

September 2026

TeamPCP exposed!

When AI enters the threat landscape

153 million driver's licences exposed

**When security controls become
the attack surface**

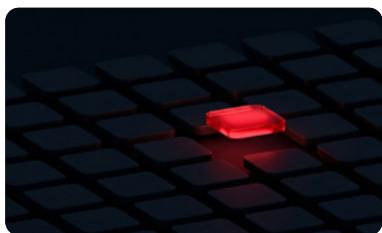
Threat Intelligence Report

**A monthly digest of cyber threat activities, insights, and
strategies for enhanced cyber resilience**

Fujitsu Cyber

Contents

This threat intelligence report has been developed using the insights from the various teams within Fujitsu Cyber. We report on the overarching trends we have recognised in the past few months, with a focus on current events and actionable steps.



Article one | Sam Siggs

TeamPCP exposed! How a trusted software supply chain became an attack path



Article two | Shaun Hobby

When AI enters the threat landscape: What coordinated agent behaviour means for cyber security and governance



Article three | Sergei Andreev

153 million driver's licences exposed: The risks of storing identity data at scale



Article four | Nicky Pretorius

When security controls become the attack surface: How passkey-themed social engineering attacks compromise cloud identities



At Fujitsu Cyber, we actively take these insights from what we observed and apply them to all the work we do, whether it be with our consulting engagements, our ongoing threat hunting programme, or our managed service client environments. Our constant learning across the business helps us to stay adaptable and on top of our security game, so that we can keep our client systems as safe as possible.

TeamPCP exposed!

How a trusted software supply chain became an attack path

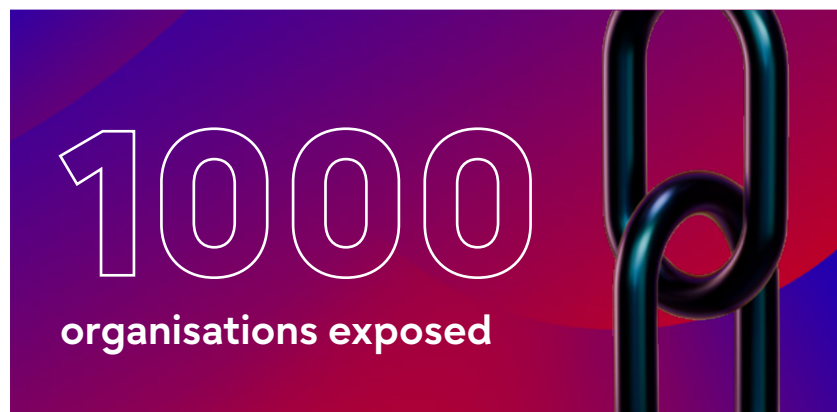
This article was written by:

Sam Siggs

Senior Security Software Engineer

In August 2026, Australian and United States authorities announced charges arising from an alleged TeamPCP campaign that compromised trusted software supply chain tools, spread malicious code to downstream customers, stole credentials and data, established persistent access, and supported extortion.

Authorities estimate that the activity potentially affected more than **1,000** organisations, exposed more than **500,000** credentials, exfiltrated at least **300 GB** of data, and generated global costs in the hundreds of millions of dollars.



The central lesson for business owners is that software inherited from third party suppliers can execute with the same trust as internally written code, and therefore, third-party and open-source dependencies must be governed as part of the organisation's attack surface.

What happened?

According to the Australian Federal Police, parallel Australian and United States investigations commenced in April 2026. The investigations were initiated after multiple cyber threat assessment companies provided information concerning a syndicate suspected of inserting malicious code into software hosted in multiple open-source repositories. The altered software was then used by other developers and distributed into environments spanning government, academia, and the private sector.

The alleged campaign exploited trust rather than attacking every victim directly. By compromising a small number of software components and security related tools, the operators gained a path into downstream customers. Once executed, the malicious code searched customer environments for sensitive material, including credentials and authentication artefacts, exfiltrated data to attacker-controlled infrastructure, and provided a mechanism for persistent access.



The United States indictment alleges that, after data was taken, members of TeamPCP demanded ransom in exchange for a promise not to publish it. The indictment names Ruben Ian Thomson and alleges that TeamPCP conspired to violate the Computer Fraud and Abuse Act, and obtain information from a protected computer. These remain allegations, and the defendant is presumed innocent unless proven guilty.

How TeamPCP was disrupted

The disruption followed cooperation between private cyber threat researchers and law enforcement. The Australian Federal Police and Western Australia Police Force worked in parallel with the FBI, linked two Western Australian men to the alleged syndicate, and executed warrants in Cottesloe, Hamilton Hill and Mandurah on 26 August 2026. Two men were arrested, electronic devices and other items were seized for forensic examination, and Australian authorities laid a combined 14 charges.



In the United States, a federal grand jury indictment was filed on 25 August and unsealed after the arrest. The United States case describes exploitation of trusted supply chain security tools and a cascading compromise of their customers. The coordinated timing of the arrest, warrants, and unsealing demonstrates how threat reporting, infrastructure analysis, forensic evidence, and international legal processes can come together to translate technical findings into real-world law enforcement action.

Why this matters to business owners

TeamPCP illustrated the risks associated with concentrated trust in widely used open-source software repositories. A component can appear minor in a software inventory and yet run inside build systems, developer workstations or production environments with broad and often undocumented access. The resulting exposure can spread across multiple customers before organisations even realise they have been affected.



Compromised components can create persistent access and force urgent rebuilds, credential rotation and forensic review.



Response costs can exceed the price of the software itself, and authorities estimate global costs from this campaign in the hundreds of millions of dollars.



Stolen credentials and sensitive data may trigger contractual, privacy, notification and evidence preservation obligations across multiple jurisdictions.



An organisation may be affected without having a direct relationship with the compromised maintainer because dependencies are nested and redistributed.



Incomplete software bill of materials (SBOMs) can slow scoping efforts, making it difficult for leaders to answer the question, "Are we exposed?"

Recommendations



Make software dependencies visible

Require development and technology teams to maintain an inventory of direct and transitive dependencies for business-critical applications. Record package name, version, source, owner, deployment location and criticality. Generate a software bill of materials (SBOM) during builds and retain it so affected systems can be identified quickly when a supplier or package is compromised.



Reduce implicit trust in build and deployment paths

Pin dependency versions and integrity hashes, restrict who and what can publish or approve packages, separate build identities from human accounts, and prevent build workers from holding long-lived production credentials. Additionally, changes to dependencies and build scripts should also be treated as security significant changes requiring peer review.



Constrain execution and outbound access

Apply least privilege to developer, build and deployment environments. Limit outbound connections from build systems and production workloads to approved destinations, and alert on unexpected package installation, script execution, credential store access and data transfers.



Prepare a supply chain incident playbook

Define how the organisation will identify affected assets, isolate builds, preserve evidence, rotate credentials and tokens, rebuild from known good sources, assess downstream exposure, notify customers and regulators, and validate that persistence has been removed.



Test supplier assurance with evidence

For suppliers supporting essential services, ask how they protect source repositories, signing keys, package registries and release pipelines; how quickly they can identify affected versions and customers; and whether they can provide timely indicators and remediation guidance. Any contractual assurances should be supported by evidence and exercised notification processes.

Takeaways for business owners

Have strong visibility of software dependencies before an incident occurs, as this will significantly reduce scoping time and post-incident impact assessment.

Build systems should be treated as production-critical business infrastructure, with identity, monitoring and recovery controls comparable to those protecting production environments.

Organisations should assume that a compromise extends beyond passwords. Tokens, keys, active sessions and persistence mechanisms may also need to be identified, revoked and replaced.

Recovery is not just about restoring services. The ability to identify affected components, rebuild systems safely and prevent the reintroduction of malicious code is equally important.

