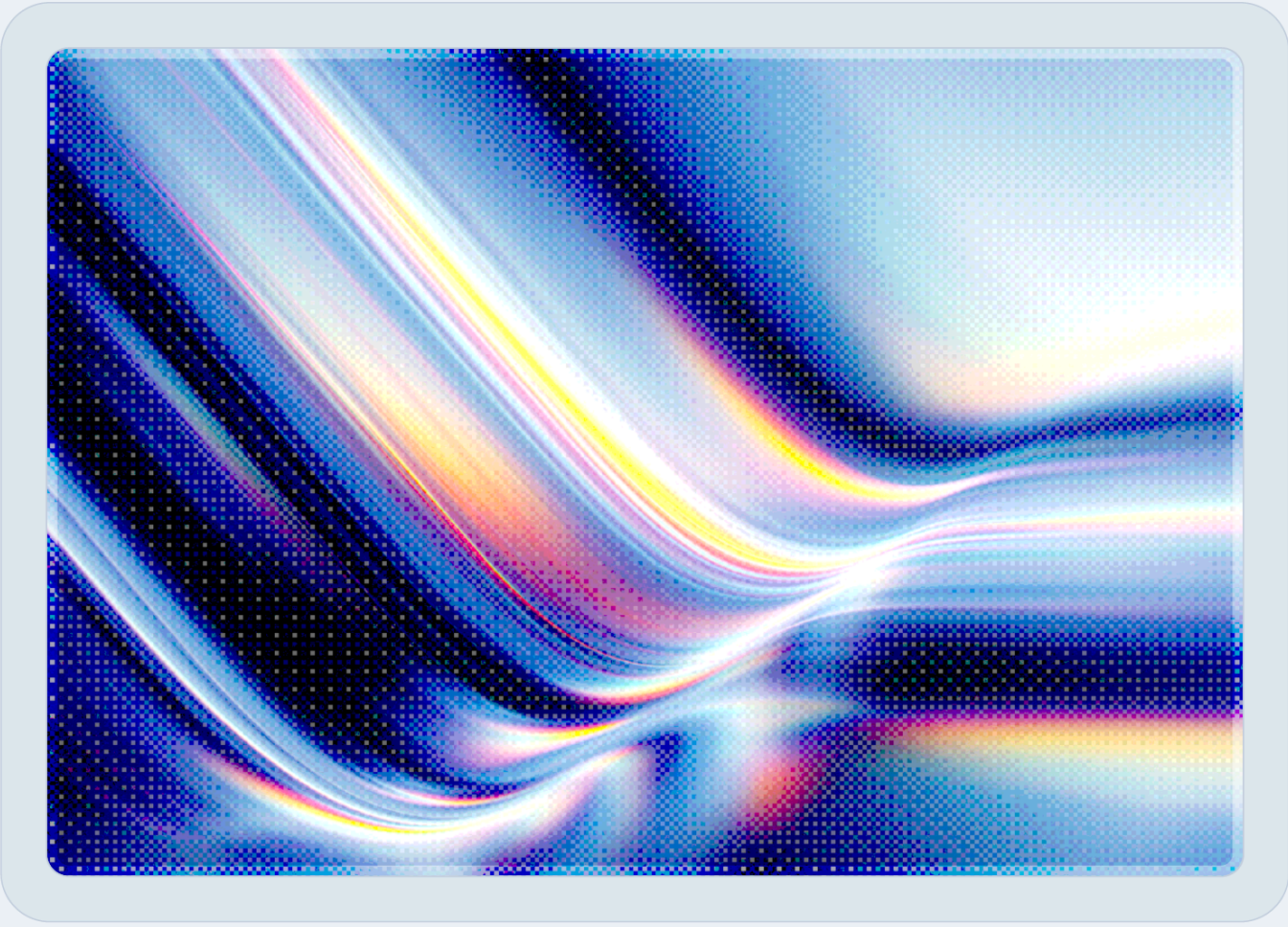
CSAI Foundation | Cloud Security Alliance

