iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VI **Month of publication:** June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72574

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Automation and Orchestration in Cyber Threat Intelligence (CTI): A Survey

Anauksa Das
Computer Science & Engineering Dept. Sikkim Manipal Institute of Technology

Abstract: ***Cyber Threat Intelligence (CTI) has become an essential part of security operations with the sudden rise in cyber threats. To manage the collection, enrichment, correlation, and response workflows for CTI, several automation and orchestration technologies, such as SOAR (Security Orchestration, Automation, and Response) platforms and CTI pipelines are utilised. This survey reviews open-source as well as commercial SOAR platforms e.g. Splunk SOAR, Cortex XSOAR, IBM QRadar SOAR, TheHive, Shuffle) as well as frameworks used in CTI such as MISP and OpenCTI, comparing their features, integration capabilities, and limitations. Research done in the last 5-7 years on automated CTI processing, including methods for improving and linking data, AI and machine learning-driven analysis, and full-system architectures, is highlighted in this paper. Standard formats like STIX/TAXII and related sharing protocols help ensure different systems can communicate effectively. A comparison table shows the difference between major platforms based on key aspects such as data formats, integrations, response capabilities, and AI/ML support. Common challenges like data compatibility, source reliability, data quality, processing speed, and scalability are also discussed. Finally, a classification of automation components along with an example orchestration architecture, illustrated (Figure 1) is presented. The survey concludes with an overview of current challenges and potential future developments in CTI automation and orchestration.***

Index Terms: ***Security Orchestration, Automation, and Re- sponse (SOAR); Cyber Threat Intelligence (CTI); STIX; TAXII; automation; pipeline; threat intelligence sharing; AI/ML.***

I. INTRODUCTION

Cyber Threat Intelligence (CTI) helps organizations detect and prevent cyber-attacks by collecting, analysing, and sharing threat information. Security Operations Centres (SOCs) are overwhelmed with alerts, logs, and external data, making man- ual threat analysis impractical [1]. To address this, Security Orchestration, Automation, and Response (SOAR) systems have emerged, bringing together different security tools, au- tomating repetitive tasks, and accelerating incident response [1]. By integrating data from global threat databases, open- source intelligence (OSINT), and internal monitoring systems, SOAR-based CTI workflows enhance alerts with valuable con- text and trigger automated actions to mitigate risks. However, traditional SOAR implementations have limitations. Many rely on rigid, no-code systems that lack flexibility and scalability, have difficulty capturing the complexity of evolving threats, and require constant playbook updates, increasing operational workload [1]. Recognizing these challenges, researchers and industry experts have been exploring AI-driven orchestration frameworks to improve CTI automation. This survey provides a comprehensive review of the current state of automation in Cyber Threat Intelligence, covering commercial and open- source platforms, recent academic research, industry standards, and the end-to-end architectures that power automated CTI workflows.

II. SOAR PLATFORMS FOR CTI

A. Commercial SOAR Solutions

Leading security companies provide SOAR platforms that help security teams manage incidents more efficiently. These platforms come with built-in playbooks, app integrations, and case management tools. Splunk SOAR (formerly Phantom) connects with more than 300 third-party security tools and supports over 2,800 automated actions [3]. It uses frameworks like MITRE ATT&CK and D3FEND, along with machine learning for threat scoring which is backed by Splunk's Threat Research Team [3]. The platform can be deployed on-premises, in the cloud, or in hybrid environments, and integrates seamlessly with Splunk Enterprise Security to unify security workflows [3]. Palo Alto Cortex XSOAR offers over 900 prebuilt integration packs and thousands of actions for custom playbooks [4]. Its visual playbook editor allows secu- rity teams to automate processes, while its "war room" feature supports investigations with chat and command-line interfaces. The platform leverages machine learning for triage, automated documentation, and enrichment from Unit42 threat feeds [4].

One case study suggests that around 60% of incidents could be automatically closed using its 200+ prebuilt playbooks. IBM QRadar SOAR (formerly Resilient) focuses on ease of use, featuring a low-code playbook designer and strong case management tools. Its dashboards provide analysts with rich context on incidents, while its drag-and-drop dynamic playbooks allow security teams to build automated responses in minutes without coding [5]. The platform offers more than 300 integrations through its App Exchange, with many sample playbooks for fast deployment [5]. Recent updates add support for structured data formats like JSON and advanced playbook looping [5]. While these commercial SOAR platforms pro- vide broad integrations, support for data formats like JSON and STIX and powerful playbook engines, they also have limitations. These platforms face challenges such as vendor lock-in, high costs, and the ongoing need for human expertise when dealing with unique or evolving threats. Despite their automation capabilities, they still rely on security teams to handle unpredictable scenarios and adapt playbooks to new attack patterns.

