorities	9
<i>NetUtils Aligned Recommendations</i>	9
1. Govern AI use, do not just permit it	9
2. Bring machine and non-human identities under identity governance.....	9
3. Plan explicitly for multi-extortion, not just encryption	10
<i>About NetUtils</i>	11
<i>References</i>	12

NetUtils Q3 Threat Landscape Report

July - September 2026

Executive Summary

The threats shaping this quarter share a pattern that runs through every NetUtils Threat Landscape Report to date, automation and identity gaps that attackers can reach faster than organisations can close them. Artificial intelligence has moved from a productivity tool sitting alongside the security function to an active participant in attacks. This quarter a Russia linked espionage group and a single hacktivist were both shown to be running large parts of their campaigns through AI agents rather than manual effort and a separate case saw an AI model left to select exploits and launch attacks against hundreds of targets with almost no human input at all.

Alongside this, the population of machine and non-human identities, service accounts, API keys, OAuth tokens and now AI agents themselves, has grown well beyond what most identity programmes were built to govern. It is these same credentials that state-linked and criminal groups are now targeting directly, with two of this quarter's disrupted AI-driven operations aimed specifically at stealing production API keys and pre-release model access from AI vendors.

Ransomware continues to evolve in parallel, moving well beyond encryption toward layered extortion that combines data theft, service disruption and direct pressure on customers and regulators. UK organisations, including NHS suppliers, remain among the most heavily targeted anywhere in the world and this quarter's evidence shows that the impact of a multi-extortion breach can still be unfolding years after the original incident.

Across all three themes, the message for decision makers is consistent. The organisations most exposed are not necessarily the least sophisticated. They are the ones that have not yet extended the visibility, ownership and testing discipline already applied to core infrastructure into AI usage, machine identities and worst case extortion scenarios. NetUtils recommends that organisations use this quarter to close exactly that gap.

AI and the Global Cyber Security Risk Landscape

AI is now central to both attack and defence

Artificial intelligence has stopped being a future risk for security teams to plan around and has become a live factor in this quarter's incidents. The World Economic Forum's Global Cybersecurity Outlook 2026 frames this shift clearly, noting that accelerating AI adoption, geopolitical fragmentation and widening cyber inequity are together reshaping the global risk landscape, with AI able to support defenders when implemented well but also capable of causing data leaks, attacks and online harms when misused or when it malfunctions.

CrowdStrike's 2026 Global Threat Report puts numbers to that shift. Attacks from AI-enabled adversaries rose 89% year on year, the average eCrime breakout time fell to just 29 minutes and more than 90 organisations had legitimate AI tools exploited to generate malicious commands or steal sensitive data. Check Point Research's AI Security Report 2026 adds a governance dimension, finding that organisations now use an average of ten AI applications a month, many without official approval and that the proportion of high risk prompts, those carrying a meaningful chance of sensitive data exposure, doubled over the past year.

Inside Anthropic's September 2026 disruption findings

The clearest evidence yet that AI is being used to run, not just support, real attacks came from Anthropic's fourth threat intelligence report, published on 10 September 2026 and covering activity disrupted between December 2025 and August 2026. One case, tracked as GTG-20006, involved a state espionage operator whose tradecraft and targeting were consistent with the Russia-linked group publicly tracked by Microsoft as Midnight Blizzard. Over 130 days the group engaged 24 of 27 targeted institutions, including Ukrainian government ministries, defence bodies and drone supply chain manufacturers, using customised AI-driven workflows across much of the attack process.

