CSAI Foundation

Cloud Security Alliance AI Safety Initiative

OAuth Device Code Phishing: 37x Surge in Enterprise ATO

PhaaS Tooling Commoditizes Device Authorization Grant Abuse Against Microsoft 365

Unofficial AI-assisted Research

2026-04-05

Key Takeaways

- Push Security's threat research team documented a 37.5x surge in device code phishing pages by April 4, 2026, attributed to the commercial launch of EvilTokens — a Phishing-as-a-Service platform that appeared on Telegram in mid-February 2026 and within weeks had lowered the skill barrier for conducting token-theft campaigns at scale [1].
- Device code phishing exploits a legitimate OAuth 2.0 standard (RFC 8628), meaning no vulnerabilities are patched and no indicators of compromise appear in traditional credential-theft telemetry — no stolen credentials, no intercepted session cookies. The attack does, however, generate detectable `DeviceCodeFlow` events in Microsoft Entra ID Sign-In Logs that organizations can alert on; the victim completes their own MFA challenge, and the authentication succeeds on Microsoft's real infrastructure [2, 6].
- A single EvilTokens campaign active from February 19 through March 2026 compromised more than 340 Microsoft 365 organizations across the United States, Canada, Australia, New Zealand, and Germany, spanning construction, financial services, healthcare, government, and manufacturing sectors [3].
- Russia-aligned state actors — including Storm-2372, UTA0304, UTA0307, and UNK_AcademicFlare — adopted device code phishing against governments, defense entities, NGOs, and energy organizations across Europe, North America, Africa, and the Middle East beginning in August 2024; financially motivated actors followed in October 2025 [4, 5, 13].
- Harvested refresh tokens persist for up to 90 days and self-renew on each use; in advanced attack scenarios, attackers convert them into Primary Refresh Tokens (PRTs) that enable single sign-on across all Microsoft 365 services and survive password resets [3, 5].
- The highest-priority mitigation for organizations that do not legitimately use device code flow is a Microsoft Entra ID Conditional Access policy that explicitly blocks the Device Code Flow authentication condition — a control Microsoft itself began automatically provisioning for eligible tenants in February 2025 [7, 8].

Background

The OAuth 2.0 Device Authorization Grant, codified in RFC 8628, was engineered to solve a specific usability problem: how does a resource-constrained device — a smart television, a CLI tool, a printer — authenticate a user when it cannot render a full browser session or accept keyboard input? The protocol's answer is an elegant delegation: the constrained device requests a short-lived code from the authorization server, the user is directed to a companion URL on a device with a browser (typically a smartphone or laptop), and after the user authenticates and approves the request there, the constrained device receives its tokens through a polling loop. Microsoft's implementation exposes this flow through the `microsoft.com/devicelogin` endpoint, making it recognizable to most enterprise users who have authenticated to Teams on a television or configured Azure CLI on a headless server [2, 6].

The protocol's security model depends on a critical assumption: that device codes reach users through out-of-band, trustworthy channels — displayed on the screen of the device requesting authentication. When a device code arrives instead in a phishing email, the entire trust chain inverts. The attacker, not a legitimate device, has initiated the authorization request. The attacker controls the polling loop. The victim, presented with a familiar-looking prompt on a familiar Microsoft domain, completes every authentication step normally — including whatever MFA factors their organization requires. From the authorization server's perspective, the session is entirely legitimate. From the victim's perspective, nothing unusual has occurred. The attacker receives valid access and refresh tokens the moment the victim clicks approve [4, 5].

Security researchers identified this abuse potential as early as 2021 [6], and Secureworks published proof-of-concept tooling (SquarePhish) in 2022 documenting the attack chain in detail. For several years, device code phishing remained a niche technique — operationally straightforward but requiring manual setup and lacking the commoditized delivery infrastructure that drives phishing campaigns at scale. That constraint effectively dissolved in February 2026. The emergence of EvilTokens as a subscription Phishing-as-a-Service platform — offering pre-built Microsoft 365 phishing pages, AI-generated lure content, email harvesting, BEC automation, and a built-in webmail interface — removed the last friction point. By April 4, 2026, Push Security's detection infrastructure had recorded a 37.5x increase in device code phishing pages over the pre-2026 baseline [1, 9].

