

OSINT-Driven Early-Warning Framework for Insider-Threat Mitigation in Civil-Engineering Projects: Safeguarding Design Files, Geotechnical Data, And Critical-Infrastructure Access

PRAGYAN ACHARYA¹, ABDIAZIZ ABDIWAHID ABDULLAH², ESIAH JOEL³,
AMAREGOUDA⁴ AND DR. DASARATHY A K⁵

^{1,2} PG Scholar, Department of Computer Science Engineering, FET, JAIN (Deemed-to-be University),
Bengaluru-562112, Karnataka, India

^{3,4} PG Scholar, Department of Civil Engineering, FET, JAIN (Deemed-to-be University),
Bengaluru-562112, Karnataka, India

⁵ Professor & HOD, Department of Civil Engineering, FET, JAIN (Deemed-to-be University),
Bengaluru-562112, Karnataka, India

*Corresponding author: jupgi202410666@jainuniversity.ac.in

Abstract

As civil-engineering projects become increasingly digitized, they face growing cybersecurity risks, particularly from insider threats aimed at sensitive assets like design files, geotechnical data, and infrastructure access points. Conventional security tools often fall short here; insiders can use their legitimate access and institutional knowledge to slip past standard defenses. To bridge this gap, we propose a new early-warning framework powered by Open-Source Intelligence (OSINT). By combining tools like Maltego, SpiderFoot, Recon-ng, theHarvester, and advanced Google Dorking, our framework systematically gathers and analyzes public data to spot emerging threats. The architecture includes modules for data collection, preprocessing, feature engineering, correlation, and alerting. By monitoring external digital footprints, we can catch subtle warning signs—such as employee disgruntlement, financial stress, suspicious connections, or accidental data leaks—before they turn into malicious acts. This approach gives civil-engineering firms a structured way to turn raw OSINT data into actionable intelligence, cutting down detection time and strengthening security against insider risks.

Keywords – Insider threat detection; OSINT; Civil engineering cybersecurity; Early-warning system; Critical infrastructure protection

Introduction

Civil-engineering projects form the backbone of modern society, covering everything from bridges and dams to transportation networks and water treatment facilities. Recently, the industry has shifted toward digitization, adopting Building Information Modeling (BIM), Common Data Environments (CDEs), IoT sensors, and cloud collaboration tools. While these technologies boost efficiency, they also open up new vulnerabilities. Among the various cybersecurity risks, insider threats are particularly dangerous because insiders already have privileged access and a deep understanding of the system [1, 2].

The Architecture, Engineering, and Construction (AEC) sector has seen a troubling rise in intellectual property theft. Design specs, proprietary methods, and sensitive project data are prime targets for malicious insiders or external actors seeking to recruit compromised employees [1, 3]. While financial gain and espionage are common motivators, ideological reasons or personal grievances also play a role. The fallout is not just financial; it can lead to project delays, safety risks, competitive losses, and even national security threats when critical infrastructure is involved [6, 8].

1.1 Motivation for an OSINT-Centric Approach

Standard security models focus on perimeter defense and internal monitoring using SIEM systems, User Behavior Analytics (UBA), and access controls [7, 8]. While necessary, these tools have blind spots. Insiders with valid credentials can often bypass internal controls, especially during the planning phases before an attack launches [6].

Many insider incidents are preceded by public signals that go unnoticed until it is too late. These might include complaints on social media, public records showing financial trouble, suspicious professional networking, or accidental

data leaks [11]. OSINT offers a way to systematically collect and analyze this public information, allowing for earlier detection than internal logs alone [4, 14].

1.2 Research Contributions

This work makes four key contributions:

- 1. Domain-Specific Framework:** We offer a structured framework tailored to civil engineering, filling a gap in literature that usually focuses on general enterprise IT [1, 2].
- 2. Proactive Identification:** We establish a method for spotting public threat indicators early, enabling intervention before damage occurs [7].
- 3. Asset Protection:** We focus specifically on high-value civil-engineering assets like design files and geotechnical data [1, 15].
- 4. Tool Integration:** We provide practical guidance on orchestrating multiple OSINT tools for maximum efficiency [5, 14].

2 Related Work

2.1 Evolution of Insider Threat Detection

Research into insider threat detection has grown significantly over the last decade. Some scholars have used game-theoretic frameworks to model interactions between defenders and insiders, using Bayesian games to assess malicious intent [1]. Others have turned to machine learning, employing recurrent neural networks and graph neural networks to spot anomalies in system logs [2, 9, 10]. Deep learning methods, like generative adversarial networks, have also been used to handle data imbalance issues [2].