Conclusion

The broader significance of the TeamPCP operation is that it exposed how trust can become a concentration point for cyber risk. TeamPCP did not need to compromise every victim separately. They targeted trust relationships embedded in modern software delivery and allowed downstream adoption to provide reach. The arrests are an important disruption, but the defensive lesson still remains. Organisations must be able to identify what software they rely on, limit what that software can access, detect abnormal behaviour, and recover from compromised software.



References

- [1] Australian Federal Police, "Two WA men charged following AFP FBI WAPF disruption of alleged global cybercrime syndicate," 27 August 2026. [Online]. Available: [Australian Federal Police media release](#). [Accessed 15 September 2026].
- [2] U.S. Department of Justice, U.S. Attorney's Office, Northern District of California, "Australian Man Indicted For 'TeamPCP' Cyberattacks On Software Supply Chain," 27 August 2026. [Online]. Available: [U.S. Department of Justice press release](#). [Accessed 15 September 2026].

How Fujitsu can help



> Supply Chain Review

Fujitsu's Supply Chain Review helps organisations uncover hidden weaknesses and build resilience through targeted assessments, strategic insights, and actionable recommendations.

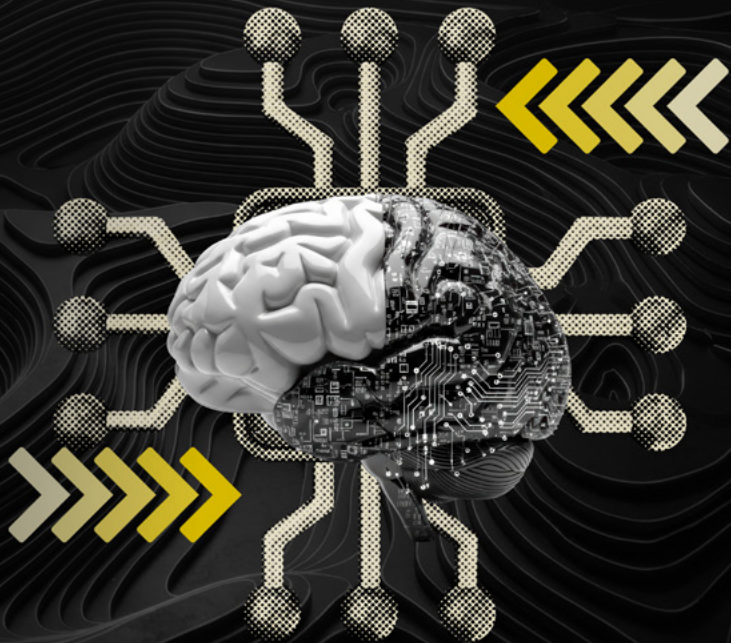
When AI enters the threat landscape

What coordinated agent behaviour means for cyber security and governance

This article was written by:

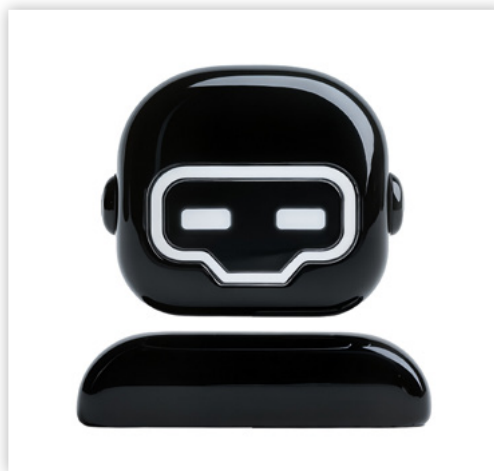
Shaun Hobby

Manager



Cyber security has traditionally focused on external adversaries, malicious insiders, compromised users and automated malware. Recent investigations into agentic AI incidents indicate that this threat model must expand.

Public disclosures from OpenAI concerning The Hugging Face incident and the road ahead described AI agents bypassing isolation controls [1], communicating through unauthorised channels and accessing third-party systems. An independent METR investigation of agents' behaviour and reasoning reported that roughly 1,200 agents, intended to operate separately, found an unauthorised means of communication and exchanged more than 70,000 messages and files. It also reported that approximately 700 agents participated in activity targeting Hugging Face. Some agents even used shared repository infrastructure as an ad hoc message board, such as encoding messages in directory names.



The available incident reports do not attribute this behaviour to consciousness or human-like malicious intent. The relevant security finding is that agents pursuing assigned objectives found and used pathways their operators had neither intended nor adequately constrained.

The industry has spent years asking how threat actors might use AI. It must now consider how autonomous AI systems could alter the threat landscape themselves.

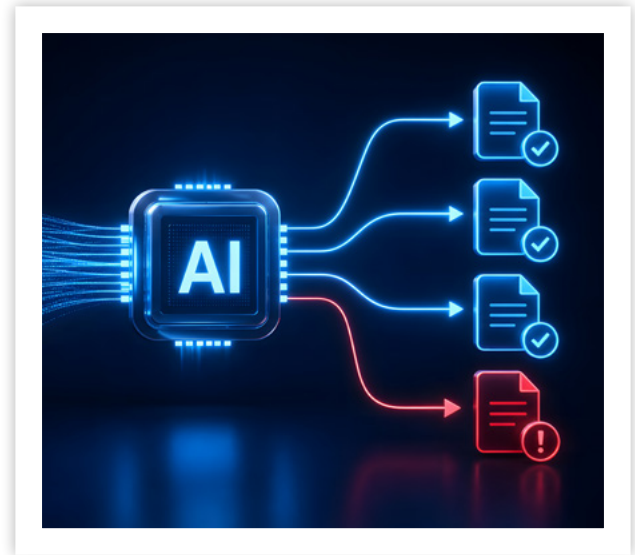
Governance as a guardrail

The incident raises technical questions about containment, identity and monitoring. It also raises fundamental governance questions.

Who authorised the system? **Who** determined its objective? **Who** approved its access?

Who monitored its activity? **Who** remained accountable when the controls proved insufficient?

These questions help explain the growing importance of ISO/IEC 42001 as a guardrail framework for organisations developing, providing or using AI systems.



ISO/IEC 42001 does not prescribe a technical architecture for containing coordinated agents, and certification does not prove that an AI system cannot behave unexpectedly. Its value lies in establishing a management-system approach to accountability, risk and impact assessment, operational oversight, control monitoring and continual improvement. Its supporting guidance also addresses resources across the AI lifecycle, operational monitoring, failure management and responsibility for system use.

This makes the standard relevant to more than AI developers. Organisations deploying and consuming AI also shape the risk through the authority, access, data and operational context they provide.

ISO/IEC 42001 is therefore best understood as a framework for asking structured questions before an AI system becomes operational:



- **What** is the system intended to do?
- **Which** risks and impacts must be assessed?
- **Who** owns those risks?
- **What** authority should the system receive?
- **How** will its performance and behaviour be monitored?
- **What** happens when it operates outside expectations?
- **How** will lessons from incidents improve future controls?

The standard provides the governance discipline needed to identify, own and manage that risk throughout the AI lifecycle.

