THREAT INTELLIGENCE

LRQA
Your Risk Management Advantage

2025 YEAR IN REVIEW

Threat intelligence insights from 2025 and expectations for the year ahead

CONTENTS

FOREWORD

Every year we investigate hundreds of incidents and every year we see the same thing. Cyber criminals are not masterminds. They are opportunists. They want quick wins with minimal effort and minimal risk. When something works they copy it. When it stops working they tweak it just enough to slip past the latest defence and carry on.

In that sense the criminal ecosystem is not that different from the legitimate one. Good ideas get borrowed. Bad ideas get dropped. No one is reinventing the wheel if they can keep using the old one.

The technology landscape keeps shifting, sometimes dramatically. Yet the nature of cybercrime barely changes.

Attackers go after the same weaknesses because those weaknesses keep paying off. Identity processes that do not quite join up. Unpatched or end-of-life devices. Cloud services left open by accident. Overworked support teams that are pushed to action requests quickly rather than safely.

The trends in this report reflect that reality. We have highlighted the campaigns and intrusion methods that shaped 2025 and shown how familiar techniques have moved into new environments. It is important to understand what has changed, but it matters just as much to recognise what has stayed the same.

The lesson for defenders is just as familiar: Do the basics well. Layer your controls. Maintain awareness of how threats evolve. Cyber security is rarely about chasing the latest thing. Most of the time it is about making sure yesterday's problems don't become tomorrow's breach.

Everything changes, but everything stays the same.

Stephen Robinson
Lead Threat Intelligence Researcher | LRQA

PURPOSE

This document will explore the threat landscape of 2025 through the critical incidents and campaigns of the year.

The analysis will look for causes and trends and attempt to extrapolate predictions for the year ahead.

TRENDS

1.0 NETWORK INFRASTRUCTURE COMPROMISE

Edge network infrastructure is an excellent target for attackers; It is intentionally exposed to the internet, it offers critical services that provide high value access to attackers, and it typically provides minimal access to the underlying operating system's internal functionality. As such, it often cannot be closely monitored, and it is very unlikely to support local installation of EDR software.

Not only is it an excellent dwell point for targeted, stealthy, high dwell-time attackers who wish to avoid security monitoring tools; it is also an easily accessible target for mass exploitation by high volume, indiscriminate attackers that seek access for rapid detonation or smash and grab attacks. This has been the case for some time, but this trend has continued at pace during 2025. It is generally acknowledged that total prevention of compromise is not always possible, both because zero-days exist, and because differentiating malicious or benign behaviour often requires behavioural context. However, network infrastructure appears to have a zero-day problem while also not running EDR software, making it perfect for attackers, and a nightmare for defenders.

That zero-day problem is not caused by complex or unusual vulnerabilities. Time and again zero-days in network infrastructure have been caused by things that really should be solved problems by now: Path traversal, buffer overflows, or just running all functionality in a single binary as root. Basic protections such as Data Execution Prevention or Address Space Layout Randomisation have been missing. Bigger vendors have tended to do better on this front, but the words "Watchtowr blogpost" still cause concern for every network vendor.

Unfortunately, network infrastructure devices do not just have a zero-day problem, they also have an n-day problem.

While a zero-day is a vulnerability for which there is no patch, meaning that administrators have "0 days" to patch before an exploit becomes available, an n-day is a vulnerability for which a patch is available, and where administrators have "n" number of days to patch before an exploit becomes available.

In 2025 zero-day and n-day exploitation of vulnerable firewalls, VPNs, and routers has been discovered affecting Cisco, Fortinet, WatchGuard, Palo Alto, Juniper and Ivanti, to name a few. Hundreds of thousands of Internet exposed devices from these vendors have been found to have actively exploited zero-days. Exploitation has been observed by both state-sponsored and financially motivated actors. While zero-days are typically the reserve of actors with the technical ability and financial backing to discover and weaponise them, the slow pace of infrastructure patching means that there are many n-day vulnerabilities with pre-written, publicly available exploits that are still viable for use by low-capability actors.

By this, we mean there are lots of unpatched devices on the internet.

2.0 CLOUD COMPROMISE AND EXTORTION

Modern enterprise systems and technology environments run in a hybrid-cloud model by default. Reinventing the wheel is not cost effective. As such, many solutions which organisations require are provided as cloud services, and most of the time it is a sensible business decision to use them.

However, cloud SaaS solutions create an attack surface problem. Traditionally, your HR systems, payroll, file storage, and development environments would only be accessible from inside your network. You controlled your perimeter, so you decided who had access and from where. Now, however, cloud services are often accessible to anyone with an internet connection. On top of this, cloud services intentionally abstract away their internal workings. They want to be seen as simple, and with a shallow learning curve, so they will hide any complexity of configuration or operation. However, the effort of making a complex system appear simple almost always results in the system itself becoming even more complex. These are complex software systems running on top of further complex software and hardware systems, i.e. the cloud. This hidden complexity provides a multitude of opportunities for actors who understand a system better than its users.

During 2025, actors typically referred to as Shiny Hunters, who coordinate via the diffuse chat network known as The Com, have been executing a long running campaign of attacks targeting Salesforce cloud instances.

These attackers often socially engineer their victims through voice phishing calls pretending to be IT support. The actors then persuade the victim to link the attacker's malicious data loader app to the victim organisation's Salesforce instance, an operation which merely requires a logged-in user with sufficient permissions to enter an 8-digit number into a Salesforce config screen. After that, the attacker can simply exfiltrate the data from Salesforce cloud through their malicious data loader integration app. In these cases, the actors have been observed using this access and data to then move laterally to other cloud platforms then repeat the process there.

GitHub has been ripe for attacks by actors who understand the complexity hidden within the platform. Sometimes those attacks are as simple as just looking for (and finding) credentials accidentally uploaded to GitHub repositories by developers, but attackers are also abusing GitHub workflow actions to compromise repositories and steal credentials /secrets. In some cases these compromises have relied on something as simple as repository maintainers using the GitHub Actions pull_request_target trigger instead of the pull_request trigger, inadvertently giving malicious external contributors the ability to execute code with elevated privileges and modify the victim repository.

3.0 CLOUD SUPPLY CHAIN

Supply chains are force multipliers for attackers, amplifying a single hack into a multitude of compromises. They offer an easier route into a hard target, as even a well-defended organisation's supply chain will have a weaker link where the defences are simply not as strong or the organisation simply has no visibility or control.

Cloud "as a Service" environments are supply chains that create explicit trusts between customers and service providers. Customers must trust the service provider and any additional services the provider uses. In fact, customers may well be using the cloud service to offer their own cloud services to downstream customers. Attackers are keenly aware of this and intentionally target *aaS providers as stepping stones to access their customer chains.

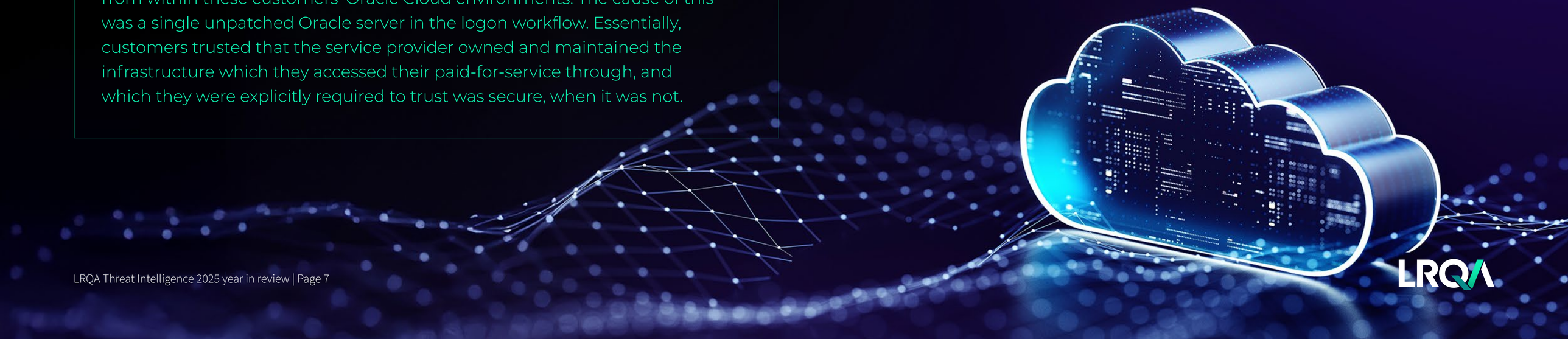
In 2025 Q2, a server used in the provisioning of Oracle Cloud was compromised, allowing the attacker to harvest up to 140,000 credentials from the logon process for Oracle Cloud.