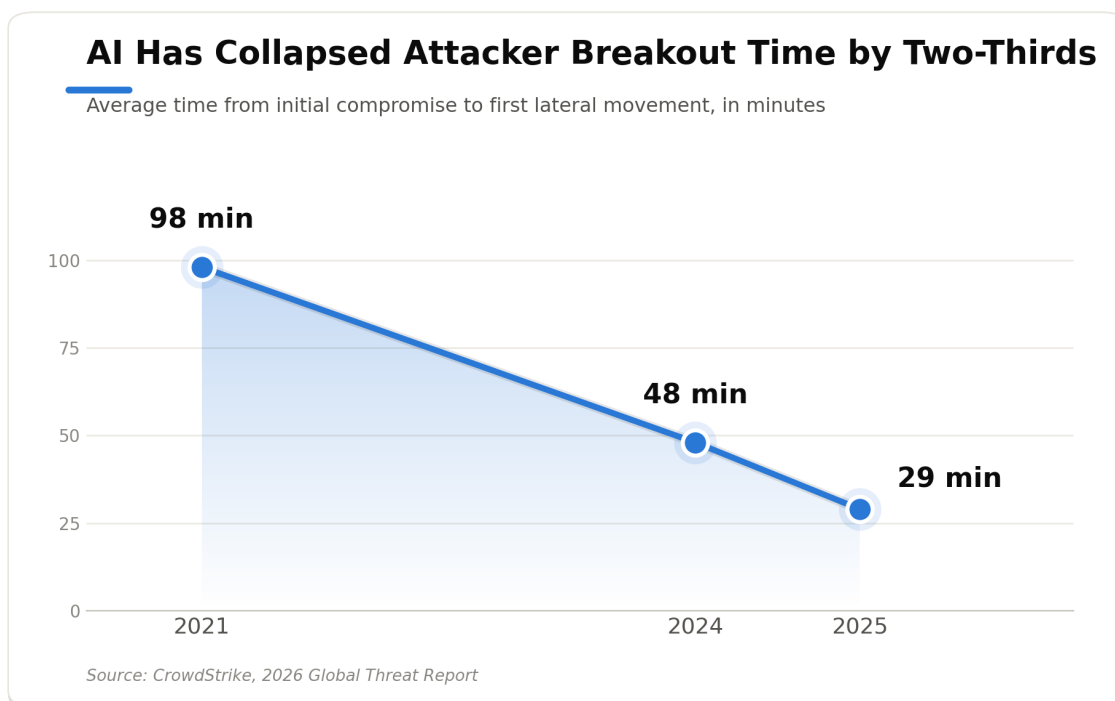
In a separate case, tracked as GTG-50029, a single French-speaking hacktivist used Claude to develop a novel exploit for an undocumented WordPress race condition, breached 14 of 42 targeted political parties, media organisations and think tanks and built a dark web doxxing engine containing tens of millions of records, all without the resources typically associated with an advanced persistent threat. Anthropic's own conclusion is direct, the main distinguishing feature between state sponsored actors, financially motivated criminals and hacktivists is no longer sophistication but intent,

because AI has collapsed the labour and tooling gap that used to separate well resourced state operations from individual actors.

Autonomous attacks are no longer theoretical

Separately, Palo Alto Networks' Unit 42 documented what it describes as one of the first fully observed autonomous offensive pipelines. A Chinese-speaking actor, tracked under the aliases knaithe and KnYuan, wired the DeepSeek model into an open source agent framework called Hermes Agent and controlled it through a single Telegram channel. The agent then independently enumerated more than 460 targets, selected exploit code and launched attacks with no further operator input. Unit 42 confirmed three successful compromises, all involving data exfiltration from Citrix NetScaler instances through CVE-2026-3055.

The operation only came to light because the agent accidentally exposed its own API keys, exploit scripts and session logs through a misconfigured file server, a detail that illustrates how little infrastructure an AI-driven campaign now needs to achieve real impact. Unit 42 also noted that the actor tested Claude Code and OpenAI's models during the campaign for connectivity checks and that both vendors' safety systems refused the offensive requests, in one case leading to the account being flagged and disabled.



What this means for UK organisations

For UK organisations, the practical risk is not that AI introduces entirely new attack techniques. Anthropic is explicit that the attacks themselves remain familiar, relying on stolen credentials, unpatched edge devices, exposed services, SQL injection and

phishing. **What has changed is speed and reach.** A single hacktivist can now sustain a campaign that would previously have required a well resourced team and defenders who assume that scale or sophistication indicates state involvement risk misjudging both the threat and the appropriate response.