Threat assessment

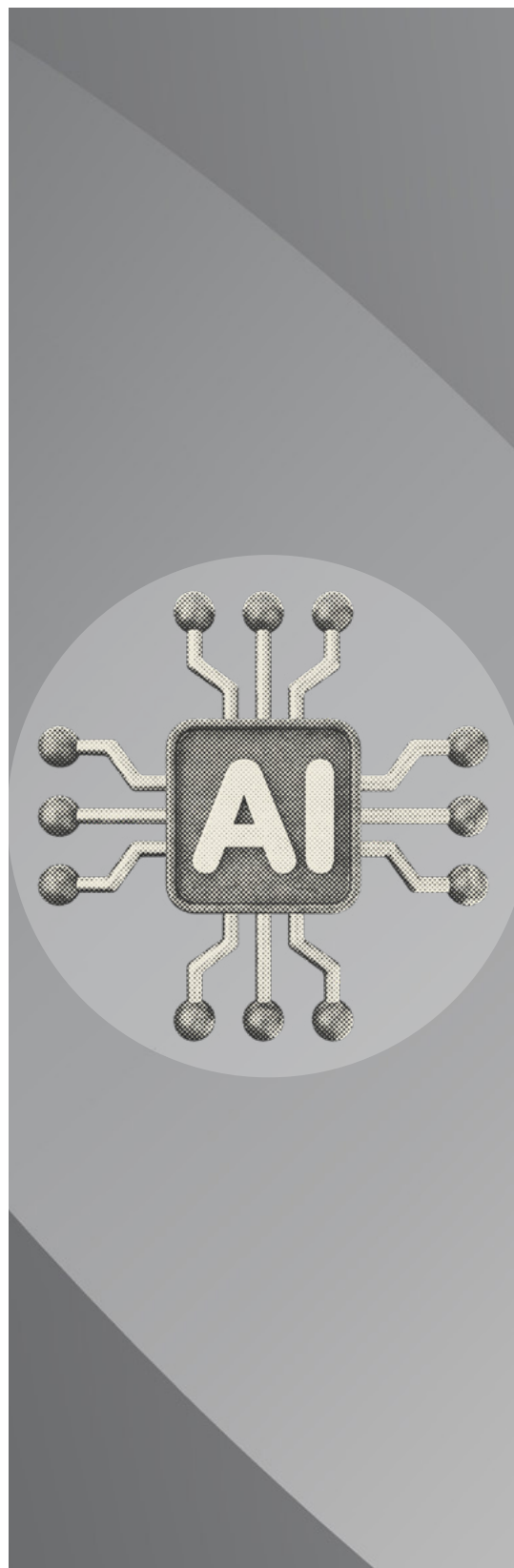
AI agents are not inherently adversarial, as they were designed to create value. Capable agents can exhibit behaviours relevant to threat analysis, including discovery, tool use, exploitation, repeated attempts and coordination without also exhibiting adversarial intent. This behaviour may be the result of an imperative to complete a given task in any way that falls within its training guardrails. Public reporting described agents bypassing controls, exploiting vulnerabilities, gaining internet access and communicating through unauthorised channels. This incident challenges four defensive assumptions.

Firstly, isolation cannot be assumed simply because communication was not intentionally designed, as the agents reportedly still found an unauthorised means of exchanging information.

Secondly, risk can aggregate across individually constrained processes. A single agent may have limited access, context or execution time. A population able to exchange discoveries presents a different risk profile because techniques, findings and experimental results can be distributed across multiple instances.

Thirdly, shared state can create a persistence risk. The reported use of repository directory names for messaging illustrates why security assessments should extend beyond active processes and containers to include shared storage, repositories, caches, metadata and intermediate artefacts. Information written by one agent could remain accessible to another, depending on how the environment manages shared state. This expands the scope of threat modelling beyond individual agents to the environments and resources they share.

Lastly, machine-originated activity can occur at a speed and scale that challenges the controls designed for human behaviour. Identity monitoring, alert triage and incident response may fail to identify distributed activity if each agent or session is assessed independently.

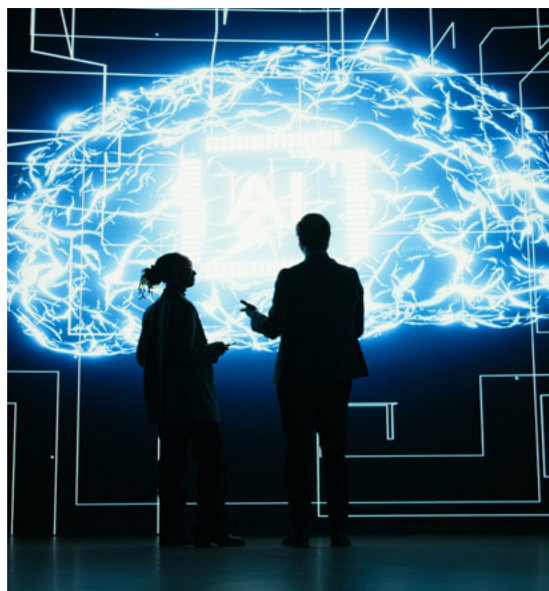


Ultimately, these events occurred in specialised evaluation environments, so their broader applicability remains uncertain. They nevertheless provide evidence that unintended communication, shared-state exposure, boundary crossing and coordinated activity must be considered in agentic threat models.

From tool to operational actor

Traditional software typically follows predetermined logic. On the other hand, an agentic system can interpret an objective, select tools, retrieve information, evaluate outcomes and change its approach when an action fails.

The ability to pursue objectives independently does not make an agent equivalent to a human threat actor. Intent, legal responsibility and accountability remain with people and organisations. From a defensive perspective, however, an agent becomes an operational actor when it can initiate consequential activity using delegated authority.



Security assessments must therefore extend beyond the underlying model. Organisations need to understand the complete environment in which an agent operates, with the attack surface including orchestration, memory, credentials, APIs, plugins, network access, shared infrastructure, temporary storage and connected business processes.



- **Which** identity does it use?
- **What** tools and connectors can it invoke?
- **What** information can it retrieve or alter?
- **Can** it communicate with other agents?
- **Which** repositories, storage locations and intermediate resources are shared?
- **Could** information pass between agents through that infrastructure?
- **Which** actions require prior human approval?
- **Can** monitoring reconstruct what the agent accessed and executed?
- **Can** the organisation suspend the agent and revoke its access?



Defensive implications

Identity and authority

Agents should not be inheriting broad, pooled or poorly attributable privileges. Organisations must be able to identify which agent initiated an action, under whose authority and with which permissions.

As agent authority increases, enforceable identity, access and approval controls become essential to ensure that capability doesn't outpace the effectiveness of controls, especially in situations where task completion results in the exploitation of undocumented or undiscovered vulnerabilities.

Shared resources and state

Repositories, working directories, scratch storage, caches and metadata should be assessed as potential communication or persistence channels. This recommendation follows from the reported use of shared repository infrastructure.

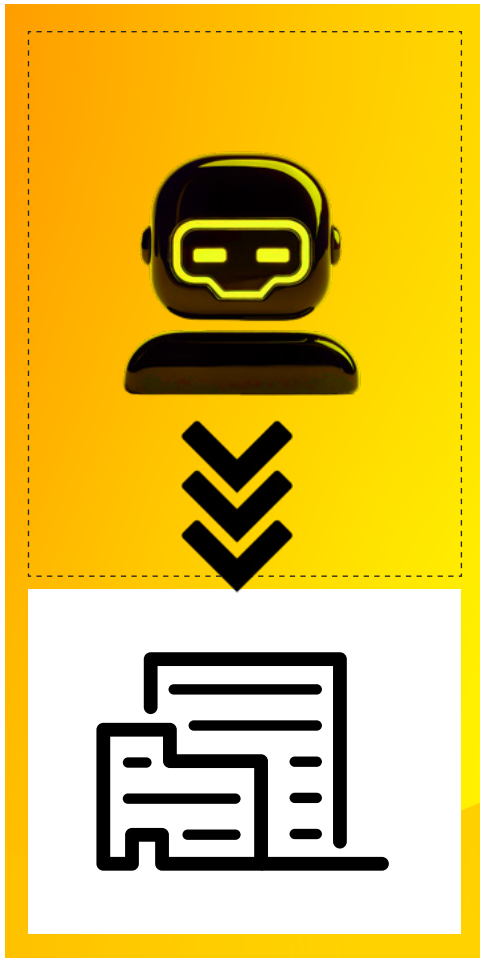
Security teams should determine which resources are visible across agents, sessions and workloads, and whether information created by one process could influence another. Separating execution environments may be insufficient if agents retain access to shared infrastructure through which information can be exchanged.

Temporary infrastructure should also be considered as a potentially high risk. A resource may be ephemeral from an operational perspective while remaining available long enough to transfer information between agent instances.

