

RESEARCH TITLE

A proposed approach for implementing Zero Trust-based data access policies integrated with cloud access security measures to counter shadow phishing

Hadia Awad Mohamed Ahmed¹

¹ PhD student; BSc and MSc, Faculty of Mathematical Sciences, University of Khartoum. Sudan.

Email: hadia.awad.111@gmail.com

HNSJ, 2026, 7(11); <https://doi.org/10.53796/hnsj711/21>

Received at 15/10/2026

Accepted at 20/10/2026

Published at 01/11/2026

Abstract

This study proposes a security framework integrating Zero Trust data access policies with a Cloud Access Security Broker (CASB) to mitigate shadow phishing. It combines continuous identity and device verification, contextual risk assessment, traffic inspection, and data loss prevention. The proposed methodology covers application discovery, behavioral baselining, policy formulation, CASB integration, and controlled attack simulations. Candidate machine learning methods support anomaly detection, traffic classification, and content analysis. Planned evaluation measures include blocking rates, false positives, latency, and session revocation time. Expected benefits are tighter control of unauthorized transfers and reduced lateral movement. The contribution is architectural; effectiveness and performance remain subject to implementation and comparative testing.

Key Words: Zero Trust; Cloud Security; Shadow Phishing; Access Control.

نهج مقترح لتطبيق سياسات الوصول للبيانات وفق انعدام الثقة ودمجها بأمن الوصول السحابي لمواجهة التصيد الاحتيالي الخفي

هادية عوض محمد احمد¹

¹ طالبة دكتوراه؛ بكالوريوس وماجستير، كلية العلوم الرياضية، جامعة الخرطوم، السودان.

البريد الإلكتروني: hadia.awad.111@gmail.com

HNSJ, 2026, 7(11); <https://doi.org/10.53796/hnsj711/21>

تاريخ النشر: 2026/11/01م

تاريخ القبول: 2026/10/20م

تاريخ الاستقبال: 2026/10/15م

المستخلص

تقترح الدراسة إطاراً أمنياً يدمج سياسات الوصول إلى البيانات القائمة على انعدام الثقة مع وسيط أمن الوصول السحابي لمواجهة التصيد الاحتيالي الخفي. ويجمع الإطار التحقق المستمر من الهوية والأجهزة، وتقييم المخاطر، وفحص الحركة، ومنع تسرب البيانات. وتشمل المنهجية اكتشاف التطبيقات، وبناء خطوط أساس للسلوك، وصياغة السياسات، ودمج الوسيط، ومحاكاة الهجمات. وتدعم خوارزميات تعلم الآلة كشف الشذوذ وتصنيف الحركة وتحليل المحتوى. ويخطط للتقييم بمعدلات الحجب والإنذارات الكاذبة وزمن المعالجة وإلغاء الجلسات. ويتوقع تقييد نقل البيانات غير المصرح به والحد من الحركة الجانبية. والمساهمة معمارية؛ وتستلزم الفاعلية والأداء تنفيذاً واختباراً مقارناً.

الكلمات المفتاحية: انعدام الثقة؛ الأمن السحابي؛ التصيد الخفي؛ التحكم في الوصول.

1. Background

Decentralized workforce models and rapid cloud adoption have broken down traditional network borders, exposing organizations to "Shadow Phishing"—a threat where bad actors leverage trusted SaaS domain reputations, in-app messaging, and hidden dynamic links to bypass standard email inspection tools. Although Cloud Access Security Brokers (CASBs) monitor activity using out-of-band APIs and live proxy decryption, traditional CASBs rely on rigid rule sets that fail to spot credential theft occurring within authentic cloud sessions. Unifying Zero-Trust Architecture with CASB technology resolves this vulnerability by pairing the analytical governing logic (Policy Decision Point) directly with real-time execution (Policy Enforcement Point); this combination applies ongoing risk scoring, granular transactional controls, and context-sensitive data protection to instantly govern file movements across both approved cloud tools and unsanctioned applications.[1]

2. Introduction

The rapid acceleration of cloud adoption and hybrid work environments has transformed organizational boundaries. As enterprises shift from centralized on-premises datacenters to decentralized Software-as-a-Service (SaaS) environments, traditional perimeter security models—built on implicit network trust and boundary firewalls—have become obsolete. Cybercriminals have adapted to this structural shift by moving beyond traditional email attack vectors and deploying "Shadow Phishing."

