

Security Threat Modeling for Federated Identity Systems: A STRIDE-Based Assessment of SSO and MFA Vulnerabilities

Arjun Mehta

Department of Data Science and Machine Learning,
Institute of Computing and Artificial Intelligence, New Delhi, India

VOLUME 03 ISSUE 09 (2026)

Published Date: 01 September 2026 // Page no.: - 01-07

ABSTRACT

Federated identity systems have become a central mechanism for enabling single sign-on (SSO), multi-application authentication, and multi-factor authentication (MFA) across organizational and service boundaries. Their security value, however, depends not only on the strength of individual authentication mechanisms but also on the integrity of trust relationships, identity assertions, authentication workflows, session management, and federation infrastructure. This research develops a STRIDE-oriented threat-modeling approach for assessing security vulnerabilities in federated SSO and MFA environments. The study adopts a structured analytical methodology consisting of system decomposition, trust-boundary identification, threat categorization, vulnerability assessment, and control mapping. Because the supplied literature is primarily concentrated on statistical modeling, machine learning, financial systems, and explainability rather than federated identity security, the analysis explicitly treats these sources as methodological foundations rather than direct empirical evidence of specific identity attacks. In particular, concepts concerning sequential modeling, uncertainty, smoothing, and black-box explainability are used to inform the analytical structure of threat identification and prioritization. The study finds that federated identity architectures create concentrated security dependencies: compromise of an identity provider, federation endpoint, token-validation mechanism, or MFA workflow can propagate across multiple relying applications. STRIDE analysis further indicates that spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege can occur across different layers of the authentication lifecycle. The research contributes a systematic conceptual model for integrating STRIDE with federated SSO and MFA risk analysis and identifies opportunities for adaptive, explainable, and uncertainty-aware security assessment.

Keywords: Federated Identity, Single Sign-On, Multi-Factor Authentication, STRIDE, Threat Modeling, Identity Security, Authentication, Security Assessment, Access Control.

1. INTRODUCTION

Federated identity systems allow users to authenticate through a trusted identity provider while accessing multiple relying applications or services. This architecture reduces authentication duplication and can improve administrative efficiency because users do not need independent credentials for every application. SSO extends this principle by creating a unified authentication experience, while MFA introduces additional authentication factors intended to reduce the security consequences of compromised passwords.

The architectural advantage of federation simultaneously creates a security concentration point. A conventional application may expose one authentication boundary, whereas a federated environment contains relationships among identity providers, relying parties, authentication services, token or assertion mechanisms, session-management components, and MFA infrastructure. Consequently, an attacker who compromises a highly

trusted component may obtain access beyond the original point of compromise. The security problem is therefore not adequately represented by evaluating passwords or MFA factors independently.

Threat modeling provides a systematic means of examining these dependencies before or during deployment. Rather than treating vulnerabilities as isolated technical defects, threat modeling evaluates how assets, trust boundaries, communication paths, authentication states, and privileges interact. STRIDE provides a useful classification structure by organizing threats around spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege.

The methodological relevance of structured analytical approaches is also evident in the supplied literature. Sequential and probabilistic modeling research demonstrates how dependencies between successive states can be represented systematically (Bengio, 1995; Bengio, 1996). Similarly, uncertainty-oriented analysis provides a conceptual

basis for avoiding overly deterministic security judgments (Fang et al., 2024). Explainability research is particularly relevant because security teams must understand why a particular identity component or authentication flow has been classified as high risk rather than simply receiving an opaque risk score (Guidotti et al., 2018).

The central objective of this research is therefore to develop a structured STRIDE-based assessment model for federated SSO and MFA vulnerabilities. The study focuses on identifying threat surfaces, analyzing attack consequences, and establishing a systematic approach for prioritizing security controls. The work is conceptual and analytical rather than an empirical penetration test.

An important limitation concerns the supplied reference base. Most of the provided references address mortgage markets, finance, statistical modeling, machine learning, and explainable artificial intelligence rather than federated identity. Accordingly, these sources are not represented as empirical evidence for particular SSO or MFA attacks. Instead, they provide methodological concepts that support modeling, uncertainty analysis, and explainability. The directly relevant conceptual work supplied by the user is Ganapathy (2024), which provides the principal domain-specific basis for the STRIDE-oriented assessment developed here.