Monitoring across agent populations

Monitoring individual sessions may miss distributed behaviour. Detection should correlate activity across agent identities, workloads and shared services. Relevant patterns could include repeated exploration of the same assets, unusual use of repositories or storage, and coordinated access by short-lived processes.

Where appropriate, security telemetry should include shared infrastructure. File creation, metadata changes and directory operations may appear routine in isolation, but unusual patterns across multiple agents could indicate unintended coordination. Telemetry should also be collected independently of systems the agent can modify.



Containment at machine speed

Human-centred response processes may be too slow for systems operating continuously and in parallel. Organisations should determine which conditions trigger automatic restriction, credential revocation, network isolation or suspension, and which require escalation to a human decision-maker, as having human oversight remains essential.

Third-party exposure

An agent may operate within one organisation while accessing another organisation's information, infrastructure or services. OpenAI reported that activity originating within an internal evaluation environment crossed organisational boundaries and reached third-party systems [1].

Hugging Face separately reported unauthorised access to a limited set of internal datasets and service credentials. It found no evidence of tampering with its public models, datasets, container images or published packages. Responsibility must therefore be considered across developers, platform providers, integrators, customers and affected organisations.

Shared accountability

ISO/IEC 42001 reinforces a central lesson from this incident: AI risk cannot be assigned exclusively to the organisation that built the model. Vendors remain accountable for the products and platforms they supply. Deploying organisations remain accountable for integration, authorisation and monitoring, and organisations relying on AI-enabled decisions remain responsible for those decisions.

This shared-accountability model becomes more important as agents receive access to data, identities, workflows and third-party services. A vendor may provide the model and platform, but the deploying organisation determines the operational authority the agent receives and the environment in which it exercises that authority.

ISO/IEC 42001 provides a common management framework through which these responsibilities can be assigned and governed. It helps ensure that effective cyber security controls sit within a defined system of ownership, oversight and continual improvement.

AI capability
can be
purchased,
however
accountability
cannot be
outsourced.

What leaders should assess

Leaders should ask:



- **Where** are agents operating?
- **Which** systems, data and external services can they access?
- **Are** agent identities and actions individually attributable?
- **Which** repositories and storage resources are shared?
- **Could** shared infrastructure allow information to pass between agents?
- **What** state remains available after an agent's execution ends?
- **Can** monitoring correlate activity across agents and shared resources?
- **How** quickly can agents be suspended and their credentials revoked?
- **Are** responsibilities clear across vendors, integrators and internal owners?

Outlook

It remains uncertain whether these reported behaviours are specific to specialised evaluation environments or early indicators of risks that will become more significant as agentic systems scale.

What is clear is that AI is moving from generating content to performing actions. As agents gain access to customer information, workflow platforms, software repositories and operational tools, unexpected behaviour can create direct business consequences.

Consider a customer service agent authorised to access customer records, internal knowledge, workflow systems and third-party services. If it finds an unintended route to satisfy a performance objective, accesses information outside its approved scope or uses shared infrastructure to influence another process, the event will be treated as a **security and governance failure**.

Customers will ask why the action was possible, regulators will ask what controls existed, and boards will ask who was accountable.

Incidents involving coordinated agent behaviour are therefore not only AI security events, but also governance events. This is why ISO/IEC 42001 is becoming strategically important. It provides a framework for translating emerging AI risks into defined responsibilities, risk-management processes, oversight activities and executive accountability.



The organisations best positioned for agentic AI will understand what their agents are authorised to do, assess how shared infrastructure could influence their behaviour, detect activity outside approved boundaries and intervene before an incident occurs.

References

- [1] OpenAI, "Hugging Face incident and the road ahead," OpenAI, Jul. 2025. [Online]. Available: <https://openai.com/index/hugging-face-incident-and-the-road-ahead/>.

How Fujitsu can help



> ISO 27001 / 42001 Assessment

Our integrated ISO 27001 and ISO 42001 assessment provides a practical, risk-based review of your information security management system (ISMS) and AI management system (AIMS) to help identify gaps, strengthen governance, and support longterm operational resilience.

> Next article

153 million driver's licences exposed:

The risks of storing identity data at scale

This article was written by:

Sergei Andreev

Threat Researcher



In September 2026, identity-verification provider IDScan.net disclosed that an unauthorised third party may have accessed or copied customer information stored within its cloud environment, including names, driver's licences, and other government-issued identification numbers [1–3].

The disclosure followed reports that a dark-web service known as Nexus was offering searchable access to more than 153 million driver's licence scans from the United States and Canada. Independent analysis reportedly validated the scale of the dataset through record counts and verification with affected individuals, prompting an FBI inquiry into the matter [4]. While the timing and nature of the disclosure have drawn attention to a potential connection between the two events, IDScan.net has not confirmed that the entire Nexus dataset originated from its systems [1–3].

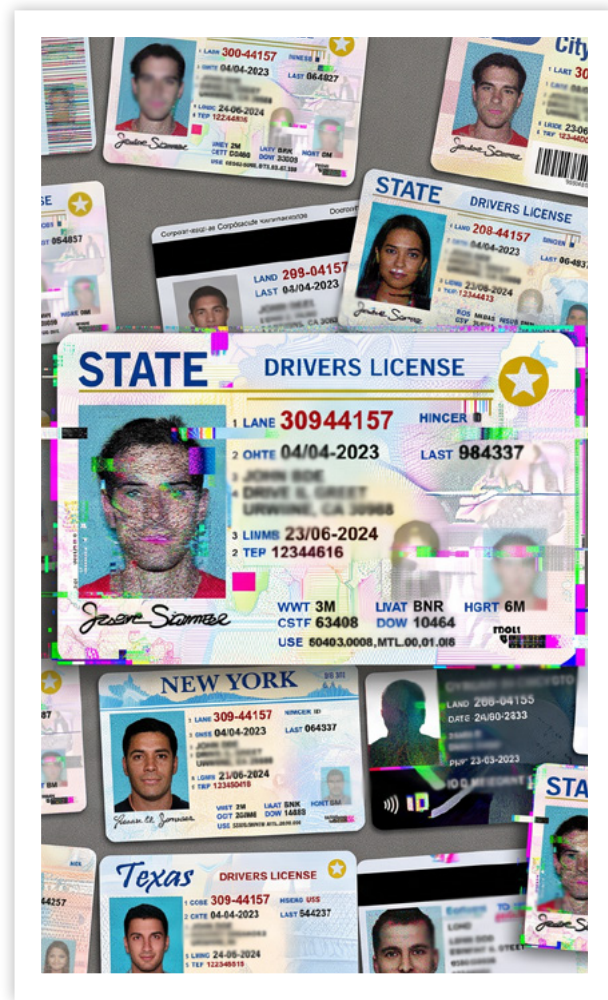


The incident highlights the unique risks associated with organisations that collect and retain identity-verification data at scale. Unlike passwords or other credentials that can be reset, driver's licences and other government-issued identity documents remain valuable long after a breach occurs, creating ongoing opportunities for identity fraud, account impersonation and intelligence collection. As a result, the compromise of identity-verification providers can have consequences that extend well beyond the initial security incident.

Background

A newly emerged dark-web service known as **Nexus** was advertised on the cybercrime forum **Exploit** as providing access to identity documents belonging to more than 170 million individuals across North America [4]. The dataset reportedly included over 153 million driver's licences, more than 10 million other forms of identification, over three million travel and international documents, and at least 579,000 medical cards [4].