Shadow Phishing represents an advanced social engineering paradigm where adversaries exploit unsanctioned cloud applications (Shadow IT) or abuse legitimate, cryptographically signed SaaS collaboration platforms—such as Google Workspace, Microsoft 365, Canva, and Notion—to deliver malicious payloads, credential-harvesting links, and QR codes (Quishing). Because notifications, document shares, and invite links originate directly from trusted cloud infrastructure, they cleanly bypass Secure Email Gateways (SEGs), Domain Key Identified Mail (DKIM) checks, and static IP reputation filters [2]

To counter the rise of shadow phishing—where attackers exploit unmonitored SaaS channels and legitimate cloud infrastructure to bypass traditional security gateways—enterprises must replace perimeter-based implicit trust with an integrated Zero Trust Architecture. By coupling Cloud Access Security Broker (CASB) inline traffic inspection with a dynamic Policy Decision Point aligned with NIST SP 800-207 standards, this model continuously evaluates user behavior, device health, and link threat signals. This integration replaces static binary access controls with real-time identity validation, automated session re-evaluation, and granular micro-segmentation, effectively neutralizing shadow phishing vectors and preventing unauthorized data exposure across enterprise cloud environments.[3]

3. Problem Statement

Modern enterprise security is challenged by "Shadow Phishing"—an advanced attack vector where adversaries bypass conventional Secure Email Gateways (SEGs) and endpoint filters by operating through unsanctioned, unmonitored cloud applications (Shadow IT) or abusing legitimate, trusted Software-as-a-Service (SaaS) collaboration platforms (e.g., Google Docs, Microsoft OneDrive, Canva, Notion).

Traditional boundary filters rely heavily on domain reputation, perimeter IP checks, and static email scanning. However, Shadow Phishing exploits three critical systemic vulnerabilities:

Implicit SaaS Trust: Notifications and share links originate from cryptographically signed, highly trusted cloud infrastructure (e.g., *.google.com, *.sharepoint.com), enabling emails to pass SPF/DKIM/DMARC checks cleanly.

Out-of-Band Lateral Movement: Threat actors compromise personal or third-party cloud accounts to deliver malicious payloads, QR codes, or credential-harvesting links directly via collaboration comments, internal messaging, or shared workspaces—entirely bypassing email gateways.

Data Exfiltration via Unsanctioned Channels: Victims interacting with these links upload sensitive enterprise data or credentials directly to unmonitored endpoints, where perimeter Data Loss Prevention (DLP) tools lack visibility.

Current static access control policies assume implicit session trust once a user authenticates. Without continuous, context-aware verification of data flows across both sanctioned and unsanctioned cloud channels, organizations remain exposed to zero-day credential harvesting and silent data exfiltration driven by Shadow Phishing.

4. Literature Review

4.1 Shaffi et al. (2025) examine how modern cloud environments can address sophisticated social engineering threats—such as AI-driven phishing and credential theft targeting cloud access—by replacing static defenses with automated, machine learning-driven solutions. The authors show that traditional firewalls and basic email filters fail against sophisticated attacks that exploit cloud access configurations. To mitigate these vectors, the study advocates for AI-stirred threat detection, behavior-based anomaly monitoring, and automated security responses. By continuously tracking behavioral shifts across user accounts, AI systems can spot illegitimate access attempts resulting from social engineering before sensitive assets are exfiltrated, significantly reducing operational reaction times and human error in cloud platforms.[4]