The attacker used these credentials to access large amounts of sensitive data from within these customers' Oracle Cloud environments. The cause of this was a single unpatched Oracle server in the logon workflow. Essentially, customers trusted that the service provider owned and maintained the infrastructure which they accessed their paid-for-service through, and which they were explicitly required to trust was secure, when it was not.

In Q3, it was disclosed that the SaaS provider SalesLoft had been compromised. In Q2 the attackers had gained access to SalesLoft's GitHub repositories from which they were able to harvest additional credentials and move laterally into its AWS environments. From AWS they were able to move into the customer tenant environments for the SalesDrift LLM support agent. Many of the tier-1 customers of SalesLoft were themselves service providers, such as Workday, Workiva, CyberArk, Proofpoint, Cloudflare and Palo Alto.

Because SalesDrift is intended to provide customer support services, it integrates with customer support CRM instances, which by definition will contain data from customer support interactions with further downstream customers. Furthermore, investigators from Google found that the attackers were specifically running searches for additional access tokens and credentials for further downstream customers, in order to compromise further victims. The attackers in this case were identified as Shiny Hunters, the same group that was targeting Salesforce instances more widely during 2025 (see the section on Cloud Compromise and Extortion).

It appears that they worked out that a supply chain compromise against a Salesforce integration provider would be a far more effective method of compromise than simply compromising individual victims one by one. The success of this supply chain attack means it is likely that they will continue this method of targeting.



4.0 TCS, RETAIL, AND THE OUTSOURCING SUPPLY CHAIN

Scattered Spider, a fluid, ill-defined group who organise via The Com, have been behind a number of high-profile, extremely disruptive attacks this year.

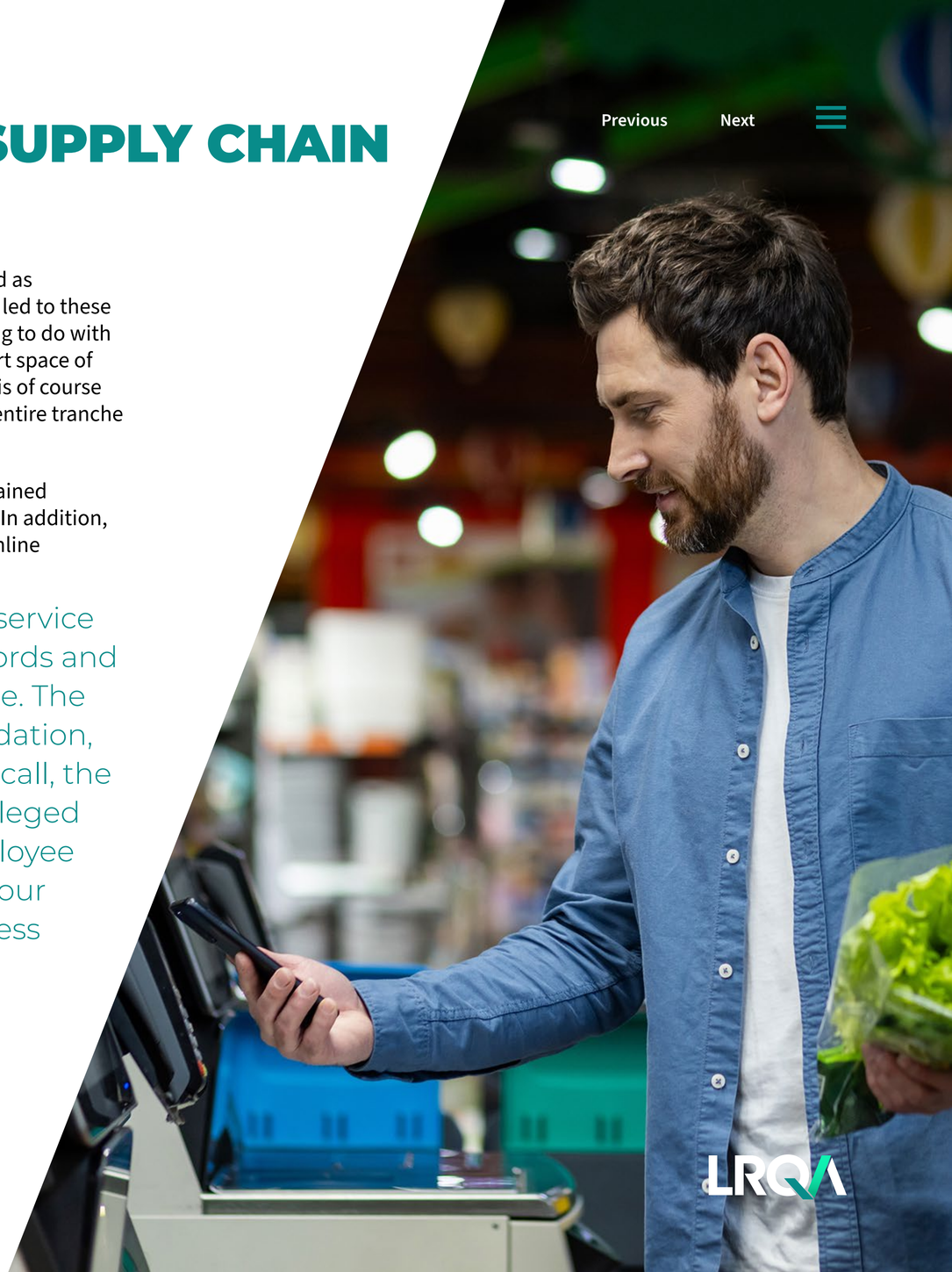
An unusual hallmark of their more well-known activity is that it is often clustered in thematically and geographically linked campaigns, with multiple targets hit in less than a week. These attack campaigns began by targeting major UK retailers, causing over £500 million in impact in total to Marks and Spencer and Co-op. After this they targeted US insurance companies Erie Insurance, AFLAC and PHLY, and then the airlines WestJet, Qantas and Hawaii Air. Then, later in the year they compromised the UK automotive company Jaguar Land Rover, causing an estimated £2 billion of impact. These are most definitely not Scattered Spider's only victims this year, but all of these victims are known to be linked by one thing: they were customers of Tata Consultancy Services (TCS).

TCS are an extremely large and successful India-based Managed Service and Managed Security Service Provider (MSP and MSSP) who have contracts with enterprise customers around the globe. They are renowned for being highly competitive with the price of their services and solutions, something they often manage to achieve by outsourcing as many functions as possible to India, and/or by importing workers from India to the locality of the customer on temporary contracts and visas. Thanks to public reporting and social media posts, we can see that all of the above listed victims of Scattered Spider were customers of TCS in at least some capacity with both Co-op and Marks and Spencer relying on TCS for their password reset service desks.

It is important to note that TCS have not been described as responsible for any failings which may or may not have led to these breaches, and indeed they may simply have had nothing to do with all of these customer breaches which occurred in a short space of time. When you have as many customers as TCS has, it is of course entirely possible that random chance could lead to an entire tranche of victims being your customers.

Marks and Spencer have stated that Scattered Spider gained access by compromising a third party run service desk. In addition, an unnamed and unverified source claimed in public online discussions about the UK retail compromises that

> "in 3 of 4 calls [to the TCS provided service desk], the service desk reset passwords and re-enrolled MFA with zero resistance. The caller simply gave a name – No validation, no callback, no check". "On the 4th call, the attacker requested access to a privileged group. The agent asked for an employee ID. The ID given didn't even match our company's format; and yet, the access was granted anyway".



4.0 TCS AND THE WIDER OUTSOURCING SUPPLY CHAIN

TCS have been very clear that none of their systems or accounts were compromised, which is almost certainly true. But if an outsourced service desk can be social engineered to hand out access to privileged accounts, there is no need to compromise the systems and accounts of the service desk itself.

This is of course not to say that TCS themselves are the problem, or that this recent rash of attacks is anything new.