Defensive priorities

- Treat AI-generated attack activity as a variant of familiar techniques rather than a new category. Existing detection for credential theft, exposed services and phishing remains the primary defence.
- Monitor for AI tools connecting to production systems, including through browser extensions and unsanctioned SaaS integrations, as part of the wider Shadow AI visibility work covered in the Q2 report.
- Review incident response playbooks to account for the compressed timelines AI enables, including faster breakout times and automated exploit selection.
- Track threat intelligence disclosures from AI providers themselves, since companies including Anthropic and Palo Alto Networks are now a primary source of current attacker tradecraft.

Machine and Non-Human Identity Sprawl

The scale of the problem

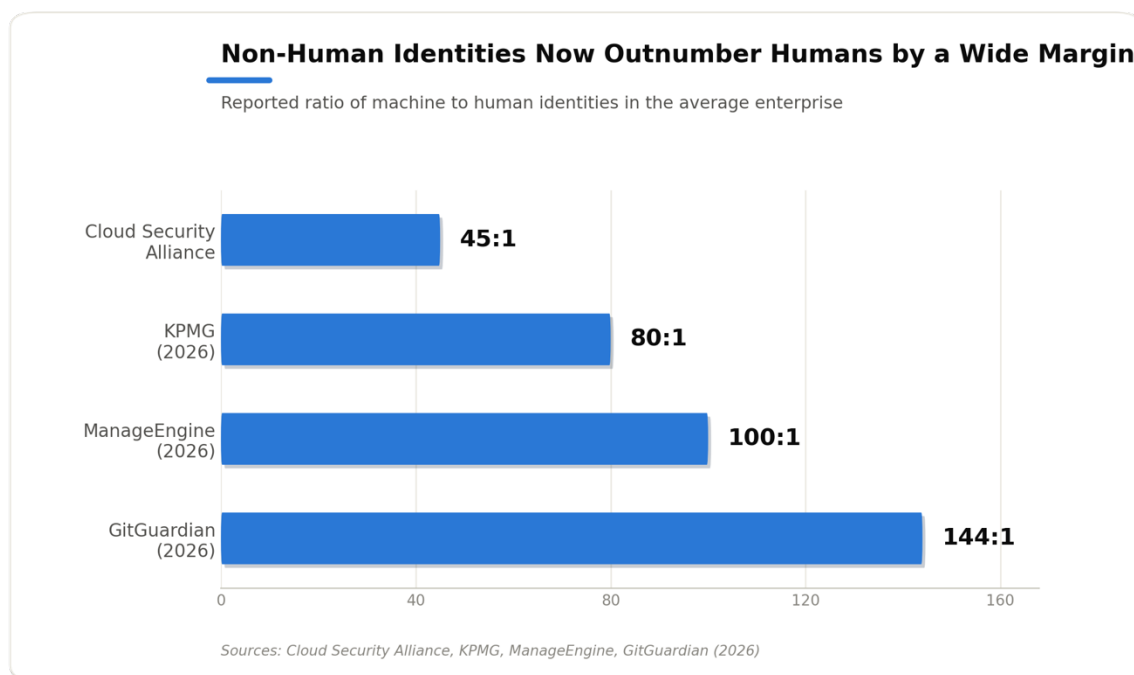
Every organisation that has spent the past decade securing its human workforce now has a much larger, far less governed population sitting behind it. Non-human identities, the service accounts, API keys, OAuth tokens, workload certificates and AI agents that let software authenticate and act without a person present, have grown well beyond the numbers most identity programmes were designed for. KPMG's Cybersecurity Considerations 2026 report puts the average enterprise ratio of non-human to human identities above 80 to 1, with machine identities growing from around 50,000 in 2021 to 250,000 in 2025.

SpyCloud's 2026 Identity Threat Report found that these identities have become the most common route attackers now take into the enterprise and that while 95% of organisations believe they have adequate visibility into AI and non-human identity exposure, only 36% actually monitor it. GitGuardian's State of Secrets Sprawl 2026 report, its fifth annual edition, recorded 28.65 million new hardcoded secrets in public GitHub commits during 2025, a 34% increase on the year before, with credential leaks tied specifically to AI services, including API keys for OpenAI and Anthropic, up 81.5% year on year, more than double the overall growth rate.