Security Analysis

The PhaaS Commoditization Curve

The surge in device code phishing is not a story about a new vulnerability or a novel technique — it is a story about infrastructure maturity. EvilTokens, operated by `eviltokensadmin` via Telegram and analyzed in depth by Sekoia's Threat Detection and Research team in March 2026, represents the point at which device code phishing crossed from skilled adversary tradecraft into commodity tooling [9, 10]. Affiliates who subscribe to the platform receive a complete attack stack: phishing page templates impersonating Adobe Acrobat Sign, DocuSign, SharePoint, OneDrive, calendar invitations, voicemail notifications, eFax, password expiry notices, and email quarantine alerts; AI-powered lure generation capable of tailoring social engineering copy to a target's role or organization; anti-analysis protections (right-click disabled, developer tools blocked, window-size heuristics); and a management console for harvested tokens [1, 9].

Push Security's April 2026 analysis identified at least eleven distinct phishing kits driving the surge alongside EvilTokens, including VENOM, SHAREFILE, CLURE, LINKID, AUTHOV, DOCUPOLL, FLOW_TOKEN, PAPRIKA, DCSTATUS, and DOLCE [1]. This trajectory resembles the commoditization pattern observed in credential phishing markets — a period of actor-specific tooling followed by dominant platform emergence, affiliate networks, and competitive differentiation on features rather than on fundamental technique — though the precise analogy depends on how market and law enforcement dynamics evolve. Sekoia assessed with high confidence that EvilTokens would become a major competitor in the phishing and BEC market, with planned expansion to Gmail and Okta targets already documented in platform communications as of March 2026 [9].

The delivery infrastructure for EvilTokens campaigns reflects the same sophistication. The March 2026 campaign that compromised more than 340 organizations routed phishing URLs through legitimate security vendor redirect systems operated by Cisco, Trend Micro, and Mimecast — exploiting the trust these domains carry in email security products — before forwarding through Cloudflare Workers and Railway.com infrastructure to attacker-controlled landing pages [3]. This layered redirection technique frustrates static URL blocklists and complicates forensic attribution.

Attack Chain and Token Persistence

The technical execution of a device code phishing attack requires no vulnerability and leaves no traditional compromise indicator. The attacker initiates an OAuth authorization request using a client ID that Microsoft recognizes as legitimate — the Microsoft Authentication Broker, Azure AD PowerShell, or

an attacker-registered Azure App Registration that has passed Microsoft's verification process. The authorization server returns a `device_code` for the attacker's polling loop and a human-readable `user_code` for victim interaction. The attacker's phishing infrastructure delivers the `user_code` to the victim embedded in a credible lure, with instructions to visit `microsoft.com/devicelogin` – a real Microsoft domain – to complete what the lure frames as routine authentication [4, 5, 6].

At `microsoft.com/devicelogin`, the victim enters the code and encounters Microsoft's authentic authentication flow. They enter their credentials, complete their MFA challenge – push notification, TOTP code, or Windows Hello – and approve the access request. Every step occurs on Microsoft infrastructure, under the victim's own account, with their own authenticators. Microsoft Entra ID logs record a successful, MFA-satisfied authentication. The attacker's polling loop, which has been checking the authorization endpoint every five seconds since the code was generated, receives valid `access_token` and `refresh_token` values the moment the victim clicks approve [4, 5].

The token properties that make this attack particularly damaging are well-documented but under-appreciated at the organizational level. Access tokens grant approximately 90 minutes of direct API access to Exchange Online, OneDrive, SharePoint, and Teams with no further authentication. Refresh tokens last 90 days and self-renew each time they are exchanged for a new access token – meaning a single successful compromise can provide persistent, silent access for months without triggering any additional authentication prompt. Advanced adversaries, notably Storm-2372 beginning in February 2025, exploit the Microsoft Authentication Broker client ID specifically because the resulting refresh tokens can be used to register attacker-controlled devices in Microsoft Entra ID and acquire Primary Refresh Tokens (PRTs). PRTs enable single sign-on across all Microsoft 365 services with no password or MFA challenge and critically survive password resets – a property that renders the most common first-response action ineffective against sophisticated actors [4, 5].