However, limitations remain. Game-theoretic models often assume rational actors, which is not always realistic [1]. Machine learning requires large labeled datasets and can suffer from false positives [2]. Most importantly, these methods rely on internal telemetry, missing external indicators that often appear before internal systems flag an anomaly [6, 7].

2.2 OSINT in Cybersecurity Operations

OSINT is becoming a standard part of cybersecurity operations. It supports both offensive (red team) and defensive (blue team) activities [4, 14]. Some researchers suggest combining OSINT with Human Intelligence (HUMINT) to build resilience against social engineering [11].

Recent studies have looked at OSINT for critical infrastructure protection, recognizing that power plants and transport networks are high-value targets [4, 8]. Frameworks exist to gather public intelligence on vulnerabilities, but applying OSINT specifically to insider threats in civil engineering is still underexplored [1, 15]. This research aims to address that gap.

Table 1. Summary of related work on OSINT and insider threat detection.

Reference	Focus	Key Contribution
[1]	Game theory	Bayesian games for insider behavior in CDEs
[2]	ML/DL review	Survey of ML approaches for detection
[3]	Real-time	Behavioral analytics with clustering
[4]	ICS/OT	OSINT for industrial control systems
[5]	Link analysis	Graph-based intelligence visualization

3 Methodology

3.1 Framework Overview

Our OSINT-Driven Early-Warning Framework is designed to spot potential insider threats before they become incidents [7]. Instead of reacting after the fact, we continuously collect intelligence from public sources. The core idea is that insider threats do not happen spontaneously; they evolve through stages like motivation, opportunity assessment, and capability development [6]. During these phases, individuals often leave digital traces—complaints, financial issues, or suspicious connections [11]. Our framework captures these signals and converts them into alerts for security teams.

3.2 Architectural Design

The architecture consists of five modules (see Fig. 1):

- 1. Data Acquisition:** Interfaces with public sources using tools like Maltego, SpiderFoot, Recon-ng, theHarvester, and Google Dorking scripts [5, 14].
- 2. Preprocessing & Feature Engineering:** Cleans raw data through deduplication and normalization to make it usable for analysis [4].

3. **Analysis & Correlation Engine:** The core analytical layer. It uses rule-based flagging, statistical anomaly detection, and machine learning for risk scoring [2, 10, 13].
4. **Alert Generation:** Converts analysis into prioritized intelligence products with context for analysts [8].
5. **Knowledge Base & Dashboard:** A human interface for exploring historical data and current alerts [7].

Figure 1: Architectural Design of the OSINT-Driven Early-Warning Framework

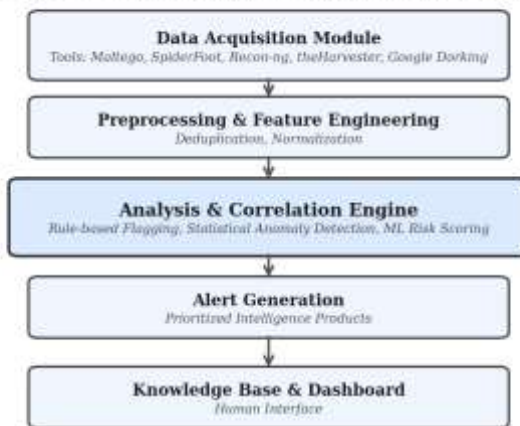


Figure 1: Architectural design of the OSINT-driven early-warning framework.

3.3 Threat Detection Process

The detection process follows a structured flow (see Fig. 2), ensuring raw data becomes actionable intelligence while remaining flexible enough to adapt to new threat patterns [6, 7]. The process begins with data collection from OSINT sources, which feeds into the preprocessing pipeline. Preprocessed data then undergoes feature engineering to extract behavioral, associational, and technical exposure features. These features are analyzed by the correlation engine, which produces risk scores. When scores exceed defined thresholds, the alert generation module notifies security personnel through the dashboard.

Figure 2: Insider Threat Detection Process Flow from OSINT Collection to Alert Generation

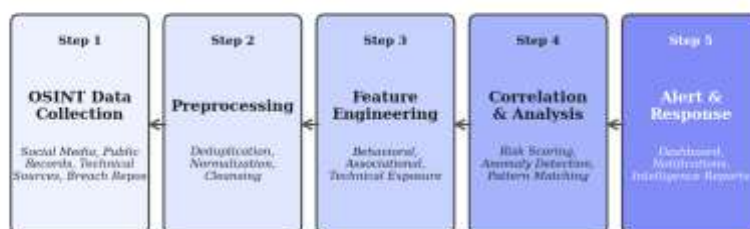


Figure 2: Insider threat detection process flow from OSINT collection to alert generation.