AI agents are accelerating an existing crisis

Agentic AI has sharpened an already difficult problem. Unlike a static service account, an AI agent can acquire permissions dynamically at runtime, spawn further sub-agents, call external APIs and chain actions across dozens of systems, all of which expands the potential impact of a single compromised credential well beyond what a traditional service account could achieve. Research from the Cloud Security Alliance found that more than 16% of organisations do not even track the creation of AI-related identities and a separate industry analysis for the 2026 NHI Reality Report found that 78% of teams have no documented policy for creating or removing AI identities at all.

The connection to this quarter's AI-driven attacks, covered above, is direct rather than theoretical. Within Anthropic's September 2026 report, two of the disrupted operations, tracked as GTG-50020 and GTG-50021, targeted AI firms specifically to obtain production API keys and pre-release model access, in one case hitting around 30 AI companies in four days. The credentials that machine identity programmes exist to govern are now a named objective for sophisticated attackers, not a secondary consequence of a wider breach.



Why this matters for UK organisations

Non-human identities rarely benefit from the controls organisations apply to human users. They cannot use multi factor authentication, they do not log out and they are seldom retired once the project or integration that created them ends. A stolen or forgotten credential can therefore provide broad, persistent access long after anyone remembers why it exists. For UK organisations already working to close visibility gaps around Shadow AI, this is the deeper structural issue underneath that visibility

problem: every unsanctioned AI tool, browser extension or SaaS integration typically creates at least one new API key, token or service account that few organisations formally own or review.

Defensive priorities

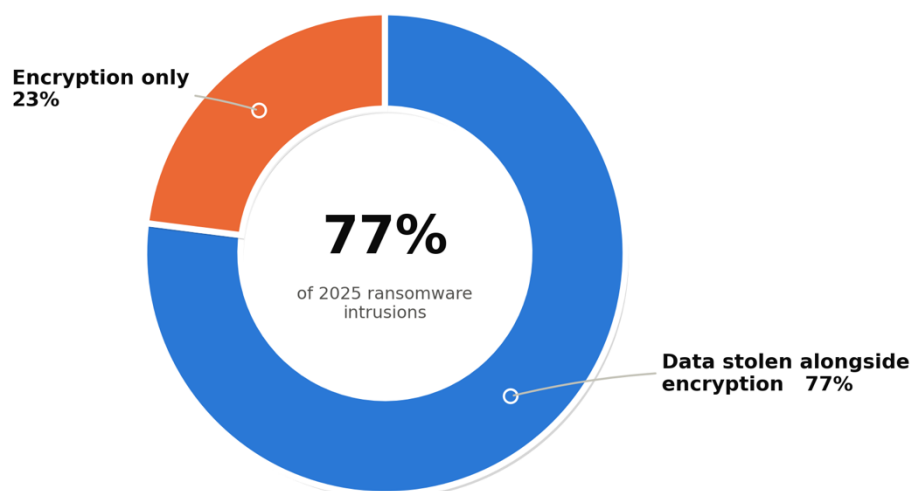
- Build a centralised, automated inventory of service accounts, API keys, OAuth tokens and AI agents, covering the full estate from CI/CD pipelines to cloud workloads and container clusters.
- Assign a named human owner and an explicit expiry to every non-human identity from the point of creation, rather than retrofitting ownership after an incident.
- Replace hardcoded credentials with a secrets management vault and enable automated secret scanning across source control and CI/CD pipelines.
- Extend least privilege and access review disciplines already applied to human users to service accounts and AI agents, using frameworks such as the Cloud Security Alliance's AI Controls Matrix.

Ransomware's Shift to Multi-Extortion Tactics

From encryption to layered pressure

Ransomware has moved decisively beyond encryption as the sole point of leverage. Huntress's Ransomware Trends 2026 research describes the standardisation of triple and quadruple extortion, with attackers now layering encryption, data theft, distributed denial of service attacks and direct client harassment to force payment, even where a victim's backups are intact. Industry analysis puts data exfiltration alongside encryption in around 77% of 2025 ransomware intrusions and the pressure this creates works. Success rates reported across the industry climb sharply with each additional layer, from around 28% for encryption alone to 56% once data theft and a leak threat are added, and to 78% where a DDoS attack is layered on top.