NIST IR 8587: Token Security Rules Extend to AI Agent Identity

2026-09-19

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

Key Takeaways

NIST and CISA finalized NIST Interagency Report (IR) 8587 on September 15, 2026, giving federal agencies and cloud service providers their first consolidated implementation guidance for protecting identity and access tokens from forgery, theft, and misuse [1][2]. According to NIST, the final report reflects roughly 250 public comments received on a December 2025 draft and direct collaboration with major cloud providers, including Google, Microsoft, Amazon Web Services, IBM, Okta, and Oracle [3]. The report responds in part to a pattern of real-world incidents in which stolen signing keys let adversaries forge authentication tokens and move laterally through federal systems undetected, most notably the 2023 Storm-0558 campaign that used a compromised Microsoft signing key to steal roughly 60,000 emails from a single federal agency [4][5]. While IR 8587 is written primarily for human single sign-on, federation, and API access scenarios, it explicitly acknowledges that AI agents and other automated workloads rely on the same token mechanisms and states that additional standards work is still needed to fully govern machine and agent identities [1][3]. Security and compliance teams should treat IR 8587 as the near-term federal baseline for token hygiene while watching NIST's parallel, more recent effort on AI agent identity and authorization for the rules that will eventually close that gap.

Background

IR 8587, formally titled "Protecting Tokens and Assertions from Forgery, Theft, and Misuse: Implementation Recommendations for Agencies and Cloud Service Providers," was authored by NIST's Ryan Galluzzo and Andrew Regenscheid, Accenture Federal Services' Stephanie Nelson, and CISA's Christine Lazcano [3]. It builds on the control baseline established in NIST SP 800-53 Release 5.1.1 and translates those controls into concrete architectural guidance for identity providers, authorization servers, and the cloud services that consume their tokens [3]. The report was issued jointly by NIST and CISA, a joint posture CISA Acting Executive Assistant Director Chris Butera made explicit: "Identity is the new perimeter, and the tokens and assertions behind it are attractive targets for sophisticated adversaries" [2].

The report's origin traces to a series of high-profile compromises in which nation-state actors forged tokens using stolen cryptographic material rather than exploiting a specific software vulnerability. The best-documented case is the 2023 Storm-0558 campaign, in which a China-based threat actor obtained a Microsoft consumer signing key that had leaked into a corporate debugging environment and

used it to forge authentication tokens for Outlook Web Access, ultimately accessing government email accounts at roughly 25 organizations, including the U.S. State Department and Department of Commerce [5][6]. Because Azure Active Directory did not distinguish consumer-tier signing keys from enterprise-tier keys at the time, a single stolen key functioned as a nearly universal pass across tenant boundaries, and reporting on the incident estimated that around 60,000 emails were exfiltrated from the State Department alone [4][6]. NIST and CISA cite this style of attack directly in IR 8587 as an illustrative example of the risks that motivate stronger key management, token verification, and lifecycle controls [1][4].

IR 8587 divides responsibility primarily between cloud providers, who are expected to secure signing infrastructure, protect private keys, and issue tokens according to secure-by-design principles, and agency customers, who are expected to configure identity services, apply available security controls, and monitor their own token usage, with additional shared duties for both parties in detecting compromises, exchanging security signals, and supporting incident investigations [4]. Ryan Galluzzo summarized the intended reach of the guidance beyond government: "Anyone who is using tokens as part of their access management infrastructure can look to this for insights, whether they are in government or commercial industry" [1]. That framing matters because the underlying token mechanisms IR 8587 addresses, OAuth 2.0 bearer tokens, SAML assertions, and OIDC identity tokens, are governed by the same protocols whether they authenticate a human employee, a backend service, or an autonomous AI agent [10].

Security Analysis

The substantive guidance in IR 8587 concentrates on four areas: key management, token verification, revocation, and lifecycle controls, all applied to systems using digitally signed or asymmetrically encrypted tokens across single sign-on, federation, API access, and workload access scenarios [3][4]. On key management, the report favors outcome-based rather than highly prescriptive requirements, reflecting the reality that cloud providers use varied key storage and rotation architectures; agencies and providers are still expected to demonstrate that signing keys are protected against the kind of exfiltration that enabled Storm-0558. On verification and lifecycle controls, the report pushes toward continuous monitoring and threat-adaptive defenses rather than static, one-time validation at token issuance, and it calls for configurable, interoperable revocation mechanisms so that a compromised token or key can be invalidated quickly across federated systems rather than remaining valid until natural expiration [4].