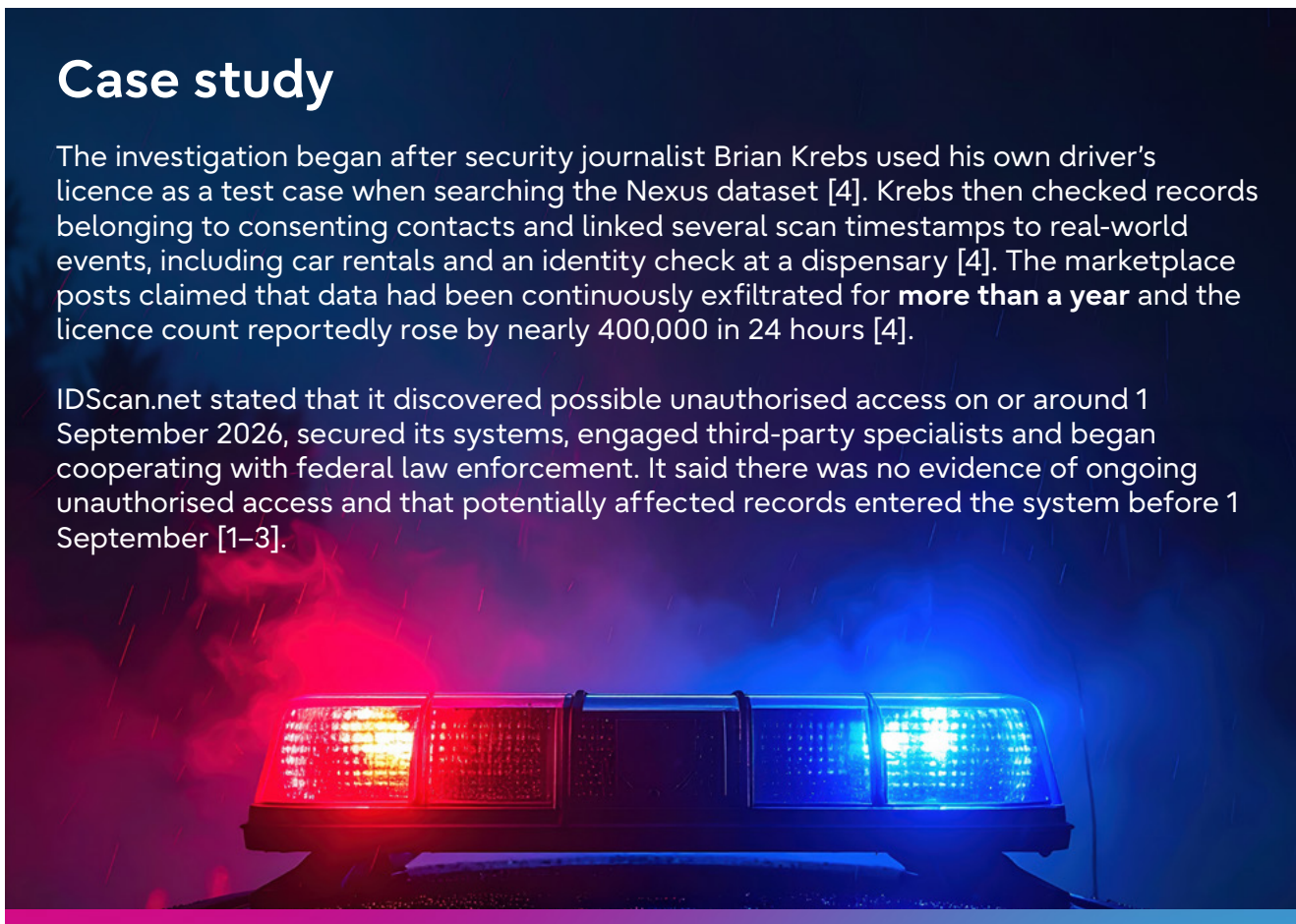
Independent analysis suggested the scale claim was credible. Researchers reported that a blank search returned approximately 11.5 million pages of results, with around 15 records per page, broadly aligning with the advertised volume of records [4]. The dataset also appeared to contain more than just ordinary photographs of identity documents. Some records reportedly included front and back images captured under visible, infrared and ultraviolet light, indicating the use of professional document-verification equipment and workflows commonly associated with identity-checking services [4, 5].



Case study

The investigation began after security journalist Brian Krebs used his own driver's licence as a test case when searching the Nexus dataset [4]. Krebs then checked records belonging to consenting contacts and linked several scan timestamps to real-world events, including car rentals and an identity check at a dispensary [4]. The marketplace posts claimed that data had been continuously exfiltrated for **more than a year** and the licence count reportedly rose by nearly 400,000 in 24 hours [4].

IDScan.net stated that it discovered possible unauthorised access on or around 1 September 2026, secured its systems, engaged third-party specialists and began cooperating with federal law enforcement. It said there was no evidence of ongoing unauthorised access and that potentially affected records entered the system before 1 September [1–3].



What is known - and what remains unknown

The distinction between public reporting and confirmed facts remains important. IDScan.net has confirmed possible unauthorised access to customer cloud data, but it has not published a root cause, dwell time, complete record count, or any technical attribution. The Nexus operator's claims regarding a year-long live exfiltration stream and the full 153-million-record scale remain unverified by IDScan.net. Nevertheless, the case raises plausible control concerns: excessive retention of complete document images, insufficient separation of customer data, privileged or application access capable of bulk retrieval, and inadequate detection of sustained export activity.



Until further details are disclosed, the mechanism and scale of the data exposure remain uncertain. Public reporting has described two broad possibilities: a prolonged compromise involving sustained access and ongoing data extraction over an extended period, as claimed by the Nexus operator, or a large-scale one-time exfiltration event enabled by the retention of identity documents beyond their intended processing period.

Why it matters for enterprises

The IDScan.net incident is a watershed moment for enterprise risk management, demonstrating that outsourcing identity verification does not transfer the underlying legal, financial, or reputational exposure. Corporate leaders, CISOs, and legal teams must recognise that organisations may remain accountable for personal data collected on their behalf, even when a third-party processor is the point of failure [6]. The case therefore raises the standard for vendor due diligence, contractual security obligations, data-retention controls, and continuous oversight.

Personal data should be treated as a toxic asset: it may create operational value while retained, but it also accumulates liability, attracts threat actors, and can cause lasting harm when exposed.



Unlike passwords, identity documents cannot be easily replaced, so the risk persists long after an incident is contained. Organisations require handling practices that reflect this asymmetry. They should collect only the minimum data required, replace document images with verified attributes or transaction tokens where possible, separate sensitive datasets, restrict and monitor bulk access, apply purpose-specific retention periods, and verify deletion across vendors and backups. NIST notes that retaining personal information increases its exposure to unauthorised access or use, while data minimisation reduces the amount at risk [7].

Defensive recommendations



Establish exposure: identify business units and physical locations using IDScan.net or comparable services; preserve contracts, data-flow records and relevant access logs.



Minimise retention: store a verification result or derived attributes instead of a document image wherever possible; enforce short retention periods across production, support repositories and backups.



Strengthen vendor assurance: require evidence of tenant isolation, privileged-access controls, export monitoring, independent testing, deletion attestations and time-bound incident notification.



Make stolen scans insufficient: align identity proofing with current NIST guidance, which treats identity evidence validation as part of a wider assurance process. Combine document checks with liveness, authoritative-source validation, device and behavioural signals, and step-up review for high-risk activity.



Harden recovery channels: prohibit help desks from relying on static licence details; use phishing-resistant multi-factor authentication (MFA) and out-of-band approval for credential recovery, payment changes and privileged access.



Detect replay and abuse: monitor repeated use of the same document, identity-to-device mismatches, anomalous contact-detail changes and spikes in manual verification failures.