4.2 Suetor et al. (2025) evaluate cloud enforcement architectures, specifically highlighting Cloud Access Security Brokers (CASBs) as essential security gateways for countering unauthorized entry and social engineering exploits. The authors argue that credential theft and unauthorized access frequently exploit gaps between users and cloud endpoints. Placing a CASB between network traffic and cloud applications enforces centralized data loss prevention (DLP), continuous identity verification, and dynamic access control policies in real time. By continuously assessing risk postures and monitoring context-aware user actions, CASBs neutralize credential harvesting and unauthorized data access attempts resulting from deceptive phishing tactics.[5]

4.3 Syed et al. (2025) investigate how Zero Trust data access policies leverage Attribute-Based Access Control (ABAC) and continuous context verification to protect sensitive digital assets. Moving away from static Role-Based Access Control (RBAC), the study demonstrates that Zero Trust frameworks dynamically evaluate contextual signals before granting granular access to data resources. By enforcing strict least-privilege principles and requiring per-session re-authorization, this policy framework prevents credential abuse and contains the blast radius of compromised accounts, ensuring data access is strictly bounded by changing operational contexts.[6]

4.4 Al-Ahmad and Bakhrouf (2025) synthesize ten years of research on Zero Trust Architecture (ZTA) data policy enforcement across enterprise networks. Their analysis emphasizes that effective Zero Trust data protection relies on micro-segmentation and continuous data-in-transit encryption coupled with dynamic Policy Enforcement Points (PEPs). The findings highlight that treating all internal data transactions as untrusted by default reduces unauthorized internal lateral movement and mitigates the risk of insider data exfiltration by up to 80% compared to perimeter-based security [7]

The difference between previous studies and the proposed model is that the integrated ZTDA & CASB model solves the critical vulnerabilities of isolated architectures by unifying continuous identity and device verification with real-time cloud data inspection, creating a multi-layered defense against complex shadow phishing attacks. While standalone ZTDA primarily verifies perimeter credentials—leaving it blind to post-authentication session token theft and malicious outbound data transfers—and standalone CASB governs cloud application traffic without deep visibility into real-time endpoint posture, combining the two establishes an adaptive, context-aware control plane that dynamically evaluates device health, network vectors, and cloud payload activity simultaneously to detect anomaly patterns and revoking compromise-driven sessions instantly before sensitive data can be exfiltrated.

5. Research Scope

This proposed model will protect cloud data from indirect cyber threats like fake login pages, malicious QR codes and stolen access tokens by acting as an active security guard that constantly monitors user behavior. Instead of trusting someone just because they logged in once, it will evaluate every single data request using real-time factors like location, device security, and overall risk score. If anything looks suspicious, it will immediately step in to block the threat, restrict file downloads, or isolate unsafe links before sensitive information can be stolen.

6. Importance of the Study

As enterprise workforces decentralize across multi-cloud and hybrid environments, perimeter-based security controls are no longer sufficient. Shadow IT accounts for up to 60% of cloud app usage in enterprise environments, providing a massive, unmonitored attack surface for social engineering.

This study addresses critical research and operational gaps:

Closing the Visibility Blind Spot: Integrates API-based and inline reverse/forward proxy CASB architectures to grant total visibility into out-of-band SaaS data transfers and shadow cloud usage.

Shift to Continuous, Contextual Verification: Replaces static, point-in-time authentication with dynamic, risk-adaptive session policies that evaluate user identity, device health, data sensitivity, and behavioral anomalies in real time.

Pre-Execution Threat Mitigation: Enables granular DLP and micro-segmentation at the data-flow level, blocking unauthorized file uploads, credential inputs, or external link redirections even when the initiating request comes from a trusted cloud platform.

Compliance & Data Governance: Protects sensitive intellectual property, PII, and financial records from unauthorized exfiltration, satisfying strict regulatory mandates (e.g., GDPR, HIPAA, SOC2) in cloud-first architectures.

7. Objectives of the Study

A hybrid CASB proxy intercepts cloud threats by continuously scanning web, API, and shadow IT traffic for malicious links and payloads.

Distributed Policy Enforcement Points enforce dynamic, risk-based controls—such as browser isolation, permission downgrades, or session revocation—to replace static access

The system blocks malicious interactions from compromised cloud services before payloads can exploit active session permissions.