Threat Actor Landscape

Device code phishing's threat actor profile has expanded from state-sponsored elite units to financially motivated commodity operators over an 18-month period, a trajectory that closely parallels the historical commoditization of other MFA-bypass techniques. Microsoft disclosed Storm-2372 in February 2025, assessing with moderate confidence that the group aligns with Russian state interests and has been active since at least August 2024. The group's targeting spans governments, NGOs, IT service providers, defense contractors, telecommunications providers, healthcare organizations, higher education institutions, and energy and oil and gas companies across Europe, North America, Africa, and the Middle East [4].

Concurrent threat intelligence research from multiple vendors identified additional Russia-assessed clusters — UTA0304, UTA0307, and UNK_AcademicFlare — operating device code phishing campaigns against overlapping target sets [5, 13]. UNK_AcademicFlare is notable for its reconnaissance discipline: the cluster conducts extended rapport-building outreach with targets before delivering phishing lures, framing device code prompts as pre-meeting document review links and often operating through compromised government or military email accounts to avoid sender-reputation detection [5, 13]. APT29 (Midnight Blizzard/NOBELIUM), Russia's SVR-linked group, has also been associated with device code phishing against government and diplomatic targets [13].

The state actor phase demonstrated that device code phishing could be operationalized with high operational security — tight targeting, slow cadence, and compromised-account delivery infrastructure — before commodity actors adopted looser, volume-focused approaches. The financially motivated phase that followed reduced that operational barrier to a subscription fee. Proofpoint documented TA2723, a financially motivated actor, beginning device code phishing operations in October 2025 using salary document and shared file lures against enterprise Microsoft 365 users [5]. By February 2026, EvilTokens had made the same capability available to any affiliate willing to pay a monthly fee, without requiring any understanding of the underlying OAuth protocol [1, 9, 11, 12, 13].

Why MFA Does Not Stop This Attack

Organizations that have invested in phishing-resistant MFA as a control against credential theft should understand that device code phishing operates in a fundamentally different threat model. Traditional phishing and adversary-in-the-middle (AiTM) attacks intercept credentials or session cookies in transit. Device code phishing does neither. The authentication is real, the MFA completion is real, and the resulting tokens are legitimate grants issued by the authorization server. There is no stolen credential to detect, no session cookie to invalidate, and no lookalike domain to block. The attack succeeds precisely because the victim does everything correctly from their perspective.

This distinction matters for security controls design. FIDO2 hardware security keys and passkeys do provide meaningful protection against device code phishing, but not through the origin-binding mechanism that protects against traditional lookalike-domain phishing. In a device code attack, the victim authenticates at the real `microsoft.com/devicelogin`, so no origin mismatch occurs at the authenticator level. The protection, where present, operates through Microsoft Entra ID server-side policy: tenants configured to require phishing-resistant authentication methods can restrict which methods satisfy the device authorization grant flow, effectively preventing device code ceremony completion when FIDO2 is the mandated authenticator [7, 8]. Standard TOTP-based authenticator apps, SMS OTP, push notification approval, and certificate-based authentication methods that lack equivalent server-side flow restrictions do not provide this protection, as they authenticate the user but

do not constrain the destination of the resulting tokens. Organizations that have deployed TOTP or push-based MFA broadly and consider themselves phishing-resistant should reconsider that assessment in the context of device code phishing specifically.

Recommendations

Immediate Actions

The single highest-impact mitigation available to most organizations is a Conditional Access policy in Microsoft Entra ID that explicitly blocks the Device Code Flow authentication condition. Microsoft documented this control in the Entra ID Conditional Access documentation, and began automatically provisioning a managed version of this policy for eligible Microsoft 365 tenants in February 2025 [7, 8]. Organizations should verify whether this managed policy is active in their tenant and confirm it is in enforced — not report-only — mode. For tenants where it is absent or in report mode, creating and enforcing a blocking policy scoped to all users is the recommended immediate action.