B. Open-Source Orchestration and CTI Tools

Some open-source projects provide similar automation fea- tures, especially for Cyber Threat Intelligence (CTI). TheHive is a free Security Incident Response Platform (SIRP) often used alongside MISP for threat investigation. It can import security alerts from SIEMs, emails, and phishing detections, and turn them into cases for analysis [7]. It makes it possible for multiple analysts to collaborate on cases, use prebuilt task templates, and document findings in lab notebooks. TheHive integrates with Cortex analysers, allowing users to attach automated analysis tasks, such as virus scans and IP reputation lookups, to indicators within a case, making threat enrichment more efficient [7]. Shuffle is a newer open-source SOAR platform designed for ease of use. It offers a visual workflow editor and encourages community-driven app sharing [8]. Users can create automation playbooks that trigger security actions using OpenAPI schemas. Shuffle includes many pre- built app connectors for Slack, email, and security APIs and supports multi-tenant organization management. Other open- source CTI frameworks include MISP and OpenCTI:

- MISP is primarily a threat intelligence sharing platform. It allows organizations to store and exchange indicators and threat attributes using structured formats such as tags and taxonomies. MISP data can be exported to STIX, OpenIOC, or IDS rules, enabling automated security responses [6]. It also provides data correlation between events and automated exports to SIEMs/IDS to enhance IoC automation [6].
- OpenCTI is a knowledge-graph-based CTI platform. It collects threat data from external feeds, vulnerability databases, and internal alerts, organizing it into an in- teractive graph. This structured data can be used for automated detection rules in security systems like SIEMs and firewalls. As of mid-June 2025, OpenCTI 6.6.16 is the latest update. It builds on the automation introduced in version 5.11 and has evolved into a robust platform for cybersecurity automation [9].

While open-source solutions are free and customizable, they often require more setup and integration effort compared to commercial SOAR platforms.

C. Capabilities and Limitations

Table 1 (see Sec. VII) summarizes the features of key platforms. SOAR solutions generally offer strong integration libraries, graphical playbook editors, and case management tools. They help security teams standardize response processes and reduce manual workloads. However, many platforms still rely on maintained connectors and struggle to handle unconventional threats. As noted in [5], traditional no-code SOAR platforms often lack advanced logic and become less efficient as playbooks grow larger [1]. Open-source CTI tools focus on knowledge sharing but may require additional or- chestration layers for automated responses. Some platforms include machine learning and AI features. For example, Splunk SOAR uses ML-based risk scoring [3], and XSOAR applies AI-assisted triage [4]. However, fully autonomous AI-driven security automation is still an active research area.

III. STANDARDS AND FRAMEWORKS FOR CTI AUTOMATION

For Cyber Threat Intelligence (CTI) automation to work efficiently, it depends on open data models and sharing pro- tocols. STIX (Structured Threat Information eXpression) is the OASIS standard used to represent CTI elements such as indicators, malware, tactics, techniques, procedures (TTPs), and attack campaigns [10]. Its structured JSON-based format allows different security tools to work together seamlessly. TAXII (Trusted Automated eXchange of Indicator Informa- tion) is the primary protocol for sharing STIX data [11]. The latest version, TAXII 2.1, introduces API endpoints that enable organizations to send and receive threat intelligence efficiently, helping scale CTI sharing across platforms [11]. The OASIS CTI Technical Committee promotes automated threat data exchange to improve security awareness, real-time defence, and advanced threat analysis [10].

Most SOAR (Security Orchestration, Automation, and Response) platforms and Threat Intelligence Platforms (TIPs) can import and export STIX data. Many security agencies also provide STIX-based feeds. Although other formats like CSV, JSON, OpenIOC, and company-specific APIs are sometimes used, STIX/TAXII is now the most common way to share structured threat intel- ligence. Additionally, frameworks like MISP taxonomies and galaxies (such as MITRE ATT&CK and TLP) help organize threat information. Another standard, OpenC2, is emerging to define automatic responses to cyber threats, which could work alongside STIX/TAXII. Despite these open standards, some challenges remain. Not all tools fully support the STIX 2.x format, and some vendors add extra custom fields, making data harder to share. Maintaining data quality and trust across different sources is another issue, as discussed in Section VI.

IV. AUTOMATED CTI PROCESSING AND RESEARCH ADVANCES

Recent academic research has focused on automating key stages of the CTI lifecycle: data collection, enrichment, cor- relation, and response. One common challenge is that raw threat data is overwhelming and often contains irrelevant or redundant information. Filtering and enriching this data intelligently is crucial. Spyros et al. highlight that many conventional CTI systems simply gather and store information without deeper analysis [13]. To address this, they propose ThreatWise AI, a framework that integrates tools such as web crawlers and parsers while leveraging MISP and AI/ML algorithms to enrich CTI records before sharing them [13].