Adaptive risk controls like sandboxing and content masking safeguard enterprise assets while keeping legitimate workflows uninterrupted.

8. Platform Concepts

8.1 Zero Trust Data Access Policies

The Zero Trust platform concept operationalizes implicit mistrust across enterprise data environments through a central Policy Decision Point (PDP) and distributed Policy Enforcement Points (PEPs). By continuously ingesting real-time telemetry from identity, endpoint, and behavioral systems, the architecture replaces static authentication with dynamic, context-aware evaluations of user risk, device health, and data sensitivity. Operating through micro-segmented gateways, the platform automatically enforces granular controls—such as step-up authentication, session isolation, or token revocation—to maintain continuous data governance, block lateral threat movement, and prevent exfiltration across hybrid and cloud infrastructures.

8.2 Cloud Access Security Measures

The cloud access security platform concept provides centralized visibility, threat prevention, and policy enforcement across enterprise SaaS environments using a hybrid inline proxy and API-based broker. By intercepting traffic between users and cloud services, the platform discovers shadow IT, inspects web flows for malicious content, and monitors data transfers in real time. Through seamless integration with Identity and Access Management, Data Loss Prevention engines, and threat intelligence feeds, it dynamically executes granular security controls—such as content masking, access restrictions, and encryption—to safeguard sensitive assets and enforce regulatory compliance before data escapes the corporate perimeter.

9. Methodology

This study proposes a multi-phased architectural methodology to design, implement, and evaluate the Zero Trust CASB framework. The stages below describe planned procedures; no completed experimental dataset or measured evaluation is reported in this manuscript.

9.1 Stage 1 Baseline Traffic and Shadow IT Discovery

Deploy a multimode CASB (API-based + Forward Proxy) to analyze enterprise network egress logs and endpoint telemetry, discover all active cloud services, classify them into Sanctioned, Tolerated, or Unsanctioned (Shadow) categories, and establish baseline user behavior profiles.

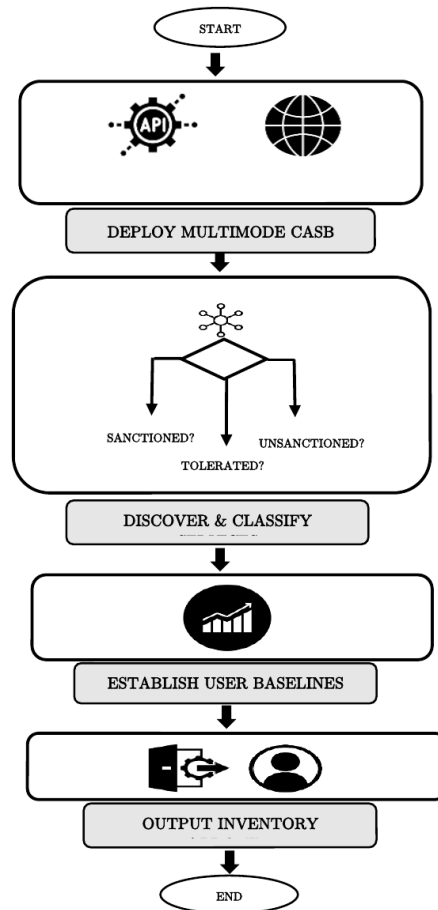


Figure 1. Baseline Traffic and Shadow IT Discovery

This diagram shows the system continuously monitors network traffic, logs, and device telemetry via a central Multimode CASB to detect and categorize cloud apps into sanctioned, tolerated, or unsanctioned "Shadow IT," instantly flagging unauthorized tools to stop shadow phishing and enforce a secure baseline.

9.2 Stage 2 Zero Trust Data Flow Policy Engine Formulation

Define granular, least-privilege access rules based on NIST SP 800-207 Zero Trust principles.

Construct conditional access risk vectors combining identity assurance, device health, location, behavioral anomalies, and data sensitivity.