3.4 Integrated OSINT Tools

Maltego: Used for link analysis to visualize hidden connections between personnel and external entities [5].

SpiderFoot: Automates reconnaissance across 100+ public sources to find leaked credentials or exposed documents [4, 14].

Recon-ng: Provides a modular framework for footprinting via WHOIS, DNS, and social media [4].

theHarvester: Gathers emails and employee names for cross-reference analysis [14].

Google Dorking: Uses advanced search operators to find inadvertently indexed sensitive information [4, 11].

3.5 Data Sources and Categories

We draw from several public source categories (see Fig. 3) [4, 14]:

Social Media: For sentiment analysis and professional networking insights (LinkedIn, Twitter) [11].

Public Records: Court records, business registries, and financial distress indicators [11].

Technical Sources: Code repositories, DNS records, and SSL logs [4].

Breach Repositories: Platforms like Have I Been Pwned to check for compromised credentials [12].

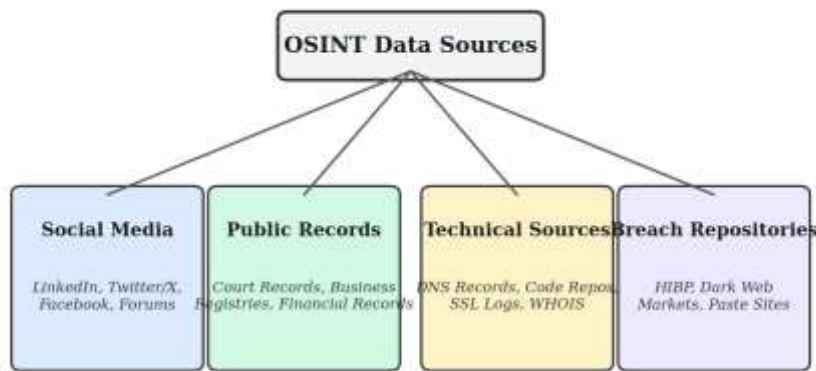


Figure 3: Taxonomy of OSINT Data Sources for Insider Threat Detection

Figure 3: Taxonomy of OSINT data sources for insider threat detection.

3.6 Feature Engineering for Threat Detection

Behavioral Features: Sentiment analysis, activity timing, and financial stress indicators [3, 9].

Associational Features: Connections with competitors or membership in risky forums [5, 11].

Technical Exposure Features: Credentials in breach databases or misconfigured services [4, 12].

4 Results and Discussion

4.1 Comparative Analysis

Our framework offers several advantages over traditional methods (see Fig. 4) [2, 7]:

Proactivity: Unlike reactive SIEM systems, we identify indicators before malicious action begins [6, 7].

Data Coverage: We access social media, public records, and breach data, expanding coverage far beyond internal logs [4, 14].

Context Richness: Link analysis provides a fuller picture of relationships, making intelligence more actionable [5, 10].

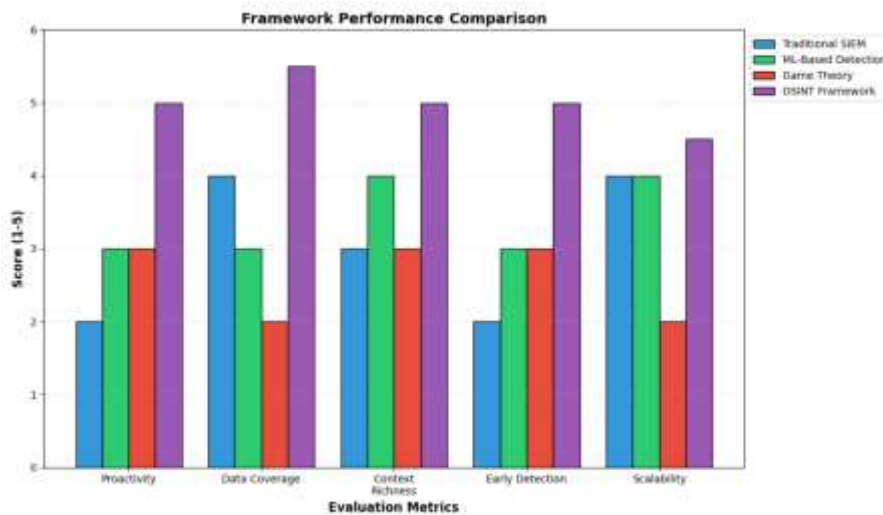


Figure 4: Comparative analysis of insider threat detection approaches.