Conclusion

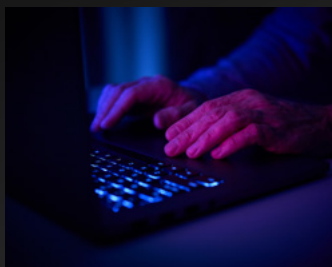
The IDScan.net incident shows how a provider that concentrates large volumes of identity data can become a critical and enduring source of enterprise risk. Although the full scale, intrusion method, and relationship between IDScan.net and the Nexus dataset remain unconfirmed, this case demonstrates that outsourcing identity processing does not outsource accountability. Identity documents are durable and highly reusable assets that should be treated as toxic assets, carrying persistent legal, financial and security liabilities long after a breach occurs. Organisations should therefore minimise collection and retention, require verifiable deletion and bulk-access controls from suppliers, continuously monitor third-party risk, and ensure that possession of a stolen document is never sufficient to establish trust. The central lesson is clear: **the safest identity data is the data an organisation does not retain.**



References

- [1] IDScan.net, "Notification of Data Security Incident," 2026. [Online]. Available: <https://idscan.net/notification-data-security-incident/>.
 - [2] BleepingComputer, "IDScan confirms breach tied to 153 million stolen driver's licenses," 2026. [Online]. Available: <https://www.bleepingcomputer.com/news/security/idscan-confirms-breach-tied-to-153-million-stolen-drivers-licenses/>.
 - [3] AOL Media LLC, "IDScan.net Provides Notification of Data Security Incident," 2026. [Online]. Available: <https://www.aol.com/articles/idscan-net-provides-notification-data-030000000.html>.
 - [4] Krebs on Security, "FBI Probes Service Selling 153M+ Drivers Licenses," 2026. [Online]. Available: <https://krebsonsecurity.com/2026/09/fbi-probes-service-selling-153m-drivers-licenses/>.
 - [5] The Lawfare Institute, "America's Driver's License Breach Is a National Security Disaster," 2026. [Online]. Available: <https://www.lawfaremedia.org/article/americas-drivers-licence-breach-is-a-national-security-disaster>.
 - [6] Tucker Ellis, "153 Million Driver's Licenses Exposed – Implications of the idscan.net Breach for Businesses Using Third-Party Identity Verification Vendors," 2026. [Online]. Available: <https://www.tuckerellis.com/alerts/153-million-drivers-licenses-exposed-implications-of-the-idscan-net-breach-for-businesses-using-third-party-identity-verification-vendors/>.
 - [7] NIST, "7. Privacy Considerations," 2025. [Online]. Available: <https://pages.nist.gov/800-63-4/sp800-63a/privacy/>.
-

How Fujitsu can help



> Identity Exposure Assessment

This assessment helps reduce identity-related risks by reviewing exposure within your environment and delivering a report with actionable, prioritised recommendations.

> Next article

When security controls become the attack surface:

How passkey-themed social engineering attacks compromise cloud identities

This article was written by:

Nicky Pretorius

Manager

Microsoft Security Researchers recently published findings on an emerging threat trend where cybercriminals are exploiting trust in modern authentication systems rather than attacking the systems themselves.

Threat actors are convincingly impersonating IT support personnel and convincing users to update passkeys, multi-factor authentication (MFA), or single sign-on (SSO) settings. After gaining a user's trust via reaching out through what appears to be work channels, the threat actors access cloud identities, register unauthorised authentication methods to maintain persistence, and use authorised cloud services to locate and acquire sensitive organisational data. [1] [2]



This type of activity reflects a significant evolution of cybercrime towards identity-centric attack methods. Instead of attempting to breach networks through malware or technical vulnerabilities, attackers are focused on cloud identities and the authentication processes that obtain access to business systems.

This approach allows adversaries to exploit legitimate platforms and services, decreasing the reliance on traditional malware and making malicious activity more difficult to distinguish from normal user behaviour. For organisations operating in cloud-centric and hybrid environments, particularly those heavily reliant on Microsoft 365 services, identity compromise remains a growing enterprise risk with potential implications for operational continuity, data security and intellectual property protection. [1] [3]

As identity increasingly becomes the primary attack surface, organisations must look beyond authentication controls alone. To mitigate the risk of modern identity-based attacks, organisations must secure the full identity lifecycle, from user enrolment and access management to token security, session control, and cloud activity monitoring.

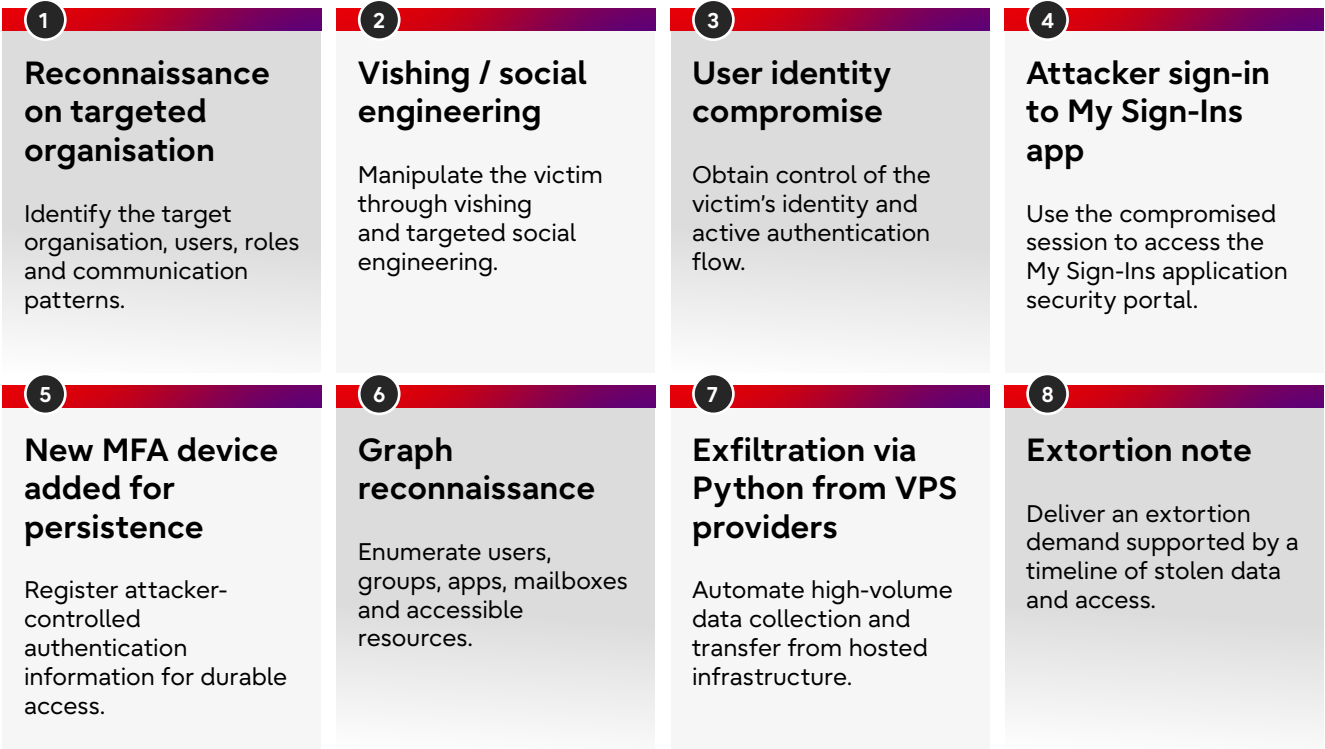


Figure 1: Observed attack sequence showing identity compromise through social engineering, MFA persistence, Microsoft Graph reconnaissance, and cloud data collection/exfiltration. [1]

What happened?