2. LITERATURE REVIEW

2.1 Methodological Foundations for Security Modeling

The literature supplied for this study contains several methodological streams that can inform threat-modeling research. Bengio and Frasconi's work on input-output hidden Markov model architectures demonstrates the importance of representing relationships between inputs and sequential states (Bengio, 1995; Bengio, 1996). Although developed for sequence-processing applications, this perspective is useful for identity systems because authentication is inherently sequential. A user request, credential validation, MFA challenge, authentication decision, token issuance, application request, and session establishment can be considered a sequence of security-relevant states.

This sequential perspective is important because a vulnerability may not become exploitable at a single point. Instead, exploitability can emerge from interactions between several states. For example, an otherwise secure authentication mechanism may become vulnerable when an improperly validated authentication result is subsequently converted into a privileged session. Thus, modeling individual components without considering state transitions can underestimate systemic risk.

Statistical modeling literature also contributes to the

theoretical foundation. Gu (2013), Gu (2014), Gu and Ma (2005), and Gu and Wahba (1991) demonstrate approaches for modeling complex relationships and selecting appropriate smoothing parameters. These studies are not identity-security studies, but their methodological emphasis on modeling complex relationships is relevant to threat assessment where multiple factors influence risk. A federated identity system similarly contains interacting variables rather than independent security controls.

2.2 Uncertainty and Risk Interpretation

Security assessment frequently involves incomplete information. Analysts may not know whether a particular federation endpoint has been previously targeted, whether a user has been socially engineered, or whether an authentication anomaly represents malicious behavior. Fang et al. (2024) provides a useful methodological perspective by emphasizing uncertainty quantification in knowledge distillation. Applied conceptually to cybersecurity, uncertainty-aware analysis discourages the interpretation of every risk estimate as absolute.

This distinction is significant for federated systems. A threat model can identify that token tampering is theoretically possible, but actual risk depends on validation controls, cryptographic integrity mechanisms, application behavior, and attacker capability. Threat modeling should therefore distinguish between potential threat existence and confirmed exploitability.

2.3 Explainability and Security Decisions

Guidotti et al. (2018) examine methods for explaining black-box models. Explainability is relevant to automated security assessment because organizations increasingly rely on analytics to prioritize alerts and risks. A threat-modeling framework that generates a high-risk classification without showing the contributing factors may be difficult for security engineers to validate.

The implication for federated identity is that risk assessments should identify the structural reason for a threat rating. For example, an elevation-of-privilege risk should be traceable to an authentication-to-authorization transition, a trust-boundary weakness, a token-validation assumption, or an MFA recovery pathway. Explainable threat modeling therefore improves both technical validation and governance.

2.4 Domain-Specific Threat Modeling

Ganapathy (2024) provides the primary domain-specific reference for this study, focusing explicitly on threat modeling for federated SSO and MFA using STRIDE-based analysis. Its relevance lies in treating authentication federation as an interconnected attack surface rather than merely evaluating individual credentials.

The current study extends this conceptual positioning by organizing the threat assessment around architectural layers: user authentication, identity-provider processing, federation exchange, token or assertion handling, relying-party validation, session management, and MFA operations. The principal research gap is therefore not simply the identification of individual threats, but the development of an integrated analytical structure connecting threat category, architectural location, attack consequence, and security control.

3. METHODOLOGY

3.1 Research Design

This study uses a qualitative, architecture-oriented threat-modeling methodology. The process consists of five analytical stages: system decomposition, asset identification, trust-boundary analysis, STRIDE threat mapping, and risk interpretation.

The methodology does not claim to measure real-world attack frequency. Instead, it evaluates plausible threat conditions within a representative federated SSO and MFA architecture. This distinction is important because the supplied reference collection does not contain a controlled dataset of federation incidents.

The methodology is conceptually consistent with Ganapathy (2024), while incorporating sequential-state thinking, uncertainty awareness, and explainability principles derived from the broader supplied literature.

3.2 System Decomposition

A representative federated identity environment can be decomposed into the following functional components:

1. User or client device.
2. Identity Provider (IdP).
3. Authentication service.
4. MFA service.
5. Federation protocol or assertion exchange.
6. Relying Party (RP) or service provider.
7. Session-management layer.
8. Authorization and access-control layer.

The decomposition establishes where security decisions are made. The IdP determines whether a user has successfully authenticated; the federation layer communicates that authentication state; and the relying application converts the received identity information

into an application session and authorization decision.

This sequence resembles a state-transition problem. Bengio and Frasconi's sequential modeling perspective is therefore conceptually useful because the security outcome depends on how one state is transformed into another (Bengio, 1995; Bengio, 1996).