In 2023, US manufacturing giant Clorox suffered a hugely disruptive cyber-attack which was at the time believed to be performed by Scattered Spider. Clorox suffered direct recovery costs of $49 million, and an additional $330 million in indirect costs due to the compromise. We know this because in 2025 they filed a lawsuit against their IT service desk provider, outsourcing firm Cognizant. This lawsuit details that the Cognizant run helpdesk reset passwords and MFA credentials multiple times, without verifying the identity of callers or notifying the account holder or account holder's line manager via email, as per their stated operating procedure. While Cognizant have not denied that the described calls and password resets took place, they have stated that they believe they were only responsible for operating the single, narrowly scoped help desk services for Clorox, not their wider cybersecurity.

Outsourcing of services to international service providers is common, particularly services which are labour intensive, such as call centres. Many organisations provide these kinds of services, and they attract customers by providing the minimum viable product at a cheaper price than the customer could provide it themselves.

Unfortunately, as with so many organisational and technological processes, people often forget to price in security. Service desks are gatekeepers, they are a point where organisational process and technological controls meet. As such, it may simply be unrealistic to expect somebody from a third-party organisation, located in a different country, who probably speaks a different language whenever they are not on a support call to understand and correctly implement your organisation's processes and technical controls, while also expecting them to perform that task as rapidly and cheaply as possible. Is a race to the bottom on price and acceptable service levels merely an expected result of the post-pandemic financial climate? Is the fact that many organisations are seeking the cheapest, most skeletal of outsourced organisational processes without the meat of cybersecurity on their bones an unintentional trend? And will the shocking disruption and costs that have resulted from these helpdesk engineering attacks cause organisations to re-assess their perceptions of risk?

Online commentary on Scattered Spider's UK retail compromises this year has pointed out that the operators behind these attacks are not necessarily technical geniuses, hacking through the latest and greatest products that the cyber security industry has to offer. Instead, they are simply sidestepping security controls by arming themselves with an understanding of organisational processes, the ability to operate a keypad-based phone menu, and the singular superpower of having an English accent while asking for a password reset.

Security is not simply technology, it is governance, policy and responsibility. Outsourcing critical security functions to a third party doesn't just outsource the function as it exists in your organisation. It outsources the function to a third party who may have very different priorities, risk tolerance, and ability to execute.

At the very least, authority and organisational rules do not transfer or translate between organisations well, which may leave you with little control over how a critical security function you rely on is provided once you have signed a contract.

5.0 CYBER INSURANCE

Cyber insurance has been seen by some as the solution to the ransomware crisis, however the question remains as to whether it is the solution for victims or for attackers.

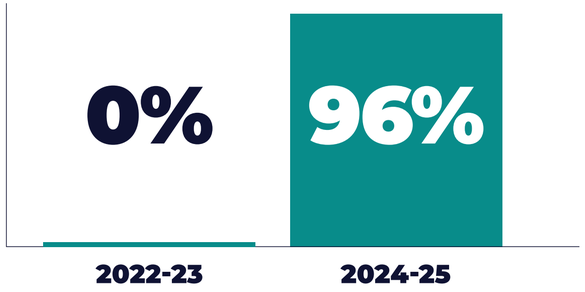
A cyber insurance policy, much like any other, requires policyholders to make certain statements and undertakings in order to get and maintain coverage. Clearly defined requirements for security solutions and processes must be met, which seems positive. However, does having cyber insurance lead organisations to take greater risks? Is cyber insurance seen as an inherent part of cyber security spending and incident preparedness, or as an alternative to it? Does it incentivise companies to outsource security functions and processes to the lowest bidder, ticking all the boxes as cheaply as possible and relying on their insurance to cover them if, or when, their bluff is called by an attacker?

Is the cyber insurance industry accurately predicting and costing in risk to their offerings? Cyber insurance has been seen as a potential growth area by some in the insurance industry, but it is also incredibly volatile. While the insurance industry does indeed deal with volatility for their clients, their main goal is to price that volatility into predictable, regular payments. If they are doing this successfully, they avoid volatility themselves. One way of measuring how successful insurance companies are at assessing risk and exposure is through the ratio of insurance premiums to claims paid, known as the loss ratio. If the loss ratio is below 100%, you're making a profit.

To maintain a balance between risk exposure, growth and profitability, the ideal ratio is generally agreed to be in the 40-60% range.

In the US National Association of Insurance Commissioners 2025 Cyber Security Insurance report of the top 20 insurers offering cyber security insurance, which together account for 76% of the US cyber insurance market, only 8 have loss ratios within that ideal range, with the maximum and minimum ratios being 96% and 9% respectively. Similar variation is seen in prior data, as in the 2024 report only 6 insurers were within the 40-60% range, while loss ratio variability was almost as high, varying between 0% and 80%. A further illustration of the volatility (if one was needed) is that the insurer with the lowest loss ratio in 2022-3 (0%), had the highest loss ratio in 2023/24 (96%).

INSURANCE PREMIUM LOSS RATIO



6.0 DATA THEFT VERSUS DISRUPTION

In Q3 2024, the number of ransomware payments made increased 5% quarter-on-quarter, yet the payment rate for ransomware attacks fell to a record low of 23%.

In the past 24 months, data theft extortion ransomware attacks have generally had a higher payment rate than encrypting ransomware attacks, but in Q3 the payment rate for data leak extortion fell from Q2's 40% to a record low of 19%. That is almost half the 3-year average of 34%, meaning that these attacks currently have a less than 1 in 5 success rate.

The success of the ransomware industry is down to the actions of both attackers and defenders. So, what has changed here? From a defender perspective, could it simply be data leak fatigue? There has been a constant, global flow of data leak disclosures being made, and it is entirely possible that public perception of the impact and severity of data leaks is shifting. When very few organisations were suffering or at least disclosing these attacks, each such attack affecting millions of individuals was shocking.

Now however, it is simply a Tuesday. It is possible that organisations are themselves becoming desensitized to data leak extortion threats, or that they can see that data leaks are not creating the same public reaction (and stock price impact) as they used to, and that is affecting decision making on whether to pay.

From an attacker perspective, are they doing something that makes them less successful? Cl0p have been highly successful data theft extortion attackers for the last several years, and while they were active in Q3 with the Oracle EBS compromises, they have not posted any victims. Are they so strongly influencing the success rate on their own that in quarters when they don't post victims the whole industry success rate goes down?

The ransomware industry is an interplay of both attackers and defenders. If data theft extortion attacks are becoming less successful, attackers will pivot to disruption. And there have been a number of ransomware attacks which have been hugely impactful to the average person on the street.