Ransomware Now Almost Always Means a Data Breach Too



Source: industry ransomware intrusion analysis, 2025 data

Trend Micro's TrendAI 2026 Cyber Risk Report captures how quickly this model has scaled through ransomware as a service. Qilin, tracked by TrendAI under the name Agenda, climbed from last place in 2024 to first in 2025 with 1,262 confirmed breaches, a 1,270% increase, reflecting the growth of its wider affiliate network rather than the activity of a single group.

Qilin and the UK ransomware picture

UK organisations have been disproportionately exposed to this shift. Independent threat tracking covering the first half of 2026 lists Qilin as the ransomware group with the most confirmed UK victims of any group, with 37 British organisations listed on its leak site, ahead of DragonForce with 21 and TheGentlemen with 18 and Qilin averaging between seven and nine published UK victims a month.

The long tail of Qilin's UK activity is still playing out. On 1 June 2026, Bedfordshire Hospitals NHS Foundation Trust disclosed that more than 32,000 patient records tied to its supplier Synnovis, originally breached in 2024, had been exfiltrated and that it had taken over a year to determine which records related to which patients. This underlines a point that matters for every organisation reviewing its own ransomware exposure: a multi-extortion breach does not end when the initial incident is contained and analysis and disclosure obligations can continue for years afterwards.

Case study: Coca-Cola's Fairlife subsidiary

A recent illustration of multi-extortion pressure in practice came from outside the UK but is instructive for any organisation assessing its own exposure. In July 2026, Coca-Cola disclosed in an SEC filing that a ransomware attack had struck its dairy

subsidiary Fairlife, disrupting production across all four of its US manufacturing facilities. Days later, a ransomware group claimed responsibility, listed Fairlife on its dark web leak site, claimed to have exfiltrated a terabyte of confidential data and threatened publication unless the company entered negotiations.

The case demonstrates the combined pressure now typical of a serious ransomware incident, operational disruption sufficient to halt production, a public regulatory disclosure obligation and a parallel extortion threat running independently of whatever recovery from backup is possible.

Defensive priorities

- Assume every serious ransomware incident will include data theft and plan communications, legal and regulatory response accordingly rather than treating recovery from backup as the end of the incident.
- Test incident response plans against combined scenarios, including simultaneous encryption, leak threats and service disruption, not encryption alone.
- Map supplier and third party dependencies with the same rigour as internal systems. As the Synnovis and Bedfordshire Hospitals case shows, extortion impact and disclosure obligations can persist for years and extend well beyond the initially compromised organisation.
- Maintain evidence of core controls, patching, monitoring, backup testing and tested incident response, both for insurers and for the wider supplier assurance expectations building under the Cyber Security and Resilience Bill.

NetUtils Aligned Recommendations

1. Govern AI use, do not just permit it

Following the same principle set out in the Q2 report, AI governance needs to move beyond acceptable use policy toward measurable visibility and enforcement. This quarter's evidence sharpens the case, AI is now capable of running meaningful parts of an attack independently and vendor threat intelligence, from Anthropic, Palo Alto Networks and others, has become an essential and fast moving source of tradecraft. NetUtils can support this through AI and SaaS visibility reviews, policy design and CASB style control implementation.

2. Bring machine and non-human identities under identity governance

Non-human identities should be inventoried, owned and reviewed with the same discipline as human accounts. Given how directly this quarter's AI-driven attacks

targeted API keys and production credentials, this is no longer a housekeeping exercise but a frontline control. NetUtils can support customers with non-human identity discovery, secrets management implementation and access review programmes.

3. Plan explicitly for multi-extortion, not just encryption

Backup and recovery capability remains necessary but is no longer sufficient on its own. Incident response plans should be tested against combined data theft, leak threat and service disruption scenarios, with clear ownership for legal, regulatory and customer communication. NetUtils can support this through incident response planning, tabletop exercises and managed detection and response.