In CSA's assessment, the report's treatment of AI and machine identity is where its scope is most consequential for CSA's constituency, and also where it remains least developed. NIST states that IR 8587 includes "new high-level considerations for handling AI" and that the guidance covers workload access by applications and automated services using the same signed-assertion mechanisms as human

sessions [1][4]. However, these considerations fall short of comprehensive tooling for AI agent identity, and NIST's own National Cybersecurity Center of Excellence (NCCoE) is running a separate, more recent track to close that gap. In February 2026, NCCoE published a concept paper, "Accelerating the Adoption of Software and Artificial Intelligence Agent Identity and Authorization," proposing a demonstration project built on OAuth 2.0, SPIFFE/SPIRE, and the Model Context Protocol to give autonomous agents distinct, attestable identities rather than treating them as generic service accounts [7]. That effort sits alongside the AI Agent Standards Initiative that NIST's Center for AI Standards and Innovation (CAISI) launched in February 2026 to develop interoperability and security standards for agentic AI systems [12].

This split matters operationally. IR 8587 gives agencies and cloud providers a compliance-grade baseline they can implement now, using existing SP 800-53-aligned control language, but it does not yet answer harder questions specific to AI agents: how to attribute an agent's action back to a human principal through a delegation chain, how to bound an agent's token scope as it invokes multiple downstream tools, or how to revoke access fast enough when an agent's credentials are exposed through prompt injection rather than a stolen signing key. Independent research CSA has published this year documents why those questions are urgent. CSA's non-human identity governance research finds that AI agent identities are frequently created dynamically, accumulate privilege beyond their operational need, and can obscure delegation chains in multi-agent workflows in ways legacy identity infrastructure was not built to handle [8]. A companion survey found that 68 percent of security teams cannot reliably distinguish AI agent activity from human activity in their logs, and that 74 percent of organizations report agents routinely receiving more access than their tasks require [9]. Read together, IR 8587's token-forgery baseline and the NCCoE's agent-identity concept paper describe a federal identity security agenda moving in two speeds: a codified minimum for token hygiene today, and an unsettled standards conversation about machine and agent identity that will take at least another development cycle to mature.

Recommendations

Immediate Actions

Agencies and cloud service providers subject to IR 8587 should begin by inventorying where digitally signed tokens and assertions are used across their SSO, federation, and API access paths, and confirming that signing-key protection meets the outcome-based standard IR 8587 describes rather than assuming existing key management is sufficient. Security teams should also verify that their identity providers support rapid, cross-tenant token revocation, since the Storm-0558 case demonstrated how a single compromised key can remain exploitable across many organizations if revocation is slow or manual

[4][6]. Organizations already running AI agents against federated identity providers should extend this same inventory exercise to agent credentials, since IR 8587's workload-access language applies to them even though dedicated agent-identity controls are still forthcoming.

Short-Term Mitigations

Over the next two to three quarters, agencies and CSPs should implement the continuous monitoring and threat-adaptive controls IR 8587 recommends, including anomaly detection on token issuance and use patterns that would have flagged the unusual cross-tenant access seen in the Storm-0558 case. For organizations with production AI agents, this is also the window to move away from shared service accounts and long-lived credentials toward the short-lived, scoped tokens and distinct per-agent identities that CSA's non-human identity research recommends, since these controls will likely align closely with whatever NCCoE's demonstration project formalizes [8]. Teams should track the NCCoE concept paper's public comment process and pilot results, since its OAuth 2.0 and SPIFFE/SPIRE-based approach is a reasonable planning assumption for where federal agent-identity requirements are heading [7].