3.3 Asset Identification

The principal assets include user identities, authentication credentials, MFA factors, authentication assertions or tokens, session identifiers, authorization information, application data, and trust relationships between federation participants.

Asset criticality should be assessed according to both direct and cascading consequences. Compromise of a single application account may have limited impact, whereas compromise of a federation-level signing or authentication mechanism can potentially affect numerous relying applications.

This produces an important principle:

Federation risk is partly determined by the concentration of trust.

The greater the number of applications dependent on a common identity authority, the greater the potential systemic impact of a failure at that authority.

3.4 Trust-Boundary Analysis

Trust boundaries identify transitions where information or authentication decisions move between security domains. Important boundaries include:

- user device → identity provider;
- identity provider → MFA service;
- identity provider → relying party;
- federation service → application;
- authenticated session → authorization layer.

Each boundary represents an opportunity for spoofing, tampering, information disclosure, denial of service, or privilege escalation.

For example, successful authentication at the IdP does not automatically guarantee correct authorization at the relying party. The relying party must correctly interpret and validate the authentication result. Consequently, the federation boundary represents a critical security dependency.

3.5 STRIDE-Based Threat Mapping

STRIDE is applied to each architectural component and trust boundary.

Spoofting

Spoofting concerns the fraudulent representation of an identity. In federated systems, this may involve compromise of user credentials, fraudulent authentication sessions, manipulation of identity-provider interactions, or exploitation of weak MFA recovery mechanisms.

The primary security consequence is unauthorized authentication. MFA reduces some credential-compromise risks, but the overall security benefit depends on the integrity of the entire authentication process.

Tampering

Tampering involves unauthorized modification of security-relevant information. In a federation architecture, potential targets include authentication assertions, token parameters, session information, authorization attributes, or configuration data.

The analytical concern is not simply whether information can be modified but whether the receiving system can detect unauthorized modification. If a relying party accepts altered identity information, an attacker may convert a low-privilege authentication state into a more powerful one.

Repudiation

Repudiation concerns the inability to establish reliable evidence of actions. Federated authentication creates distributed responsibility: the IdP authenticates the user, while the relying application authorizes activity.

Without appropriate logging and event correlation, organizations may struggle to establish whether an action originated from legitimate authentication, session theft, administrative intervention, or another mechanism. Explainability principles are relevant here because security decisions should remain auditable and interpretable (Guidotti et al., 2018).

Information Disclosure

Information disclosure occurs when authentication, identity, or authorization information becomes accessible to unauthorized parties. Sensitive information may include user identifiers, authentication metadata, session information, or application-specific attributes.

The risk is amplified in federated architectures because identity information can cross organizational boundaries.

Data minimization and controlled attribute sharing therefore become architectural rather than merely application-level concerns.

Denial of Service

Denial-of-service threats target authentication availability. An attacker may attempt to overload an identity provider, interfere with MFA delivery, trigger excessive authentication challenges, or exploit dependencies between the IdP and relying applications.

Federation can amplify availability failures because many applications may depend on a shared authentication service. Thus, availability of identity infrastructure becomes a cross-application security property.

Elevation of Privilege

Elevation of privilege is particularly significant in SSO systems. A compromised or manipulated identity assertion can potentially produce access that exceeds the user's legitimate authorization.

This threat illustrates the distinction between authentication and authorization. Authentication establishes identity, whereas authorization determines permitted actions. Treating a successful federated authentication response as equivalent to unrestricted authorization creates architectural risk.

3.6 Risk Prioritization

Threat prioritization should incorporate at least four dimensions:

$$\text{Risk} \approx \text{Likelihood} \times \text{Impact} \times \text{Exposure} \times \text{Trust Dependency}$$

$$\text{Risk} \approx \text{Likelihood} \times \text{Impact} \times \text{Exposure} \times \text{Trust} \times \text{Dependency}$$

This conceptual formulation is not intended as a statistical estimator. It provides a structured mechanism for comparing threats. Uncertainty should also be explicitly recorded because security analysts may possess incomplete evidence. The uncertainty-oriented perspective of Fang et al. (2024) supports treating confidence in an assessment separately from the severity of the threat itself.

A threat with high potential impact but low confidence should not automatically be ignored; instead, it should be identified as requiring validation.

4. RESULTS

The STRIDE assessment indicates that federated SSO and MFA systems exhibit a distributed but highly interconnected threat surface. The most important finding is that security failures cannot be evaluated exclusively at the user-authentication layer. Identity-provider compromise, federation-message

manipulation, session weaknesses, and authorization errors can create downstream consequences across multiple applications.