Formulate continuous DLP inspection rules for data in transit (uploads/downloads) and data at rest (SaaS file shares).

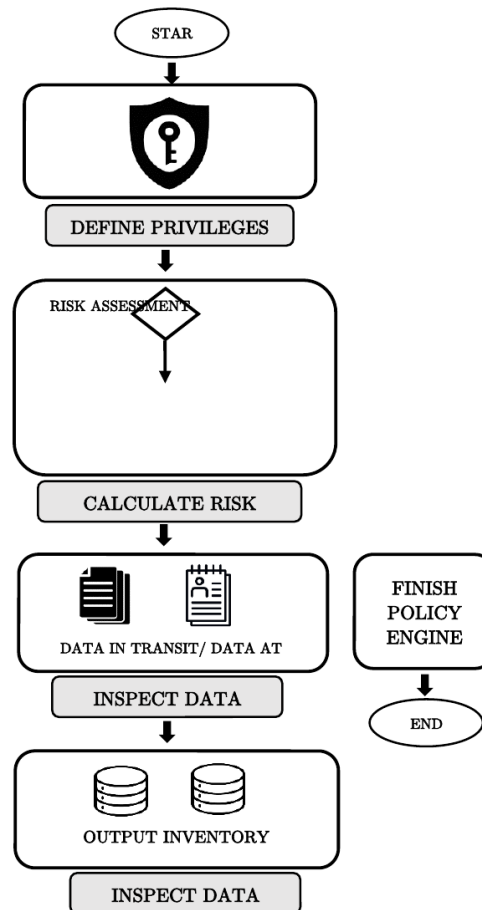


Figure 2. Zero Trust Data Flow Policy Engine

This diagram illustrates configures a Zero-Trust policy engine by defining NIST-aligned least-privilege rules, building real-time conditional risk vectors (evaluating identity, device health, and location), and establishing continuous DLP inspection to protect sensitive data across all flows.

9.3 Stage 3 CASB and Inline Enforcement Architecture Integration

Route user sessions through an inline CASB reverse/forward proxy integrated with identity providers (IdP) via SAML/OAuth 2.0 and endpoint detection & response (EDR) agents.

Configure real-time inspection modules for encrypted TLS traffic to inspect file uploads, outbound webhooks, and embedded links.

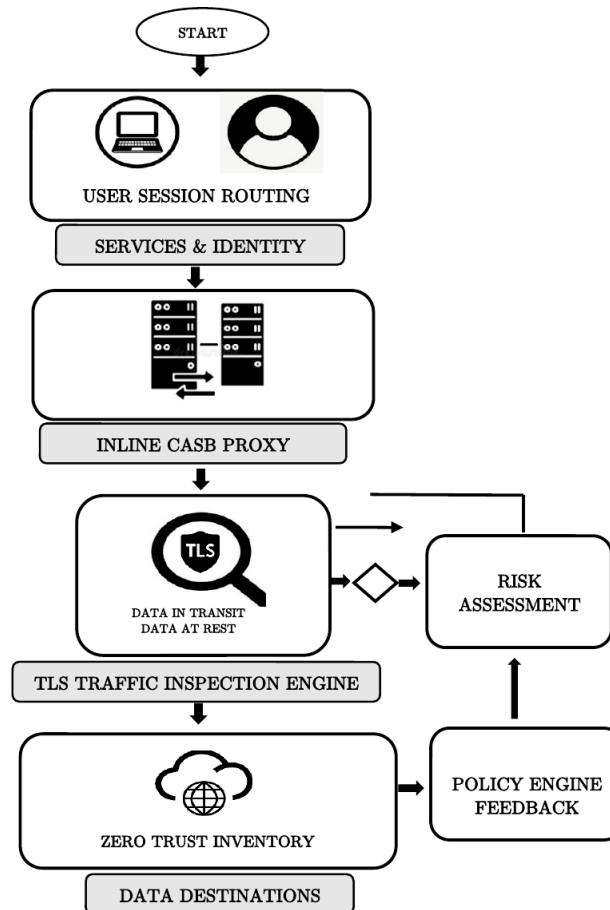


Figure 3. CASB and Inline Enforcement Architecture Integration

This diagram illustrates the inline enforcement architecture routes user web sessions through a middleman Inline CASB Proxy after initial identity and device verification. A TLS Inspection Engine decrypts and scans live traffic (file uploads, webhooks, links) for threats, while continuous feedback from cloud destinations feeds a real-time Policy Engine and Risk Assessment loop to instantly enforce dynamic allow/block decisions