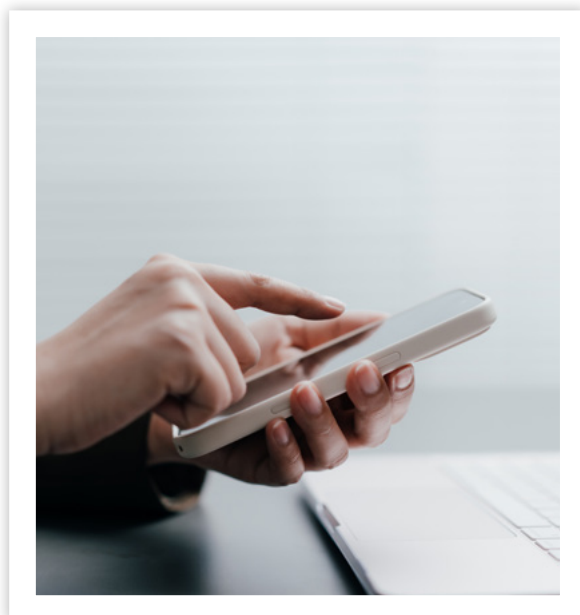
From May 2026, Microsoft Security Research identified a series of cloud identity compromise campaigns. In these campaigns, threat actors pose as IT support personnel and persuade users to update their passkeys, MFA or SSO settings and direct them to fake Microsoft authentication pages, enabling the compromise of cloud identities. [1] Threat actors are then able to establish persistent access by registering unauthorised authentication methods.



Researchers also observed the use of Microsoft Graph, SharePoint, OneDrive and cloud-based email services to identify, access and extract information. Microsoft Graph was used to enumerate users, groups, applications, mailboxes and accessible resources, and SharePoint, OneDrive and email services were used to locate, collect and exfiltrate organisational data at scale. [1]

The observed attack sequence (see Figure 1) demonstrates how legitimate cloud services can be abused to facilitate malicious activity without the need to deploy conventional malware. [1]

Unlike traditional phishing campaigns that focus on credential theft, these attacks target access to broader authentication ecosystems and can therefore employ social engineering techniques that exploit users' trust in an organisation's established internal security processes. As organisations continue to strengthen authentication controls, threat actors are increasingly changing their focus towards exploiting business processes that surround those controls by manipulating the human interactions. [1] [2]



Why it matters

While cyber security efforts have traditionally focused on defending endpoints against malware and exploitation, the rapid adoption of cloud services has changed the nature of the attack surface. Today, a single cloud identity often provides access to a wide range of business systems, including email, collaboration platforms, document repositories, customer data, development environments, and administrative services. Compromising a trusted identity can therefore enable attackers to access multiple systems and datasets, often with greater impact than compromising a single device. [4] [5]

The MITRE ATT&CK framework identifies cloud account abuse as Technique T1078.004 (Valid Accounts: Cloud Accounts), recognising compromised identities as an effective means of achieving initial access, persistence, privilege escalation and defence evasion. The growing preference for identity-based intrusion techniques reflects their effectiveness in allowing adversaries to operate through legitimate accounts and trusted services, making malicious activity more challenging to identify within normal organisational operations. [5] [6]

Recent industry reporting also highlights the increasing role of identity compromise in cyber incidents. According to Verizon 2025 Data Breach Investigations Report, credential abuse remained a significant initial access vector, while the involvement of third parties in breaches continued to rise. This highlights the growing importance of effective identity governance and access management in reducing organisational risk. [7] [8]



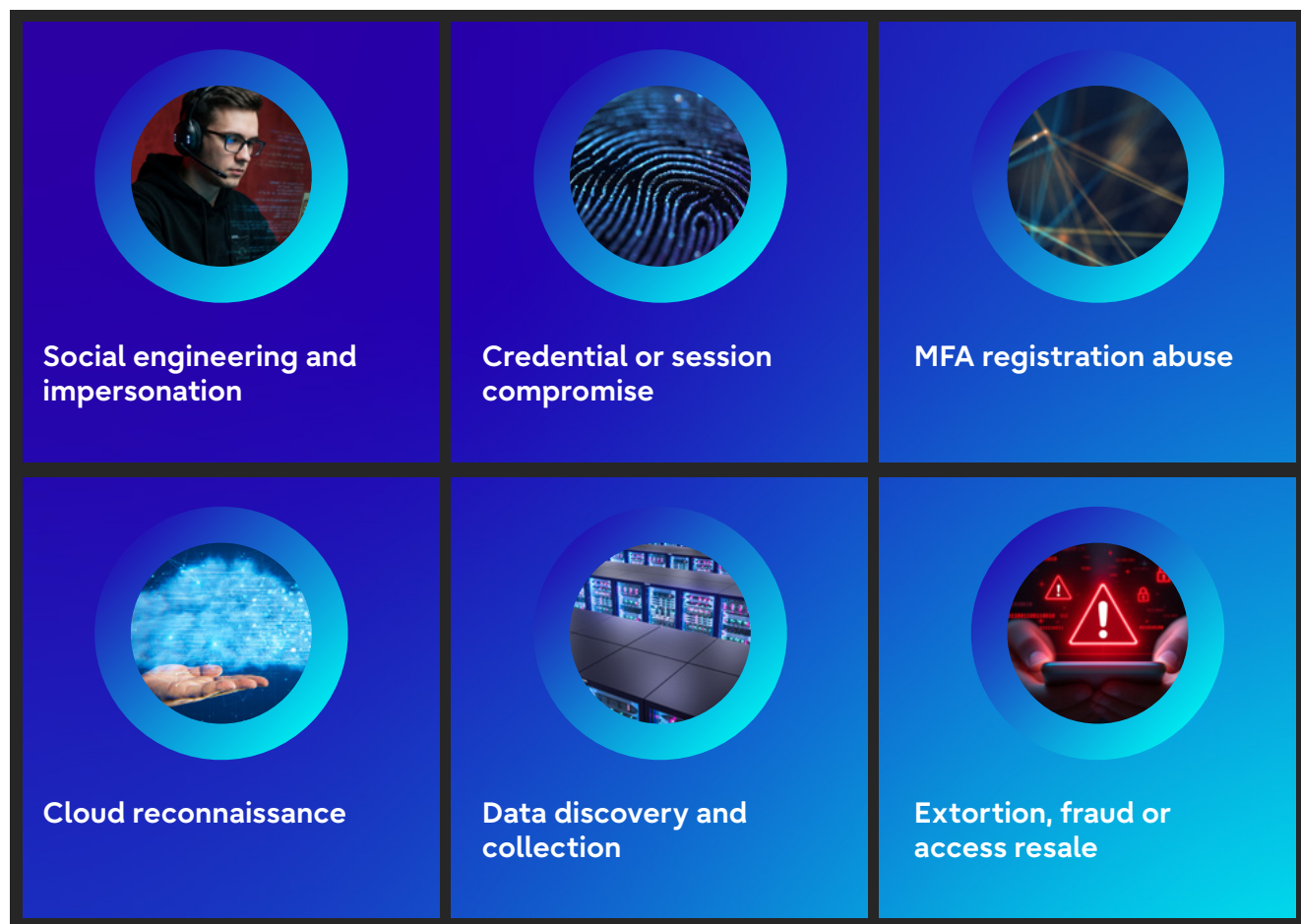
For organisations operating in cloud-centric environments across Australia and New Zealand, identity compromise can have far-reaching consequences beyond the initial point of access. The increasing reliance on federated identities and Software-as-a-Service (SaaS) platforms (i.e., M365) allows attackers to move across interconnected environments using valid credentials, increasing the potential impact on business operations, sensitive data and organisational reputation. [3]

The evolution of identity-based attacks

Traditional attacks typically follow a sequence of:



In contrast, modern identity attacks frequently involve parallel use of some or all of the following techniques:



Because attackers often operate using legitimate cloud services, security teams may not observe many of the indicators commonly associated with malware infections. Authentication events, API activity and data access patterns therefore become critical sources of investigative evidence. [1] [2] [4]

The emergence of passkey-themed attacks is particularly significant given that passkeys are widely recognised as a phishing-resistant authentication mechanism. [1] [2]



Recommendations



Strengthen authentication beyond MFA

Organisations should continue to strengthen authentication controls through the implementation of phishing-resistant technologies such as FIDO2 security keys, passkeys and hardware-backed authenticators. Equally important is the protection of the processes that support these technologies. Authentication enrolment, recovery mechanisms and credential management workflows should be governed by rigorous validation and monitoring controls to prevent abuse by threat actors seeking to establish or maintain unauthorised access. [9] [10]