Q3 2024

5%
INCREASE
RANSOMWARE PAYMENTS

DECREASE
RANSOMWARE PAYMENT RATES TO
23%

1 IN 5
SUCCESS RATE
FOR THIS TYPE OF ATTACK

6.0 DATA THEFT VERSUS DISRUPTION

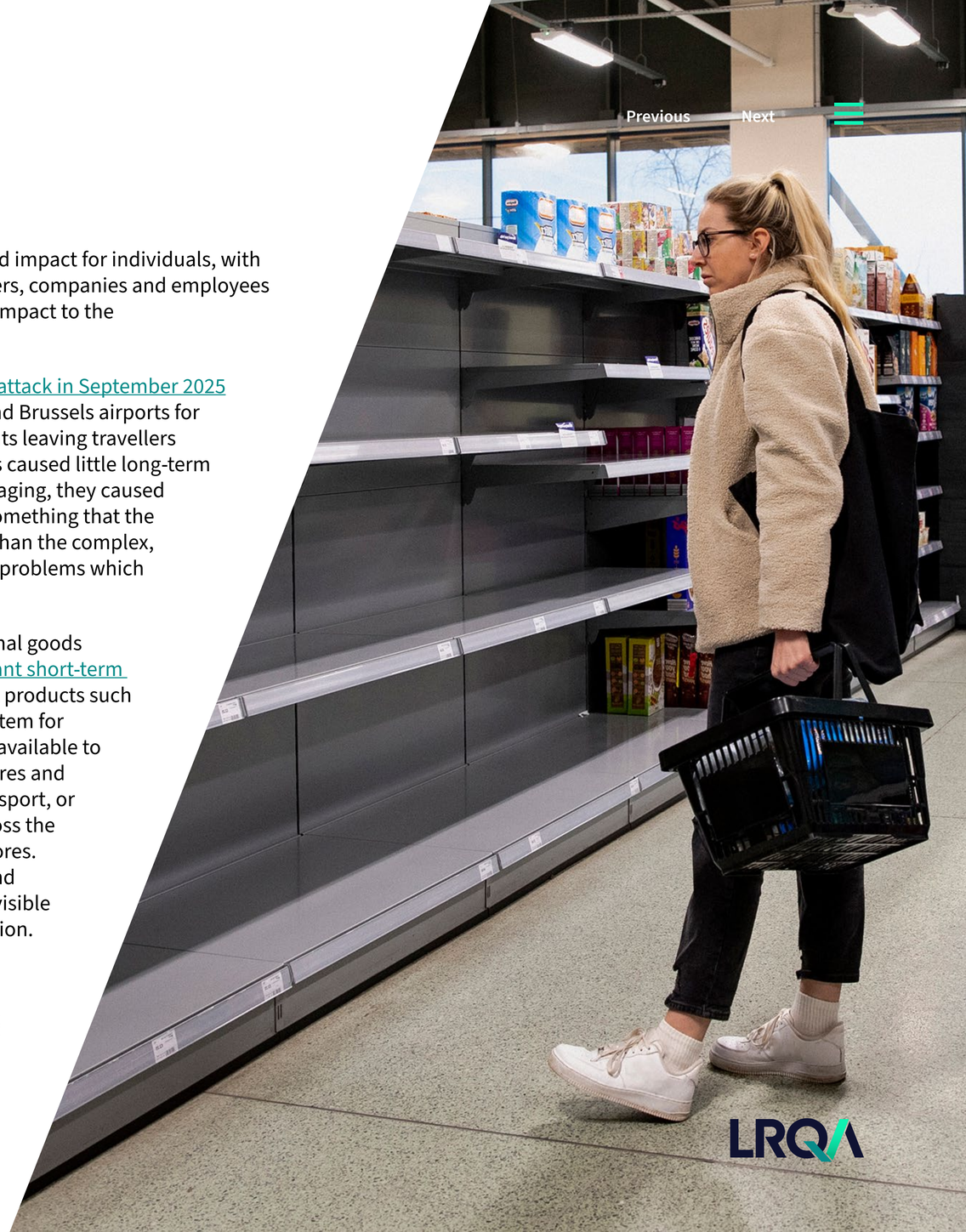
As covered in more detail elsewhere in this report, Scattered Spider caused significant disruption during their thematic and geographically linked campaigns in mid-2025. They compromised:

- UK high street grocery retailers Marks and Spencer and Co-op, with both ransomware attacks occurring over the course of a week, causing over £500 million in losses
- North American wholesale grocery supplier United Natural Foods Inc (UNFI), the major supplier of the Whole Foods Market chain and a service provider to many other stores
- Jaguar Land Rover (JLR), the UK-based car manufacturer who had to halt all manufacture, impacting their employees and customers and their entire manufacturing supply chain (estimated to employ at least 30,000 people), saddling the UK government and public with an economic impact of at least £2 billion.
- Airlines WestJet, Qantas and Hawaii Air, directly interrupting flights and travel for individuals across North America.

In each case, the cyber-attack led to real-world impact for individuals, with empty supermarket shelves, stranded travellers, companies and employees not being paid, and even long-term financial impact to the UK government.

In Europe, a cloud supply-chain ransomware attack in September 2025 impacted London Heathrow, Berlin, Dublin and Brussels airports for several days, with delayed and cancelled flights leaving travellers stranded. While the airport and airline attacks caused little long-term disruption and were technically not that damaging, they caused massive personal disruption to individuals, something that the public are far less likely to forgive and forget than the complex, behind-the-scenes and often delayed-impact problems which follow on from a mass data leak.

In Russia, multiple attacks on the central animal goods certification system, Mercury, caused significant short-term disruption to the transport and sale of animal products such as meat, dairy and eggs. With the Mercury system for electronic certification and origin tracking unavailable to verify the sourcing of goods, distribution centres and retailers were unable to receive, process, transport, or sell these products, leading to disruption across the supply chain and empty shelves in grocery stores. Once again, these attacks led to short-term and relatively minor technical impact, but highly visible and affecting outages for the Russian population.



7.0 FRAGMENTATION: A WEAKNESS IN MODERN CYBERCRIME

While we have already discussed the trend of attackers performing supply chain attacks against service providers in the modern, highly interconnected technology and service environments, there is a flipside to that.

Legitimate businesses use homogenised, highly interconnected *aaS environments because they are efficient, and efficiency is highly prized in the pursuit of profit. Cybercrime attackers are also pursuing profit, and so over time their organisations, economics, and processes have changed so that they now closely resemble those of legitimate businesses.

The difference is that legitimate companies share their achievements on LinkedIn, while cybercriminals advertise theirs on dark-web forums.

Cyber criminals often now specialise as service providers offering malware, initial access, credential dumps, phishing kits, ransomware and more. They advertise and compete for market share, they offer customer support helplines and consultancy, they maintain license servers and offer cloud hosted services. They have become more professional and in doing so created an interlinked service economy. This is referred to as the professionalisation of cyber-crime, and that interlinked service economy means that cyber criminals have become increasingly vulnerable to disruption.

Disruption to this industry has come from both external and internal market actors:

- Externally, law enforcement actions such as Operation Endgame have disrupted the operations of the dark web forum BreachForums, the Lumma and Rhadamanthys Malware as a Service (MaaS) operations, the Phobos/8BASE ransomware operation, and arrested an individual who allegedly ran the IntelBroker cybercriminal persona.
- Internally, criminals fighting for market share have also seemingly caused significant disruption such as the apparent takeover of RansomHub operations by the DragonForce brand.

These disruptions don't necessarily reduce the number of attackers, and their impact on the volume of attacks is often only temporary, however disruption of the cybercrime economy does cause fragmentation.

Large, (relatively) static groupings of criminals can co-ordinate, working together to increase their efficiencies, targeting high value targets and enjoying a higher conversion rate of attacks to payments. One of the original intentions of RaaS brands was to be iconic, creating well-known brands with known capability to both compromise, and to cause damage to victims who don't pay. A well-known, high-volume brand will get more traffic to their data leak site, ensuring greater, more immediate impact to victims who do not pay and are listed there. Breaking up those brands, reducing co-operation, may reduce the success of intimidation tactics, leading to lower ransom demands to increase success rates.

Fragmentation also seems to lead to less co-ordinated targeting. As the number of active data leak sites/RaaS brands increases, there is a greater victim diversity. This could be good for traditional ransomware targets, such as Western financial institutions, but it is bad for the sectors and geographies who suddenly find themselves in the sights of ransomware attackers.

8.0 INFOSTEALERS AND THE REALISATION OF DIGITAL VALUE

Infostealers are an interesting form of cyber-attack, because the impact of the attack is often so far divorced from the attack itself.

Credentials can be stolen and sat on for months or years before they're finally used in a damaging follow-on attack. Because credentials are stolen from a user but can be used on any of the services that user logs into, the authenticating service will likely have no indication that anything is amiss until after the attack. And, because these are legitimate credentials, the only thing that will indicate something is wrong will be the behavioural indicators. Right up until the point the user/customer suddenly complains about unexplained activity on the account.

The Snowflake account compromises were a demonstration of the impact that co-ordinated infostealer-related attacks can have, as accounts compromised over an extended time period, most likely by many completely unrelated, uncoordinated attackers were then all used in a short space of time against a single high value service. That was just the beginning however, as it appears that criminals have heard the lessons of Snowflake loud and clear – 10 million generic webmail credentials might sound impressive, but 30 compromised credentials for the domain of an obscure financial service can make you rich. 2025 has seen a number of really quite interesting attacks which almost certainly derived from historic infostealer infections, and which involved co-ordinated use of compromised credentials against thematically linked services.