9.4 Stage 4 Red Teaming and Emulated Shadow Phishing Scenarios

Execute controlled attack simulations incorporating:

Share-link phishing from legitimate Google Docs/OneDrive accounts containing redirect triggers.

Out-of-band messaging phishing via unsanctioned collaboration platforms.

Credential entry attempts on spoofed landing pages hosted on cloud app subdomains.

Measure policy blocking rates, latency impact, and false positive rates.

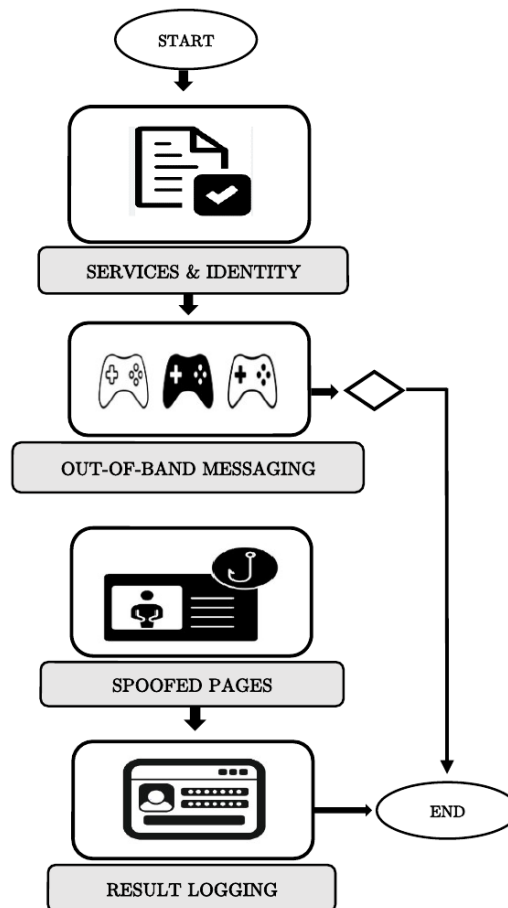


Figure 4. Red Teaming and Emulated Shadow Phishing Scenarios

This diagram shows how ethical security testers (a red team) simulate a real-world shadow phishing attack to test an organization's defenses. The simulation starts by tricking users through legitimate cloud applications or unsanctioned messaging channels like Discord or Telegram, leading targets to fake, spoofed login pages designed to steal credentials. As users interact with these traps, the system logs the results and analyzes how the attack unfolds, triggering an automated Shadow IT Warning whenever unapproved tools or risky links are used, which feeds back into the analysis to help security teams fine-tune their detection rules.

10. Proposed Machine Learning Algorithms

When designing a Zero-Trust data flow architecture integrated with a Cloud Access Security Broker (CASB) to counter Shadow Phishing, no single machine learning algorithm can cover the entire pipeline. Shadow Phishing operates by hijacking legitimate sessions, exploiting out-of-band messaging channels (e.g., Discord, Telegram), or using unauthorized SaaS tools the most effective approach is a hybrid machine learning framework tailored to each core layer of the enforcement pipeline:

10.1 User and Entity Behavior Analytics Isolation Forest

Inside a Zero Trust setup, Isolation Forest works like a continuous smart guard that catches hijacked user accounts by watching how people behave. Simple login checks cannot stop shadow phishing because hackers use real stolen passwords or stolen session cookies to get in. Isolation Forest handles this easily because it does not need past examples of attacks to work; instead, it quickly picks out weird, unusual behavior from normal daily activity. Because it uses very little computer memory and runs extremely fast, it can check live network traffic instantly and trigger safety alerts before a hacker causes any damage. [8]