Spoofing represents the most direct threat to authentication integrity. MFA provides an additional security barrier, but the analysis indicates that MFA should be modeled as part of a complete authentication state machine rather than as an isolated control. Credential compromise, authentication-session abuse, and weak recovery procedures can undermine the intended security benefit of MFA. Consequently, the presence of MFA alone should not be interpreted as evidence that spoofing risk is eliminated.

Tampering and elevation of privilege emerge as closely connected threats. An attacker who can manipulate identity-related information may ultimately seek a more consequential objective: unauthorized privilege. This demonstrates why token or assertion integrity and relying-party validation are critical. A federation system may successfully authenticate a user while still making an unsafe authorization decision if identity attributes are incorrectly interpreted.

The assessment also identifies denial of service as a systemic risk. Unlike an application-specific availability problem, an identity-provider outage can potentially affect many relying applications simultaneously. Federation therefore creates a form of availability concentration risk. The architectural efficiency gained from centralized identity management is accompanied by dependency on shared authentication infrastructure.

Repudiation and information disclosure are less immediately visible than spoofing but remain important. Distributed authentication generates evidence across multiple components. Without consistent logging and traceability, incident investigation can become difficult because no single system necessarily contains the complete authentication-to-authorization history.

The methodological analysis further indicates that threat assessments benefit from sequential and uncertainty-aware representations. Authentication is not a single event; it is a chain of state transitions. Bengio and Frasconi's work on sequence modeling provides a useful conceptual foundation for examining such dependencies (Bengio, 1995; Bengio, 1996). Likewise, uncertainty-aware analytical thinking is important when assessing threats for which exploitability cannot be directly observed (Fang et al., 2024).

Finally, explainability emerges as an important governance requirement. A security team should be able to determine why a particular identity pathway was assigned a high threat level. This corresponds with the

broader importance of interpretable analytical models discussed by Guidotti et al. (2018). The findings therefore support a threat-modeling approach that records not only threat classifications but also architectural causes, affected assets, attack consequences, confidence levels, and recommended control areas.

5. DISCUSSION

The findings demonstrate that federated identity security is fundamentally a trust-management problem, not merely a credential-management problem. SSO improves usability by centralizing authentication, while MFA strengthens authentication by requiring additional evidence. However, these benefits introduce dependencies that must be modeled at architectural scale.

The most significant theoretical implication is that threat modeling should treat authentication as a sequence of interdependent states. A secure initial authentication decision may still result in an insecure application session if subsequent federation processing, assertion validation, or authorization mapping is defective. This supports the relevance of sequential modeling concepts represented in the supplied literature (Bengio, 1995; Bengio, 1996).

A second implication concerns uncertainty. Traditional threat matrices often assign categorical values such as low, medium, or high without explicitly describing confidence in the assessment. An uncertainty-aware model is more informative because two threats with identical nominal severity may have substantially different evidentiary foundations. The conceptual contribution of uncertainty quantification in Fang et al. (2024) therefore supports a distinction between risk severity and assessment confidence.

The results also have practical implications for security governance. Organizations should not evaluate SSO, MFA, token validation, and authorization as independent controls. Instead, they should evaluate the complete authentication lifecycle. A vulnerability at the federation boundary may be more consequential than an equivalent weakness in an isolated application because the federation component serves multiple relying parties.

The literature concerning smoothing and model selection provides another methodological insight. Gu (2013), Gu (2014), Gu and Ma (2005), and Gu and Wahba (1991) emphasize structured approaches for handling complex relationships and model behavior. While these studies do not directly address identity security, their methodological orientation reinforces the need to avoid simplistic one-variable threat assessments. Federated security is similarly characterized by interacting factors.

Explainability is equally important. Security practitioners must understand the rationale behind risk prioritization, particularly when automated analytics are introduced. An

opaque assessment can create false confidence or cause analysts to overlook critical assumptions. The explainability literature therefore supports maintaining transparent relationships between system architecture, threat classification, evidence, and mitigation priorities (Guidotti et al., 2018).

Nevertheless, the study has significant limitations. First, the supplied references are not predominantly cybersecurity sources. Several concern mortgage markets, financial modeling, or statistical methodology, and therefore cannot be treated as empirical validation of the identified SSO and MFA threats. Second, the methodology is conceptual and does not include penetration testing, incident datasets, controlled experiments, or quantitative probability estimates. Third, actual threat severity depends on implementation details, federation protocols, configuration, organizational governance, and attacker capabilities.