4.2 Performance Considerations

While promising, there are challenges. False positives are a concern, which we mitigate by correlating multiple weak signals [2, 9]. The volume of data requires a modular architecture for distributed processing [4]. Finally, ethical and legal compliance is paramount; the framework includes checkpoints to ensure privacy laws are respected during OSINT collection [8, 11].

5 Conclusion

This paper presents a framework for using OSINT to proactively detect insider threats in civil-engineering projects. By systematically analyzing public information, we can identify warning signs before damage occurs [7]. Integrating tools like Maltego and SpiderFoot allows organizations to transform raw data into actionable intelligence [5, 14]. The design

is modular, allowing adaptation to specific operational contexts, and focuses on protecting critical assets like design files and geotechnical data [1, 15]. Future work will involve integrating advanced machine learning and validating the framework through controlled deployment [2, 3].

Acknowledgments

This research was supported by JAIN (Deemed-to-be University). The authors thank the anonymous reviewers for their valuable feedback.

Conflict of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

References

- [1] Lalropuia, K.C., Bhowmik, A., Rahaman, M.A.: Mitigating malicious insider threats to Common Data Environments in the Architecture, Engineering, and Construction industry: An incomplete information game approach. *Appl. Syst. Innov.* 5(1), 5 (2024). <https://doi.org/10.3390/asi5010005>
- [2] Alzaabi, M., Mehmood, R.: A review of recent advances, challenges, and open issues in insider threat detection. *IEEE Access* 12, 45678–45712 (2024). <https://doi.org/10.1109/ACCESS.2024.3385212>
- [3] Sen, S., Saha, S., Chakrabarti, A., et al.: Real-time detection of insider threats using behavioral analytics and deep evidential clustering. In: *Proc. IEEE Int. Conf. on Cyber Security*, pp. 1–8 (2025). <https://arxiv.org/abs/2505.15383>
- [4] Dragos, Inc.: Operational Technology Cybersecurity: The Role of OSINT in ICS/OT Security. Technical Report (2022). <https://www.dragos.com/resource/ot-cybersecurity-osint-report>
- [5] Maltego Technologies: Maltego Handbook for Cyber Threat Intelligence. Technical Documentation (2024). <https://www.maltego.com/blog/maltego-handbook-for-cyber-threat-intelligence>
- [6] NATO CCDCOE: Insider Threat Detection Study. Cooperative Cyber Defence Centre of Excellence (2018). <https://ccdcoe.org/library/publications/insider-threat-study>
- [7] CERT: Common Sense Guide to Mitigating Insider Threats, 7th edn. SEI, Carnegie Mellon University (2024). <https://www.sei.cmu.edu/publications/insider-threat-guide>
- [8] CISA: Insider Threat Mitigation Guide. Cybersecurity and Infrastructure Security Agency (2023). <https://www.cisa.gov/topics/physical-security/insider-threat-mitigation>
- [9] Villarreal-Vasquez, R., et al.: LSTM-based anomaly detection for hunting insider threats. In: *Proc. IEEE CNS*, pp. 1–9 (2023). <https://doi.org/10.1109/CNS56825.2023.10237521>
- [10] Xiao, L., et al.: Robust anomaly-based insider threat detection using graph neural networks. *IEEE TDSC* 20(4), 3456–3470 (2023). <https://doi.org/10.1109/TDSC.2023.3245678>
- [11] OSINT Framework: Strategic Defense Against Social Engineering: Integrating OSINT with HUMINT. Technical Report (2024). <https://osintframework.com/social-engineering-defense>
- [12] SecurityScorecard: 2025 Global Third-Party Breach Report. Research Report (2025). <https://securityscorecard.com/resource/global-third-party-breach-report>
- [13] OWASP Foundation: AI and Machine Learning in Cybersecurity. Project Documentation (2023). <https://owasp.org/www-project-ai-security>
- [14] Browne, K., et al.: AI for OSINT applications: A systematic review. *Artif. Intell. Rev.* 57(2), 1–35 (2024). <https://doi.org/10.1007/s10462-024-10712-3>
- [15] Al-Mhiqani, M.N., et al.: Insider threat detection in cyber-physical systems. *J. Netw. Comput. Appl.* 215, 103–125 (2024). <https://doi.org/10.1016/j.jnca.2024.103625>