In Japan, at least 12 different online share trading services experienced co-ordinated surges of unauthorised logons to share trading accounts. The attackers logged into the accounts, sold all held assets, then used the resulting funds in the account to purchase large amounts of a low-capitalisation stock. This resulted in a drastic rise in the value of that stock. The attacker is believed to have been pre-positioned with ownership of a large number of units of that stock, which they then sold for the inflated price. The attackers had credentials for accounts on multiple services, they were prepositioned on the target stock, and the actions they took were entirely legitimate – they did not even attempt to withdraw the funds from the accounts, they simply sold and then re-invested through the platform itself, an utterly normal and expected activity, and a fascinating way of cashing out from the attacks. It is highly possible we will see this type of attack, or other inventive cash-outs, in future.

In a very similar attack, multiple Australian pension mega-funds, a legally required form of pension investment in Australia, experienced a surge of unauthorised logins to their online management portals. The attackers in this case are reported to have identified accounts where the account holder was over 65, and so able to withdraw funds, and intentionally targeted those accounts. Some of the targeted portals required multi-factor authentication to approve fund transfers/withdrawals like that, however in many cases the default form of MFA was an email to the account holder. If, as seems very likely, these pension portal credentials were compromised by infostealer infections, it is highly likely that the email credentials were also stolen in the same attack. As such, the attackers were able to then simply approve the transfer from the recipient's email address.

Less interesting, and seemingly less impactful attacks have been seen against a number of online fashion/luxury retailers. In the case of The North Face and Cartier it appears that attackers logged into accounts with compromised credentials at volume and then copied out PII and saved information from the accounts. In other cases, such as Dior, Louis Vuitton, Adidas, and Mango, reporting is less clear, but it appears that a database or databases held by third parties were accessed in a supply chain attack. While this does not have the immediate impact of the share trading or pension account attacks, it has almost certainly allowed someone to build a valuable database of PII and financial activity for individuals who have accounts with (i.e. regularly purchase) luxury fashion brands. Individuals who regularly purchase luxury fashion brands online are likely to have above average amounts of money, as well as being known to make high value purchases over the Internet, making them targets of above average value for fraud and social engineering attacks.

9.0 POISONED PACKAGES

Pretty much all software development today is achieved through the use of software libraries, or packages. These are pre-written blocks of code which are intended to solve certain common (and often complex) problems so that programmers do not have to reinvent the wheel every time they start a new project.

From an efficiency viewpoint, this is of course a good thing. Various ecosystems sprang up in order to service the needs of programmers using libraries, with online repositories of software packages and local tools which developers can use to find, download and immediately start using any package they specify. Software packages are so prevalent that it is highly likely that any software package will itself rely on code/functionality from other software packages.

The problem with this, however, is that packages are executable code written by other people, and by downloading a package a programmer is choosing to trust both the package, the developers of that package, and the developers of any packages they are using. This extends a trust relationship, and trust relationships can of course be abused in many ways. In fact, beyond the trust between developers there is an inherent trust between the developers and whatever package repository (NPM, PyPi, etc.) they are using. Fortunately, open-source, public software libraries are available for public scrutiny. This means they can be verified to be doing what they are intended to do in a secure, trustworthy manner. This does not mean that they necessarily are regularly scrutinised and verified, just that they can be.

To recap then, software packages represent a chain of trust with multiple links, which results in software developers (who are themselves creating software for others to use) incorporating unverified executable code into their products. It should be obvious from this where the crime comes in.

ATTACKERS ARE EMPLOYING SEVERAL TECHNIQUES WHICH TARGET THESE TRUSTS:

They are attempting to compromise contributing developers to software packages so that they can insert malicious functionality into trusted packages.

They are creating and uploading their own novel packages which claim to have useful, desirable functionality, but include malicious content.

They are typo-squatting existing, popular software packages with their own malicious versions, so that if a developer mis-spells the desired package name, they get a malicious payload instead.

A key part of this type of attack is that it is inherently a software supply-chain attack. Software packages are used by software developers to create software, which is then used by users. As such, by compromising a legitimate software package, or creating a novel malicious package, an attacker can compromise not only the developers who use the package, but any of their downstream users. And of course, the developer may themselves be creating software packages which other developers will then import into their products, potentially spreading the attack even further.

There have been a number of high-impact historical instances of this, and in 2025 there have been regular, repeated attacks against multiple software development package ecosystems.

There is no reason to believe that this trend will tail off in the foreseeable future, in part because there have not yet been any changes to processes which would appreciably affect the ease or success of these attacks.

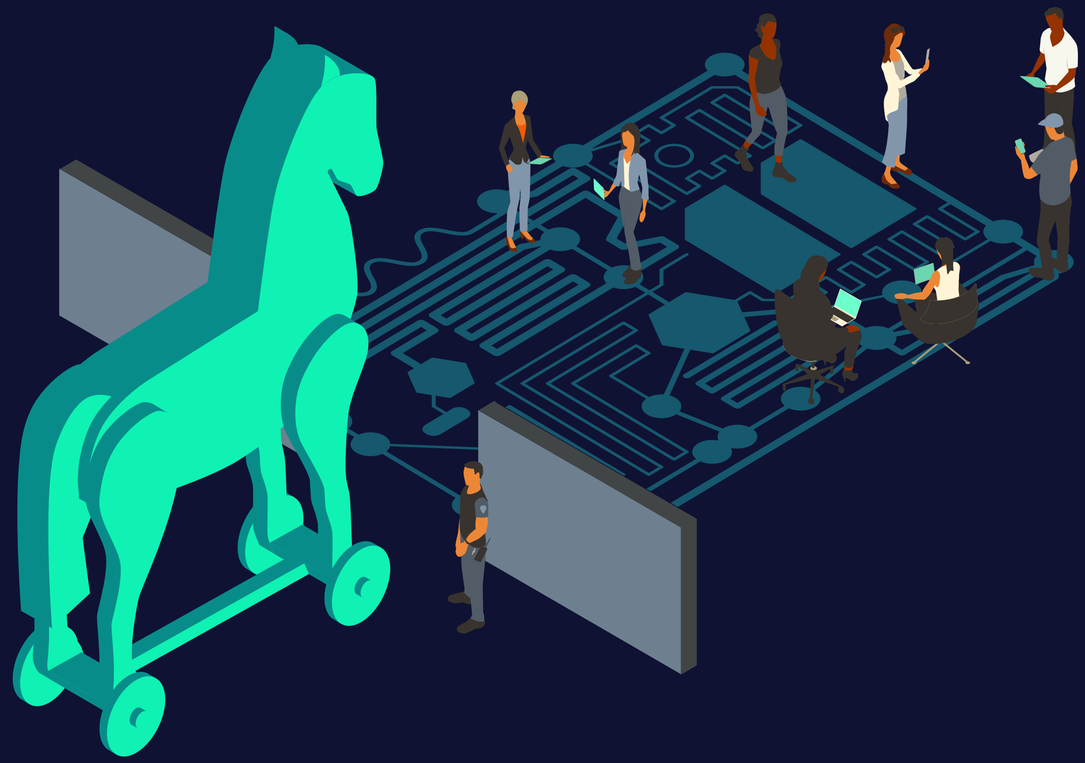
9.0 POISONED PACKAGES

In September 2025, the @ctrl/tinycolor package on the npmjs registry was compromised with self-replicating malicious code. Tinycolor is downloaded more than 2 million times per week, and by the time the compromise was identified it had propagated to a further 187 npm packages, including packages published by CrowdStrike. When a software developer/package maintainer is compromised by the malicious code, the code downloads any packages the developer has permissions to maintain, injects itself into the package, then republishes it, automatically trojanizing downstream packages. As well as propagating itself, the malicious code searches the developer's device for credentials and secrets/tokens, then sends them to a hardcoded URL/webhook.

In August 2025, the Nx project was compromised. Nx is used in enterprise scale JavaScript/TypeScript and has 5.5 million weekly downloads from the NPM package index. The attacker identified that Nx had published a GitHub workflow which was vulnerable to command injection. While Nx had identified and removed this vulnerable workflow, they had not understood how GitHub actually works, and so the workflow was still present and accessible. In addition, Nx used the vulnerable Pull_request_target trigger, which meant that the vulnerable workflow was executed with privileged permissions. As such the attacker was able to cause the Nx build process to leak their NPM token, allowing them to upload modified, malicious versions of the package on NPM. The malicious versions of the package would run an infostealer script on the local system, then create a public GitHub repository with a name containing the string "s1ngularity-repository" under the user's account and post the stolen credentials there.