Strategic Considerations

Longer term, security and compliance leadership should plan for IR 8587 and the emerging AI agent identity standards to converge into a single expectation: that every credentialed actor accessing federal or federally connected cloud systems, human or machine, has an attestable identity, a bounded and auditable scope, and a revocation path that works in minutes rather than hours. Organizations should also monitor NIST's post-quantum cryptography transition guidance referenced in IR 8587, since token signing algorithms will need to migrate on a timeline that will eventually intersect with agent-identity credentialing work [3][4]. Given that IR 8587 was shaped by direct engagement with the largest commercial cloud providers, enterprises outside the federal space should not assume its provisions are government-only; the same token-forgery risks apply anywhere OAuth, SAML, or OIDC tokens govern access, and adopting IR 8587's practices ahead of a compliance mandate reduces exposure to the next Storm-0558-style incident regardless of sector.

CSA Resource Alignment

IR 8587's core subject, protecting the cryptographic tokens that increasingly govern both human and machine access, connects directly to CSA's growing body of non-human identity research. CSA's whitepaper "The Non-Human Identity Governance Vacuum: AI Agents and the Fastest-Growing

Unmanaged Attack Surface" analyzes the same structural problem IR 8587 leaves partially unresolved: that non-human and agent identities are created dynamically, accumulate privilege beyond their operational need, and can obscure delegation chains in multi-agent workflows in ways legacy identity infrastructure was not built to handle. It recommends SPIFFE/SPIRE-based workload identity as a foundational mitigation, the same technical building block NCCoE's concept paper proposes for agent identity [8]. "AI Agent Identity Sprawl: The Enterprise Authorization Crisis" complements this with survey data quantifying the audit gap IR 8587's workload-access language implicitly targets, finding that 68 percent of security teams cannot distinguish agent activity from human activity in their logs and that 74 percent of organizations report agents receiving more access than their tasks require [9]. Finally, CSA's "Navigating Identity and Access Management (IAM)" guide provides the protocol-level reference, covering the OAuth 2.0, SAML, OIDC, and NIST SP 800-63-aligned assurance concepts that IR 8587's controls are built on, making it a useful implementation companion for teams translating the federal guidance into architecture decisions [10]. Where an organization's identity governance needs a control-framework anchor beyond these topic-specific artifacts, CSA's AI Controls Matrix (AICM) v1.1 IAM domain remains the appropriate baseline for mapping both human and non-human identity controls to a broader compliance program [11].

References

[1] NIST. "NIST Finalizes Guidelines for Protecting Online Identity and Access Tokens." NIST, September 15, 2026.

[2] CISA. "CISA and NIST Release Guidelines to Protect Federal Cloud Identity Systems from Token Theft, Forgery, and Misuse." CISA, September 15, 2026.

[3] NIST. "NIST IR 8587: Protecting Tokens and Assertions from Forgery, Theft, and Misuse." NIST Computer Security Resource Center, September 15, 2026.

[4] Help Net Security. "NIST, CISA Release Guidance on Protecting Cloud Identity Tokens." Help Net Security, September 16, 2026.

[5] SecurityWeek. "Chinese Cyberspies Used Forged Authentication Tokens to Hack Government Emails." SecurityWeek, 2023.

[6] Microsoft. "Analysis of Storm-0558 Techniques for Unauthorized Email Access." Microsoft Security Blog, July 14, 2023.

[7] NIST NCCoE. "Accelerating the Adoption of Software and Artificial Intelligence Agent Identity and Authorization." NIST Computer Security Resource Center, February 5, 2026.

[8] Cloud Security Alliance. "The Non-Human Identity Governance Vacuum: AI Agents and the Fastest-Growing Unmanaged Attack Surface." Cloud Security Alliance, May 20, 2026.

[9] Cloud Security Alliance. "AI Agent Identity Sprawl: The Enterprise Authorization Crisis." Cloud Security Alliance, 2026.

[10] Cloud Security Alliance. "Navigating Identity and Access Management (IAM)." Cloud Security Alliance, 2026.

[11] Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." Cloud Security Alliance, 2026.

[12] NIST. "Announcing the AI Agent Standards Initiative for Interoperable and Secure Innovation." NIST, February 17, 2026.