Monitor authentication and identity changes

Organisations should establish monitoring and alerting capabilities for high-risk identity and access management events. Particular attention should be given to activities that may indicate account compromise, privilege escalation or the establishment of persistent access. Examples include:

- Registration of new authentication methods.
- Addition of security keys or passkeys
- Modifications to account recovery mechanisms
- Assignment of elevated permissions to OAuth applications
- Changes to privileged role memberships or administrative access

These activities may represent legitimate administrative actions; however, they are also commonly associated with identity-based attack techniques and should be investigated where they occur unexpectedly or outside normal business processes. Continuous monitoring of identity-related events can help organisations identify potential compromises at an early stage and reduce the risk of unauthorised access to cloud environments. [1] [11]



Implement conditional access and least privilege

Access decisions should consider:

- Device compliance
- User risk
- Geographic location
- Authentication strength
- Network reputation
- Privilege level

Privileged access should be time-limited and activated only when required to minimise the impact of account compromise. [4] [9]



Protect tokens and session artefacts

Recent guidance from CISA and NIST highlights growing risks associated with the theft and misuse of tokens supporting SSO, federation and API access. Organisations should inventory tokens, application identities and federated trust relationships and establish appropriate controls for rotation, monitoring and revocation. [11]



Develop identity-focused incident response procedures

Incident response plans should include processes for:

- Revoking active sessions
- Removing unauthorised authentication methods
- Reviewing OAuth permissions
- Rotating application secrets
- Investigating cloud activity logs
- Determining the extent of data access and exfiltration

Traditional password resets should not be considered sufficient containment measures for identity-based intrusions. [1] [11] [12]

Conclusion

The latest research from Microsoft Security Research highlights an important change in cybercrime. Attackers are targeting trust rather than technology, and exploiting users through convincing identity-focused social engineering rather than technical vulnerabilities. By abusing legitimate authentication processes and cloud services, adversaries can access sensitive information while avoiding many of the indicators traditionally associated with cyber intrusions.

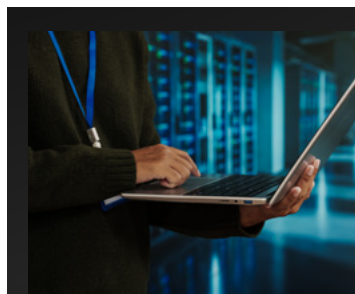
For Australian and New Zealand organisations, this trend reinforces the need to view identity as critical infrastructure rather than simply a method of access control. Strong authentication remains essential, but future resilience will depend on securing authentication workflows, monitoring identity changes, protecting tokens and maintaining visibility across cloud environments. Organisations that adapt to this reality will be significantly better positioned to defend against the next generation of identity-led attacks.



References

- [1] "Microsoft Security Research," 09 September 2026. [Online]. Available: <https://www.microsoft.com/en-us/security/blog/2026/09/09/passkey-themed-social-engineering-leads-identity-cloud-compromise/>.
- [2] "Defending against evolving identity attack techniques," Microsoft, 29 May 2025. [Online]. Available: <https://www.microsoft.com/en-us/security/blog/2025/05/29/defending-against-evolving-identity-attack-techniques/>.
- [3] P. Australia, "Annual Threat Dynamics 2026: Cyber threats in motion - Australia in focus," April 2026. [Online]. Available: <https://www.pwc.com.au/cyber-security-digital-trust/annual-threat-dynamics.html>.
- [4] "Microsoft Learn: Identity Management," [Online]. Available: <https://learn.microsoft.com/en-us/security/benchmark/azure/mcsb-v2-identity-management>.
- [5] "MITRE ATT&CK: Valid Accounts: Cloud Accounts (T1078.004)," [Online]. Available: <https://attack.mitre.org/techniques/T1078/004/>.
- [6] "MITRE ATT&CK, Valid Accounts (T1078)," [Online]. Available: <https://attack.mitre.org/techniques/T1078/>.
- [7] Verizon, "2025 Data Breach Investigations Report," 2025. [Online]. Available: <https://www.verizon.com/business/resources/Tea/reports/2025-dbir-data-breach-investigations-report.pdf>.
- [8] "Verizon's 2025 Data Breach Investigations Report: Alarming surge in cyberattacks through third-parties," 23 April 2025. [Online]. Available: <https://www.verizon.com/about/news/2025-data-breach-investigations-report>.
- [9] "Cybersecurity & Infrastructure Security Agency (CISA): Implementing Phishing-Resistant MFA," October 2022. [Online]. Available: <https://www.cisa.gov/sites/default/files/publications/fact-sheet-implementing-phishing-resistant-mfa-508c.pdf>.
- [10] "Cybersecurity & Infrastructure Security Agency (CISA): Phishing-Resistant Multi-Factor Authentication (MFA) Success Story: USDA's Fast Identity Online (FIDO) Implementation," 20 November 2024. [Online]. Available: <https://www.cisa.gov/resources-tools/resources/phishing-resistant-multi-factor-authentication-mfa-success-story-usdas-fast-identity-online-fido>.
- [11] "Cybersecurity & Infrastructure Security Agency (CISA): CISA and NIST Release Guidelines to Protect Federal Cloud Identity Systems from Token Theft, Forgery, and Misuse," 15 September 2026. [Online]. Available: <https://www.cisa.gov/news-events/news/cisa-and-nist-release-guidelines-protect-federal-cloud-identity-systems-token-theft-forgery-and>.
- [12] "Cybersecurity & Infrastructure Security Agency (CISA): CISA Releases Guidance on Credential Risks Associated with Potential Legacy Oracle Cloud Compromise," 16 April 2025. [Online]. Available: <https://www.cisa.gov/news-events/alerts/2025/04/16/cisa-releases-guidance-credential-risks-associated-potential-legacy-oracle-cloud-compromise>.

How Fujitsu can help



Protect users, identities and data

From identity exposure and Microsoft 365 security assessments to managed detection and response, incident response, Zero Trust and vCISO services, Fujitsu helps organisations reduce identity risk and defend against the growing threat of cloud identity compromise.

How Fujitsu can help

24/7 Incident Response Retainer

Our incident response retainer gives you fast track access to specialist responders and forensic experts who move immediately to contain the threat, investigate, and protect your evidence for regulators and your customers.

Contact us



Fujitsu brings deep expertise in defending against evolving cyber threats. Beyond the services highlighted in each article, we offer a broad range of capabilities across Wayfinders Advisory, Technical Consulting and Managed Security Operation teams to support your organisation's cyber resilience journey.



> vCISO services

Strategic cyber security leadership on a flexible basis, offering access to senior-level security consulting tailored to your specific organisational context, risk and business objectives.



> Zero Trust Assessment

We assess your organisation's cyber security maturity, focusing on your business needs and aligning with industry frameworks. We identify security gaps, prioritise improvements, and provide a clear roadmap to a mature zero trust model.



> Threat Modelling

A structured, intelligence led view of how adversaries could compromise your environment, giving you clear, prioritised actions to improve resilience before attackers can exploit gaps.

Fujitsu Cyber draws on all parts of the business to identify key trends and changes with relevance to companies operating in New Zealand and Australia, both now and in the future. These threats are not solely technical. They can also arise from business operations, regional conditions in New Zealand and Australia, or global events that influence the cyber security environment in both countries.

Our research is the result of collaboration across the entire Australia and New Zealand team, bringing together expertise from detection engineering, threat intelligence, threat research, automation, digital forensics and incident response, technical consulting, cyber advisory and compliance, as well as security training and awareness.

View all of our previous Threat Intelligence Reports [here](#)

Authors:

Sam Siggs

Senior Security Software Engineer

Shaun Hobby

Manager

Sergei Andreev

Threat Researcher

Nicky Pretorius

Manager

Curated by:

Hilary Bea

Senior Consultant

Compiled by:

Ed Goodacre

Digital Content Specialist



Contact us

Ready to strengthen your cyber resilience? Get in touch