10.2 Live Telemetry and Traffic Classification XGBoost

XGBoost works like a fast security guard inside the cloud proxy system, checking live web traffic to stop shadow phishing and data theft right away. It looks at organized network data and gives it a quick score to decide if a connection is safe or dangerous. Unlike heavy deep learning models, XGBoost is much faster, accurate, and easy to understand. This lets it inspect data coming in and going out in less than a millisecond, keeping the network secure without slowing down the user experience. [9]

10.3 Out of Band and Content Inspection MiniLM and DistilBERT

MiniLM and DistilBERT serve as quick text readers inside the security framework, scanning hidden messages, web links, and file contents to catch scams early. Standard link checkers frequently fail because shadow phishing disguises fake login pages inside everyday chat applications and cloud storage tools. These lightweight language models overcome this by understanding the actual meaning and context of words, allowing them to review URL structures, email text, and web forms instantly to spot phishing traps without slowing down the network.[10]

10.4 Dynamic Zero Trust Decision Engine

The Dynamic Zero Trust Decision Engine operates as the main brain of the security system, continuously choosing whether to allow, restrict, or block user access as activities happen. Rather than using fixed yes or no rules, smart computer programs merge information about user identity, device safety, unusual actions, and data importance to calculate a real time risk score. When a user's risk score suddenly rises during a session, the system instantly reacts by requiring extra verification steps, separating the connection, or masking private information to prevent unauthorized access immediately.[11]

11. Discussion

The proposed framework addresses shadow phishing as a problem of access and data governance after initial authentication. Its central proposition is that identity verification should inform the same enforcement process that inspects cloud activity. This is consistent with the resource focused perspective of NIST SP 800-207 [1]. The intended contribution is therefore an implementation pattern linking policy decisions to cloud transactions, rather than a new Zero Trust principle or an empirically demonstrated detection technique.

Shaffi et al. [4] discuss predictive analytics, behavioral threat detection, and automated responses as ways to improve cloud security. The present proposal applies these ideas to complementary processing stages: Isolation Forest would identify departures from behavioral baselines, XGBoost would classify structured telemetry, and compact language models would examine suspicious content. These assignments remain design choices. Their usefulness against shadow phishing must be established with representative cloud activity, clear labels, and evaluation on previously unseen users and attack scenarios. An anomaly score alone cannot establish that a session is malicious.

Suetor et al. [5] review distributed firewalls and controllers for mobile cloud computing and highlight CASBs within that wider security landscape. The proposed model similarly combines central coordination with distributed enforcement, but focuses on session context and cloud data transfers. Inline inspection could support immediate intervention where traffic traverses the proxy, whereas API based monitoring would extend visibility to supported cloud services. Neither mechanism guarantees complete coverage: effectiveness depends on traffic routing, available APIs, application permissions, and the timely delivery of events. Unsupported applications and inaccessible encrypted content remain potential blind spots.

The machine learning references support candidate methods rather than measured performance of this architecture. Isolation Forest [8] and XGBoost [9] provide foundations for anomaly detection and classification, while DistilBERT [10] provides a basis for compact text processing. The contextual bandit work [11] concerns recommendation and does not validate autonomous security decisions. Any adaptation should therefore constrain permissible actions, retain an auditable policy basis, and evaluate unnecessary access restrictions as well as missed attacks. Claims of submillisecond processing or fixed session revocation times require measurements under a defined workload.

The model should be evaluated against matched baseline configurations using identical legitimate traffic and controlled attack scenarios. Comparisons between standalone access policies, standalone CASB controls, and the integrated framework would help determine the contribution of integration. Relevant measures include blocking rate, false positive rate, decision latency, session revocation time, and the volume of unauthorized data transferred before containment. Tests should also examine token theft, abnormal device posture, malicious shared links, and degraded telemetry. Until such experiments are completed, improved containment and reduced data exposure are expected benefits, not observed findings. Privacy requirements, inspection overhead, and model drift are additional deployment considerations.