For organizations that legitimately rely on device code flow — for IoT provisioning, specific CLI tools, or legacy integrations — a blocking policy with a narrow allow-list scoped to specific approved users, device compliance states, or named IP ranges is preferable to leaving the flow unrestricted. The granularity of the Conditional Access Authentication Flows condition supports this scoping without requiring a single organization-wide exception [8].

Incident response for suspected device code phishing compromise should begin with immediate token revocation using the `revokeSignInSessions` action in the Microsoft Graph API for affected accounts, followed by auditing Entra ID for unexpected device registrations — the primary indicator of PRT conversion — and reviewing OAuth app consent grants for unauthorized application registrations. Password resets alone are insufficient for accounts where PRT conversion has occurred.

Short-Term Mitigations

Entra ID Sign-In Logs provide visibility into Device Code Flow authentication events and should be reviewed for anomalies. Key indicators include `DeviceCodeFlow` authentication events originating from unexpected IP addresses, geographies inconsistent with the user's normal access patterns, or known EvilTokens infrastructure. Railway.com IP ranges 162.220.234.41, 162.220.234.66, and 162.220.232.57 are documented indicators from the March 2026 campaign and should be treated as point-in-time; Railway.com infrastructure can be repurposed rapidly after public disclosure, and

defenders should prioritize behavioral detection on `DeviceCodeFlow` log events rather than IP-based blocking for durable coverage [3]. Security information and event management (SIEM) rules that alert on device code flow authentications from users who have no documented legitimate need for the flow will surface compromises that would otherwise appear as normal successful logins.

User awareness training for device code phishing differs meaningfully from traditional anti-phishing guidance because there is no malicious link to identify and no credential entry box to avoid. The phishing indicator is the device code itself — specifically, a device code that arrives through email or messaging platforms rather than appearing on the screen of a device the user is trying to authenticate. Users should be trained to treat any device code received via email, Teams, Slack, or SMS as a strong phishing indicator requiring verification through a separate channel, regardless of how the surrounding lure is framed [4, 9, 14]. This behavioral norm runs counter to most users' existing mental models of authentication flows and requires deliberate reinforcement.

Organizations should audit their Azure AD App Registrations and OAuth application consent grants for applications using client IDs associated with device code flow abuse — particularly the Microsoft Authentication Broker client ID that Storm-2372 exploited to enable PRT conversion [4]. Any registrations that appear unfamiliar or lack business justification should be reviewed and removed. Disabling device code flow at the application registration level for all applications that do not require it provides a second layer of defense beneath the Conditional Access policy.

Strategic Considerations

The commoditization of device code phishing through platforms like EvilTokens signals that MFA bypass capability — once the domain of state actors with significant technical resources — is now accessible to any financially motivated adversary willing to pay subscription fees. Organizations whose identity security architecture was designed around the threat model of credential phishing should evaluate whether that architecture adequately addresses token theft as a distinct attack class. The critical difference is persistence: stolen credentials are remediated by password reset, while stolen refresh tokens and PRTs require token revocation plus device deregistration and may have already enabled lateral movement or data access before discovery.

Migrating high-value accounts and administrative users to phishing-resistant MFA methods — specifically FIDO2 hardware keys — addresses the root capability that makes device code phishing effective: the ability for an attacker to receive tokens for authentication they did not perform. Microsoft Entra ID's Certificate-Based Authentication (CBA) with hardware-bound certificates may offer equivalent protection if the tenant is configured to require phishing-resistant methods for device code flow completion, though organizations should verify this behavior in their specific Entra ID configuration before treating CBA as equivalent to FIDO2 for this threat scenario. While FIDO2 deployment requires

hardware acquisition and enrollment workflows, the threat landscape as of April 2026 justifies prioritizing this migration for accounts with access to sensitive data, administrative capabilities, or privileged communications.

The EvilTokens platform's documented plans to expand beyond Microsoft 365 to Google Workspace and Okta targets suggest that device code phishing will not remain a Microsoft-specific concern. Organizations with heterogeneous identity environments should evaluate whether device authorization grant flows are enabled and necessary in non-Microsoft identity providers and apply equivalent blocking controls where they exist.