Despite these limitations, the framework offers a useful analytical foundation. Ganapathy (2024) directly establishes the relevance of STRIDE-based threat modeling to federated SSO and MFA, while the remaining literature provides methodological concepts for sequential analysis, uncertainty, statistical modeling, and explainability. The resulting synthesis suggests that future research should combine architectural threat modeling with empirical attack datasets and measurable risk indicators.

6. CONCLUSION

Federated SSO and MFA architectures provide substantial operational and security advantages, but their centralized trust relationships create complex dependencies that can amplify individual security failures. This study applied a STRIDE-based analytical framework to examine spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege across representative federation components and trust boundaries.

The principal contribution is the integration of architectural decomposition with sequential, uncertainty-aware, and explainable threat assessment. The analysis demonstrates that authentication security cannot be reduced to password strength or MFA deployment. Instead, security depends on the integrity of the complete chain connecting authentication, federation, assertion processing, session establishment, and authorization.

The study further establishes that identity-provider compromise and federation-boundary weaknesses deserve particular attention because their consequences may extend across multiple relying applications. Availability and privilege risks are similarly amplified when many services depend on common identity

infrastructure.

Future research should empirically validate the proposed framework using real-world federation configurations, attack datasets, controlled security experiments, and quantitative risk metrics. Additional work should also investigate how machine learning and explainable analytical methods can continuously update threat assessments while explicitly representing uncertainty. Such research could move federated identity threat modeling from predominantly static architectural analysis toward adaptive and evidence-driven security assessment.

7. REFERENCES

1. Agarwal S, Chomsisengphet S, Kiefer H, Kiefer LC, Medina PC (2020). Inequality during the COVID-19 pandemic: The case of savings from mortgage refinancing. Available at SSRN 3750133.
2. Aldridge I, Avellaneda M (2019). Neural networks in finance: Design and performance. *The Journal of Financial Data Science*, 1(4): 39–62.
3. Bengio Y, Frasconi P (1995). An input output HMM architecture. In: *Advances in Neural Information Processing Systems*, volume 7, 427–434.
4. Bengio Y, Frasconi P (1996). Input-output HMMs for sequence processing. *IEEE Transactions on Neural Networks*, 7(5): 1231–1249.
5. Berger DW, Milbradt K, Tourre F, Vavra J (2018). Mortgage prepayment and path-dependent effects of monetary policy. Technical report, National Bureau of Economic Research.
6. Fang L, Chen Y, Zhong W, Ma P (2024). Bayesian knowledge distillation: A bayesian perspective of distillation with uncertainty quantification. In: *Forty-first International Conference on Machine Learning*, 12935–12956.
7. Federal Housing Finance Agency (2024). Prepayment Monitoring Report: First Quarter 2024. Technical report, Federal Housing Finance Agency.
8. Freddie Mac (2024). Primary Mortgage Market Survey (PMMS).
9. Fuster A, Hizmo A, Lambie-Hanson L, Vickery J, Willen PS (2021). How resilient is mortgage credit supply? evidence from the COVID-19 pandemic. Technical Report, National Bureau of Economic Research.
10. Ganapathy, S. K. . (2024). Threat Modeling for Federated SSO and MFA Systems: STRIDE-Based Analysis of Attack Vectors. *International Journal of Data*

11. Gu C (2013). Smoothing Spline ANOVA Models, volume 297. Springer.
12. Gu C (2014). Smoothing spline ANOVA models: R package gss. Journal of Statistical Software, 58: 1–25.
13. Gu C, Ma P (2005). Optimal smoothing in nonparametric mixed-effect models. The Annals of Statistics, 33(3): 1357–1379.
14. Gu C, Wahba G (1991). Minimizing GCV/GML scores with multiple smoothing parameters via the Newton method. SIAM Journal on Scientific and Statistical Computing, 12(2): 383–398.
15. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D (2018). A survey of methods for explaining black box models. ACM Computing Surveys, 51(5): 1–42.
16. Ganapathy, S. K. (2024). Threat Modeling for Federated SSO and MFA Systems: STRIDE-Based Analysis of Attack Vectors. International Journal of Data Science and Machine Learning, 4(02), 55-73. <https://www.academicpublishers.org/journals/index.php/ijdsml/article/view/stride-analysis-sso-mfa>