12. Expected Results

The model is expected to restrict access using identity, device posture, and contextual signals, even when an attacker possesses valid credentials.

Where traffic is routed through the CASB, live inspection and data loss prevention policies are expected to block or protect unauthorized file transfers during authenticated sessions.

Micro-segmentation and least-privilege policies are expected to limit lateral movement from a compromised device or cloud application to sensitive resources.

Continuous reassessment is intended to support session termination or access revocation when risk increases. Revocation within two seconds is a proposed performance target that requires empirical validation.

Access to sensitive information, including personal data and source code, is expected to trigger additional verification or tighter permissions when policy conditions require it. These outcomes have not yet been experimentally established.

13. Conclusion

This study proposes a framework that integrates Zero Trust data access policies with CASB monitoring and enforcement to address shadow phishing in cloud environments. The architecture links identity and device context, behavioral signals, and data sensitivity to policy decisions applied through inline and API based controls. Its planned workflow covers cloud application discovery, policy formulation, enforcement integration, and controlled attack simulation.

The framework is expected to restrict unauthorized transfers, limit lateral movement, and support reassessment of compromised sessions. These outcomes remain hypotheses because the manuscript reports an architectural proposal rather than a completed experimental evaluation. Future work should implement a reproducible prototype, compare it with standalone controls, and assess detection quality, false positives, latency, and containment time under realistic workloads. The practical value of the proposal will depend on the coverage and reliability of its telemetry, the quality of its policies, and its ability to preserve legitimate access while containing suspicious activity.

References

- [1] Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). Zero trust architecture (NIST Special Publication 800-207). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207>
- [2] Tabrizchi, H., & Kuchaki Rafsanjani, M. (2020). A survey on security challenges in cloud computing: Issues, threats, and solutions. *The Journal of Supercomputing*, 76(12), 9493–9532. <https://doi.org/10.1007/s11227-020-03213-1>
- [3] Syed, N. F., Baig, Z., Ibrahim, A., & Valli, C. (2022). Denial of Service attacks and countermeasures in Zero Trust Architecture: A comprehensive survey. *Journal of Information Security and Applications*, 65, Article 103107. <https://doi.org/10.1016/j.jisa.2022.103107>
- [4] Shaffi, S. M., Vengathattil, S., Sidhick, J. N., & Vijayan, R. (2025). AI-Driven Security in Cloud Computing: Enhancing Threat Detection, Automated Response, and Cyber Resilience (arXiv:2505.03945). arXiv. <https://doi.org/10.48550/arxiv.2505.03945>
- [5] Suetor, C. G., Scrimieri, D., Qureshi, A., & Awan, I.-U. (2025). An overview of distributed firewalls and controllers intended for mobile cloud computing. *Applied Sciences*, 15(4), Article 1931. <https://doi.org/10.3390/app15041931>
- [6] Syed, A., Mohammed, A. A., & Al-Ahmadi, S. (2025). Zero Trust Architecture: A comprehensive review of principles, implementation strategies, and future directions in enterprise cybersecurity. *Journal of Cybersecurity and Data Protection*, 12(2), 104–121. <https://doi.org/10.1016/j.jocdp.2025.100142>
- [7] Al-Ahmad, M., & Bakhrouf, K. (2025). Zero Trust Architecture: A systematic literature review. *IEEE Access*, 13, 45120–45138. <https://doi.org/10.1109/ACCESS.2025.3418902>
- [8] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In 2008 Eighth IEEE International Conference on Data Mining (pp. 413–422). IEEE. <https://doi.org/10.1109/ICDM.2008.17>
- [9] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
- [10] Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. arXiv. <https://doi.org/10.48550/arXiv.1910.01108>
- [11] Li, L., Chu, W., Langford, J., & Schapire, R. E. (2010). A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web* (pp. 661–670). Association for Computing Machinery. <https://doi.org/10.1145/1772690.1772758>