Some developers who had not used Nx found that they had been compromised simply by having the NxConsole extension for the VisualStudio Code IDE installed. Investigation by the Nx maintainers found that this occurred because NxConsole would automatically install the latest version of the Nx package. As such, when a VisualStudio Code editor was opened while the NxConsole extension was active, the malicious code would be automatically executed.

Another interesting part of this compromise is that the infostealer used installed command-line tools for the Claude, Qq and Gemini LLMs, instructing them to search for credentials and private keys and write them to a file in /tmp. The prompt used changed in successive versions of the malicious code, indicating that the actor was trying to improve its success rate. Researchers note however that the attempt to use LLMs for this meant that the attack was less effective than it would have been if the actor had simply used command-line tools. As a result of this attack over 2,000 GitHub users were compromised.



10.0 LLMs – THE S STANDS FOR SECURITY

Although commonly referred to as AI, the technology attracting the most attention today is specifically large language models and LLM chatbots rather than wider machine-learning techniques.

An LLM is of course a highly complex accumulation of code and data, with a long, complex supply chain, typically involving at least one cloud supplier. Almost everything that goes into making AI models is complex, yet this is often hidden because their output is typically in formats which people inherently understand, so it seems simple. The primary use of LLMs is to apply them to complex data and situations in an attempt to simplify them for human interaction or consumption.

The major capability of an LLM is to generate output which is plausible based on statistical analysis of its training data set. Plausible does not necessarily mean correct or accurate, however. Unfortunately, this means that LLMs are ideal for generating phishing data such as emails, websites, or social media profiles, but less useful in situations where you need accurate output.

LLMs have found some success at generating code. The reason for this is the ease with which the output from coding tasks can be automatically determined to be correct or incorrect. The success of the LLM can be measured at machine speed by executing the code, and so LLMs can be put through high volumes of reinforcement training cycles to improve coding ability in a short space of time. This is not possible with other fields of effort which do not have outputs that can be programmatically determined to be correct or incorrect.

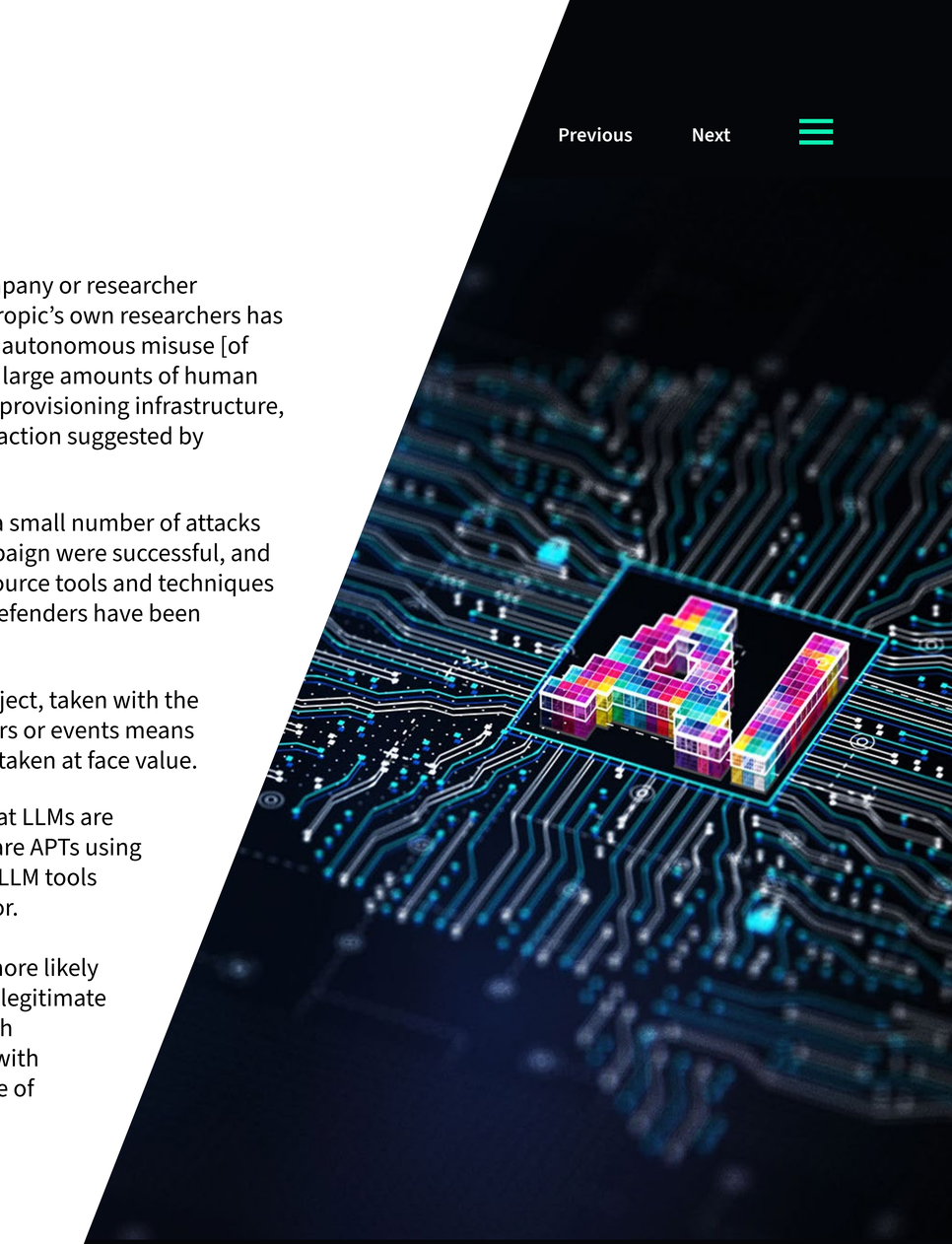
There have been reports from both Google and Anthropic of malicious attackers using LLM functionality, either directly to create tools and perform malicious actions, or implementing LLM functionality within their tools. However, in each case these reports about an LLM being used to enable allegedly groundbreaking malicious functionality are released by the same company that makes and sells the LLM in question. Google reported about malware which prompts their Gemini LLM to generate code, while Anthropic reported on malware that prompts their Claude LLM. In neither case however has any other company or researcher observed this type of advanced behaviour. One of Anthropic's own researchers has stated that while the activity they describe is "the most autonomous misuse [of Claude LLM]", it still wasn't that autonomous, involving large amounts of human effort to first create an entire orchestration framework, provisioning infrastructure, and validating every piece of information generated or action suggested by an LLM.

In addition to this, Anthropic themselves say that only a small number of attacks performed by this highly engineered LLM-abusing campaign were successful, and the LLM simply used the same readily available open-source tools and techniques which attackers have been using for years, and which defenders have been defending against for years.

The overlap of reporter and financially incentivized subject, taken with the lack of any credible examples from any other researchers or events means these reports are slightly too marketing-adjacent to be taken at face value.

These reports seem to want to create the perception that LLMs are definitely useful to malicious attackers, and that there are APTs using them to do APT things which only the creators of these LLM tools are aware of, and which no one else has any evidence for.

Instead, the activity described in these reports seems more likely to indicate that whether you are a nation state APT or a legitimate business, efficiency is key, and many organisations, both legitimate and otherwise, are currently experimenting with integrating LLM tools into their workflows with the hope of greatly increasing their efficiency and output-per-FTE.



10.0 LLM – THE S STANDS FOR SECURITY CONT.

One seemingly unintentional result of the rapid implementation of LLM-based tools is data leaks. In June, it was discovered that Meta-AI conversation logs which users had intended to remain private were being posted to a public feed by Meta, intended to advertise the service. In August, it was found that hundreds of thousands of Grok conversation logs were publicly accessible and searchable in Google, and thousands of OpenAI chat logs were also made publicly searchable. In each case this was seemingly a result of the AI companies adding in the ability for users to make their chats visible in web searches, but without enough pre-release testing to ensure that the functionality was labelled in a way that the average user could understand.

In November, it was disclosed that OpenAI had been sending entire ChatGPT prompts to Google, and that these allegedly private prompts were then being shown to Google Analytics users. The Google Search Console allows site admins to see what search terms led users to visit their site. Because ChatGPT was sending the entire prompt to Google, that was then being shown to site administrators. And because of the nature of search engine indexing combined with the format of ChatGPT's queries, highly personal and off-topic prompts were being sent to and shared with utterly unrelated web sites.