CSA Resource Alignment

This research note connects directly to several Cloud Security Alliance frameworks and working group outputs relevant to organizations responding to the device code phishing threat.

The AI Controls Matrix (AICM), as a superset of the Cloud Controls Matrix (CCM), is the primary CSA framework applicable to identity and access management risks in AI-enabled environments. CCM controls IAM-02 (Strong Authentication) and IAM-04 (Separation of Duties) are directly implicated: device code phishing demonstrates that strong authentication controls must extend beyond credential strength and MFA enrollment to encompass token lifecycle governance and authentication flow restriction. The AICM provides additional guidance on identity security in AI-assisted environments where agentic workflows may legitimately use device code flow for non-human authentication — creating a governance tension that organizations should resolve explicitly rather than defaulting to blanket policy.

The CSA STAR (Security Trust Assurance and Risk) program's Consensus Assessments Initiative Questionnaire (CAIQ) includes identity and access control questions that cloud service providers should address with respect to device authorization grant support. Organizations evaluating cloud providers or SaaS products should ask whether device code flow is enabled, whether it can be restricted by tenant policy, and what logging and alerting capabilities are provided for device code flow authentication events.

CSA's Zero Trust guidance emphasizes continuous verification and the principle of least privilege access — both of which apply directly to token lifecycle management. The device code phishing threat underscores that Zero Trust architectures must account for token-based access as a distinct trust artifact from authenticated sessions. Refresh token lifetimes and PRT generation policies represent Zero

Trust control points that many organizations have not explicitly evaluated against their trust requirements. CSA's Zero Trust working group outputs provide a useful framework for organizations reassessing their identity plane security posture in light of this threat.

The MAESTRO (Agentic AI Threat Modeling) framework's Layer 2 (Data Operations) and Layer 5 (Infrastructure) considerations are relevant where agentic AI systems use OAuth device code flow for non-human authentication — a legitimate use case for headless AI agents that access Microsoft 365 on behalf of users. Organizations deploying AI agents that authenticate via device code flow should evaluate whether those flows can be migrated to client credential grants with managed identities, reducing the attack surface while preserving automation capability.

References

[1] Push Security / BleepingComputer. "Device code phishing attacks surge 37x as new kits spread online." BleepingComputer, April 4, 2026.

[2] IETF. "RFC 8628 – OAuth 2.0 Device Authorization Grant." Internet Engineering Task Force, August 2019.

[3] The Hacker News. "Device Code Phishing Hits 340+ Microsoft 365 Orgs Across Five Countries." The Hacker News, March 25, 2026.

[4] Microsoft Security Blog. "Storm-2372 conducts device code phishing campaign." Microsoft, February 13, 2025.

[5] Proofpoint. "Access granted: phishing with device code authorization for account takeover." Proofpoint Threat Insight, December 18, 2025.

[6] Sophos. "OAuth's Device Code Flow Abused in Phishing Attacks." Sophos, June 3, 2021.

[7] Microsoft Learn. "Block authentication flows with Conditional Access policy." Microsoft, 2025.

[8] Microsoft Learn. "Authentication flows as a condition in Conditional Access policy." Microsoft, 2025.

[9] Sekoia. "New widespread EvilTokens kit: device code phishing as-a-service - Part 1." Sekoia Threat Intelligence, March 30, 2026.

[10] BleepingComputer. "New EvilTokens service fuels Microsoft device code phishing attacks." BleepingComputer, April 1, 2026.

[11] Infosecurity Magazine. "OAuth Device Code Phishing Campaigns Surge, Targets Microsoft 365." Infosecurity Magazine, December 18, 2025.

[12] Cybersecurity Dive. "State-linked and criminal hackers use device code phishing against M365 users." Cybersecurity Dive, December 19, 2025.

[13] The Hacker News. "Russia-Linked Hackers Use Microsoft 365 Device Code Phishing for Account Takeovers." The Hacker News, December 19, 2025.

[14] Help Net Security. "EvilTokens ramps up device code phishing targeting Microsoft 365 users." Help Net Security, March 31, 2026.