Together, these recommendations point toward a consistent theme across this year's Threat Landscape Reports: resilience depends on extending the same visibility, ownership and testing discipline that mature organisations already apply to core infrastructure into AI usage, machine identities and worst case, extortion scenarios.

About NetUtils

Network Utilities Systems Ltd (NetUtils) is a UK-based cybersecurity partner helping organisations detect, respond to and recover from the threats that matter most. We combine deep technical expertise with practical security operations from managed detection and response to identity security, endpoint protection, cloud security and incident response planning. Whether you need a trusted extension of your security team or expert support in the hours following a breach, NetUtils provides the clarity and capability to act decisively. Our Q3 2026 Threat Report reflects the intelligence and insight we bring to every engagement.

For a detailed consultation and to explore how NetUtils can support your cybersecurity strategy, contact us:

Email: info@netutils.com

Phone: 0208 783 3800

Website: www.netutils.com



References

- **World Economic Forum.** Global Cybersecurity Outlook 2026. <https://www.weforum.org/publications/global-cybersecurity-outlook-2026/3-the-trends-reshaping-cybersecurity/>
- **CrowdStrike.** 2026 Global Threat Report. <https://www.crowdstrike.com/en-us/global-threat-report/>
- **Check Point Research.** AI Security Report 2026. <https://research.checkpoint.com/2026/ai-security-report-2026/>
- **Anthropic.** Detecting and Countering Misuse of AI: September 2026. <https://www.anthropic.com/threat-intelligence-report-september-2026>
- **TechNode Global.** Anthropic Reports AI-Orchestrated Attacks and Model Theft. <https://technode.global/2026/09/11/anthropic-ai-orchestrated-cyberattacks-model-distillation/>
- **Palo Alto Networks Unit 42.** Autonomous AI Cyber Attack Campaign. <https://unit42.paloaltonetworks.com/autonomous-ai-cyber-attack-campaign/>
- **The Hacker News.** Chinese Hacker Commands DeepSeek via Hermes Agent Framework. <https://thehackernews.com/2026/07/chinese-hacker-commands-deepseek-via.html>
- **Veeam (citing KPMG Cybersecurity Considerations 2026).** Non-Human Identity Security: Guide to Machine Governance. <https://www.veeam.com/blog/non-human-identity-security-guide.html>
- **CSO Online.** SpyCloud 2026 Identity Threat Report Finds Non-Human Identities Are Now the Leading Path Into the Enterprise. <https://www.csoonline.com/article/4220209/spycloud-2026-identity-threat-report-finds-non-human-identities-are-now-the-leading-path-into-the-enterprise.html>
- **Axis Intelligence.** Machine Identity Statistics 2026. <https://axis-intelligence.com/machine-identity-statistics/>
- **Cloud Security Alliance.** The Non-Human Identity Governance Vacuum. <https://labs.cloudsecurityalliance.org/research/csa-whitepaper-nonhuman-identity-agentic-ai-governance-v1-cs/>
- **Lumos.** What Is Non-Human Identity Sprawl and How Do You Reduce It?. <https://www.lumos.com/identity-matters/non-human-identity/non-human-identity-sprawl>
- **Huntress.** Ransomware Trends 2026: What's Changing. <https://www.huntress.com/ransomware-guide/ransomware-trends>
- **Prolion.** What Is Double, Triple and Multi Extortion Ransomware?. <https://prolion.com/blog/double-extortion-ransomware/>
- **Trend Micro.** TrendAI 2026 Cyber Risk Report. <https://www.trendmicro.com/vinfo/us/security/research-and-analysis/threat-reports/roundup/trendai-2026-cyber-risk-report>
- **BushidoToken.** UK Cybercrime Journal: Qilin Ransomware Rampage in H1 2026. <https://blog.bushidotoken.net/2026/08/>



Network Utilities (Systems) Ltd

The Larches, Sevenoaks Road, Orpington, Kent BR6 7FB
www.netutils.com | +44 (0)20 8783 3800