Iranian state-sponsored faux-hacktivist group CyberAv3ngers were observed to be using ChatGPT for research and reconnaissance, as did China-linked threat actor SweetSpectre. However, the attacks which came out of this research were still just traditional attacks, with ChatGPT acting as a stand-in for Google during the research and reconnaissance phase.

Researchers from Volexity observed UTA0388 sending 50+ unique phishing emails in 5 different languages in a short space of time, with the content in each language coming across as fluent and natural.

In the same campaign, UTA0388 rapidly deployed multiple simple,novel malware samples, and while each one had additional features or complexity compared to the previous version, they were not the same code base, an unusual choice which suggests LLM coding. This does suggest that use of LLMs can rapidly scale phishing attacks and aid with existing attack vectors, although because they can only generate variations based on their training data, they cannot generate new attack techniques or vectors, and they require close human supervision and direction. The researchers in this case also note that some of the strongest indications of the use of LLM generated content during this campaign was non-sensical decisions with no logical basis. For example, while it is impressive that multiple unique phishing emails were sent in multiple languages, the languages did not necessarily align with the desired attack or social engineering lure. In one case the actors sent a phishing email to an English-speaking recipient, with a subject line in Mandarin and a body text in German, which claimed to be from an English-speaking American persona. The malicious payloads and phishing emails were also found to contain purposeless additional content and files, with repeated pornographic references in unusual locations.

11.0 GEOPOLITICAL DISRUPTION/CYBER SOVEREIGNTY

Because cyber activity is ultimately driven by people, major real-world events tend to produce noticeable changes in the threat landscape.

The EU and other European countries are heavily focussed on Russia and its invasion of Ukraine. This focus is very understandable and quite sensible, as Russia is similarly focussed on Ukraine and Europe, is known to operate extensive state-sponsored hacking campaigns and has historically been seen as the spiritual home of ransomware and cyber-crime. Indeed, cyber-disruption has been a key weapon of both war and propaganda during this conflict. Russian cyber-attacks have at times been co-ordinated alongside kinetic attacks to increase impact or impair responses. Beyond the geographical borders of the war zone there have been cyber-disruption operations against supply chains, state and political entities, and anybody showing support for one side or the other.

The Middle East is once again in turmoil, focussed on Israel, Iran, Gaza and Syria. All sides have employed active cyber-attacks, with state-operated and tacitly acknowledged/supported "patriotic hackers" performing disruption and propaganda. America, meanwhile, has stepped back from its historically significant international outreach, become more heavily focussed on internal politics and cultural friction.

This has had dual impacts.

Firstly, it has led to concerns about the reliability of internationally recognised and relied upon services such as the National Vulnerability Database (NVD), which while operated by the US is relied upon internationally. In response to concerns about service disruption, multiple other vulnerability databases explicitly intended for international use were stood up, with two being launched in EU countries within a matter of days. That's how seriously the industry has taken the potential loss or instability of the NVD.

Secondly, America's stepping back from international outreach and leadership has left a soft-power vacuum which China has sought to fill. In Africa and South America Chinese technology, infrastructure, and services have been enthusiastically offered by Chinese state-aligned companies, which has led to concerns around security and independence by Western countries and businesses.

12.0 THE EXPLOSION OF *FIX ATTACK VECTORS

Aside from gaining legitimate credential access, we have seen a surge in techniques that fundamentally reimagine payload delivery by transforming users into unwitting execution vectors. These attacks do not attempt to bypass traditional external to internal controls, they convince users to bypass them on the attacker's behalf.

This paradigm shift began with Browser-in-Browser (BitB) attack research in 2022, which demonstrated that spoofing trusted UI elements could manipulate even security-conscious users. BitB proved that the psychological barrier to clicking within a familiar interface was far lower than traditional phishing methods.

By creating a fake browser window complete with a spoofed URL bar that was "basically indistinguishable" from legitimate authentication pop-ups, researchers showed that visual trust could be weaponised at scale.

We believe this research laid the conceptual groundwork for what would become the "*Fix" family of attacks, ClickFix, FileFix, and their variants, which evolved the principle from visual deception to behavioural manipulation. Rather than simply mimicking trusted interfaces, these techniques exploit our ingrained response to "fix" problems, transforming helpful user behaviour into an attack vector.

BITB (2022)

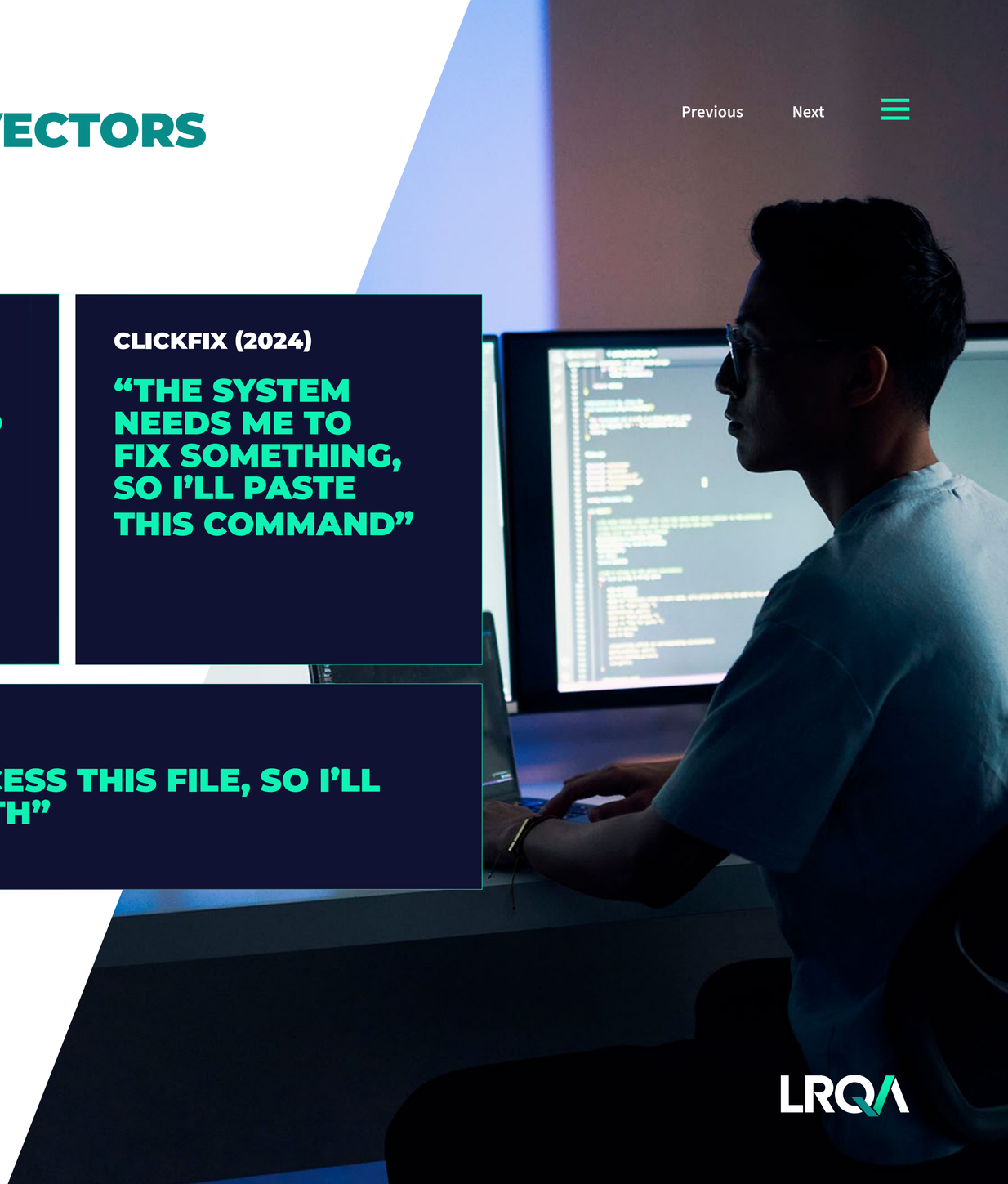
"THIS LOOKS LEGITIMATE, SO I'LL ENTER MY CREDENTIALS"

CLICKFIX (2024)

"THE SYSTEM NEEDS ME TO FIX SOMETHING, SO I'LL PASTE THIS COMMAND"

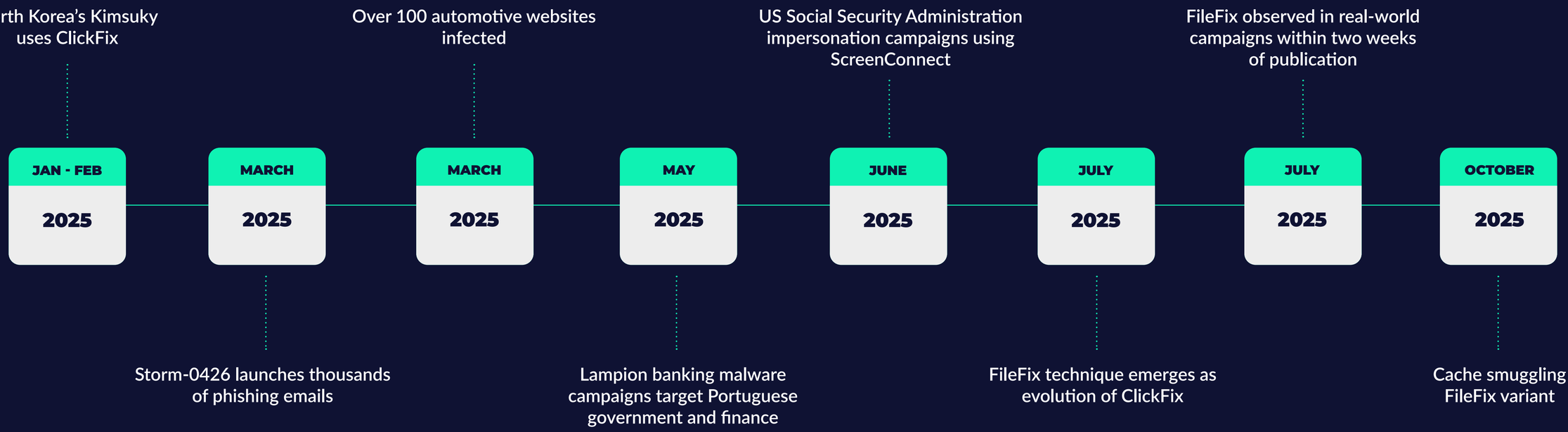
FILEFIX (2025)

"I NEED TO ACCESS THIS FILE, SO I'LL PASTE THIS PATH"



12.1 *FIX TIMELINE

Each evolution reduced the cognitive friction between user and malicious execution, exploiting increasingly routine behaviours. Notice the absence of weird and wonderful zero days?



12.2 THE RISK

The *Fix family of deception techniques pose a significant real-world threat to organisations globally because they leverage social engineering to bypass conventional security measures.

No technological link is observable between whatever social engineering content was presented to the user, and the action the user then takes. The subtlety of ClickFix, using legitimate user interaction to perform malicious activity, makes it difficult for automated security systems to detect, amplifying its potential impact.

Given the rapid evolution of *Fix techniques in 2024 and 2025, including their adaptation to target macOS and Linux platforms, threat actors are expected to refine it further in 2026.

It is highly likely that AI-driven lure generation will increase the viability and ease of use of this and other social engineer-based attack.

13.0 INSIDER THREATS

The term insider threat relates to an individual who has authorised, legitimate access, who then abuses that access to perform unauthorised activities. Typically, this is seen as disgruntled employees, or just bad people taking advantage of their employer for their own profit or satisfaction. Recently however additional types of insider threat have become an issue for many organiations.

DPRK (North Korea) has tasked agents of the state with applying for large numbers of remote IT roles under false pretences. These agents use LLMs and deepfakes to generate job applications and CVs, and to pass interviews, while supplying fake personal information and lying about their location and identity. Once they are employed, they then use LLMs to perform their duties with as little effort as possible, with at least one agent known to have held down 12 jobs at one time. Simply by drawing salaries for these jobs this operation is estimated to have made at least $88 million, however the campaign appears to be expanding into other methods of monetisation. Once the agents have access to their new employers, they have been reported to steal data, to infect the network with malware, and if/when they are fired, to deploy ransomware or make extortion demands.

As well as the DPRK, organized crime groups such as Scattered Spider and the ransomware gang Medusa have been offering to pay money to employees for access to their accounts for the purpose of deploying ransomware.

> Scattered spider publicly advertised that they would offer 10% of any ransomware payment to an individual that gave them access to their employers network,

while Medusa approached a BBC employee via telegram, offering them 1 Bit Coin for access to the BBC network.

Coinbase, a cryptocurrency exchange, received extortion demands from attackers threatening to leak data stolen from a large number of Coinbase customers. Coinbase did not pay the extortion demand, and it was soon identified that the data had been stolen by customer support agents at one or more of Coinbase's outsourced customer support call centres in India. These outsourced employees had been approached by an organized crime group offering payment if they would take photos of customer account information pages while they were on calls and send them to the group. Because of the huge difference in value between the employee's salaries and the cryptocurrency accounts they had access to, the criminals were able to offer them sums of money that were significant to them, without significantly impacting their potential profits in any way.

THE THEME OF 2025: AVOIDANCE

Cyber defence is a constant race to shore up defences against new techniques and newly discovered vulnerabilities.

Endpoint Detection and Response (EDR) solutions attempt to provide a backstop so that if attackers bypass other defences and controls, malicious activity attempts on the desktop will be detected and ideally prevented. At the very least they will be responded to before they can become critical incidents. EDR has been so very effective at this that attackers have begun to adopt techniques and attack types which avoid competing with EDR entirely. If a defence is effective, often the best thing you can do is not to attempt to overcome it, but to attack in such a way that you do not engage with it.

Side-stepping defences, particularly EDR, has been the motivator and unifying theme of the trends and critical incidents of 2025.

Infostealers provide access to legitimate account credentials, with the use of these credentials completely divorced from the act of compromise and data theft.

Social Engineering persuades legitimate, authorised individuals to perform ill-advised activity on behalf of attackers, whether this is performing password resets on accounts, or executing commands via ClickFix and FileFix.

Compromising and dwelling on network infrastructure gives attackers access to one of the few places on the network that EDR cannot be present, and the almost omnipresent nature of edge network security infrastructure means that these devices can be easily accessible. Cloud **environments** cannot be monitored via EDR, and often they are administered and owned by an entirely separate organisation.

These environments intentionally do not reveal all of their internal workings to users or organisational administrators, and as a result the logging that is provided to customers is intentionally obtuse. This can make these environments perfect targets for evasive, stealthy attackers looking to abuse chains of trust in environments of low surveillance.

Business Email Compromise attacks can take a single compromised mailbox and leverage that to compromise multiple organisations without ever touching a desktop or server.

Modern security controls such as EDR have been very effective, but attackers are just as aware of this as defenders, if not more so. As such, competent attackers will intentionally **avoid** EDR where possible. Traditional cyber-attacks which compete against EDR can still be successful, if only because many organisations and individuals are still not security focused and are not running EDR effectively. However, the current and future trend is that increasingly effective cyber defences on endpoints lead to evasive tactics by advanced attackers, who will avoid defensive controls wherever possible.

This has been seen throughout 2025, and since it has been so successfully and publicly demonstrated, the popularity of these attacks will grow until a critical mass of potential victims implement successful defences.

ABOUT LRQA

LRQA is a leading global risk management partner.

Through our connected risk management solutions, we help you navigate an evolving global landscape to keep you one step ahead.

From certification and cybersecurity, to safety, sustainability and supply chain resilience, we work with you to identify risks across your business. We then create smart, scalable solutions, tailored to help you prepare, prevent and protect against risk.

Through relentless client focus, backed by decades of sector-specific expertise, data-driven insight and on-the-ground specialists across assurance, certification, inspection, advisory and training, we support over 61,000 organisations in more than 150 countries.

LRQA – Your risk management advantage.

GET IN TOUCH

Visit **LRQA.com** for more information or email **enquiries.uk@lrqa.com**

LRQA
1 Trinity Park
Bickenhill Lane
Birmingham
B37 7ES
United Kingdom

LRQA
Your Risk Management Advantage