A. Enrichment and Correlation

Enrichment adds valuable context to raw threat data, such as IP geolocation, domain registration details, or related vulnera- bilities (CVEs). Research has applied machine learning (ML) and graph-based techniques to enhance enrichment: ETIP (Enriched TIP) by Gonza´lez-Granadillo et al. correlates open- source intelligence (OSINT) feeds with real-time network data, scoring and prioritizing alerts [2]. This system matches incoming indicators against known benign or malicious infras- tructure before forwarding high-quality intelligence to SIEMs and security partners [2]. HeteroCTI, proposed by Jin et al., builds a heterogeneous graph of CTI artifacts, using graph- mining techniques to assess the reliability of threat intelligence feeds based on source reputation. This approach helps reduce false positives and directs analyst attention to more relevant threats. ML techniques have been applied to CTI in several ways. Natural Language Processing (NLP) helps extract indi- cators and attack behaviours from unstructured threat reports. Some systems use supervised classifiers or neural networks to tag text with CTI entities. ML-based anomaly detection identifies unusual patterns in streaming threat data—for exam- ple, clustering unlabelled indicators of compromise (IOCs) to detect emerging cyber campaigns. Studies suggest that ML/AI improves threat detection and triage when trained correctly [14]. Alevizos and Dekker argue that AI can automate tasks “from data ingestion to resilience verification”, helping CTI adapt to evolving threats [14]. However, bias and transparency remain concerns in ML models.

B. Automated Response

Research has also explored automating defensive actions based on CTI insights. Some systems automatically block malicious IPs or domains by feeding indicators into firewalls or endpoint detection & response (EDR) blocklists. AWS prescriptive guidance outlines a workflow where ingested CTI updates intrusion prevention systems (IPS), while observed malicious activity feeds intelligence back into the system [12]. Other approaches build closed-loop automation. For example, phishing email triage systems can extract URLs, analyse them, and quarantine messages without requiring human input. Despite these advances, most academic studies emphasize the importance of human-in-the-loop safeguards, where automated actions require analyst approval before execution. In summary, research highlights the importance of fully connected CTI sys- tems that bring together data collection, enrichment, analysis, and automation to help SOCs work more efficiently. New tools like ThreatWise AI [13] and OpenCTI [9] playbooks show how different parts can work together, from gathering raw data to scoring threats and automatically sending updates to security systems. While these improvements make CTI workflows more scalable, full automation is still uncommon, and researchers are working on making AI-powered security decisions more reliable.

V. ORCHESTRATION ARCHITECTURES AND PIPELINES

A typical CTI automation system is made up of four main parts: data collection, processing, storage, and action. Figure 1 presents a general model where a central threat intelligence platform, hosted either in the cloud or on-premises, gathers data from security agencies and trusted sources and shares it with internal security tools [12]. Data collection involves receiving security feeds through TAXII or APIs, where they are structured into STIX format for consistency.

Once col- lected, the data goes through an enrichment and correlation stage, where internal logs, sandbox results, or external APIs add useful context to the threat indicators. Some systems rank threats based on risk or trust levels to help prioritize responses. The enriched intelligence is then stored in a central database that makes searching and analysis easier. Finally, security systems use this processed data to take automated actions, such as blocking malicious IPs, sending alerts to SIEMs, or creating incident reports for further investigation.

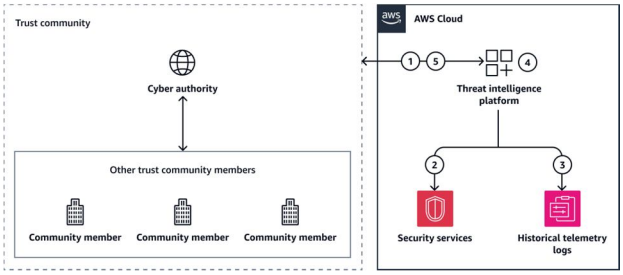


Fig. 1. Example CTI sharing and orchestration architecture on AWS [12].

Figure 1 shows an example CTI sharing and orchestration architecture on AWS [12]. It receives threat data from trusted sources, updates security tools within the organization, collects new intelligence from AWS security services, refines and analyzes the data, and then shares updated intelligence with other security communities [12].

Modern Security Operations Centers (SOCs) use similar approaches, where SOAR platforms automate tasks through playbooks, and CTI platforms like **OpenCTI** or **MISP** store and manage intelligence. Many organizations rely on cloud- based or microservices-based architectures to ensure scalabil- ity.

Some research projects incorporate real-time data stream- ing, containerized threat analysis, and graph-based storage to enhance CTI automation. For example, ThreatWise AI uses Docker-based crawlers to collect threat intelligence, Elastic- search for storage, and TensorFlow for AI-powered classifi- cation [13]. OpenCTI pipelines rely on GraphQL queries and Python-based automation to improve threat intelligence when new data arrives [9]. A well-structured CTI automation system follows key design principles such as modularity, allowing new tools or AI models to be easily integrated, and event-driven triggers, ensuring that any new threat data can automatically activate security responses while still allowing for human oversight. These advancements help security teams work faster and more efficiently by connecting data, analysing threats, and automating defences.

VI. CHALLENGES IN CTI AUTOMATION

Even with advancements, automating Cyber Threat Intelli- gence (CTI) still faces several challenges:

1) Data compatibility: Many security tools use different formats and don't always support STIX/TAXII. Con- necting CTI with older SIEM systems often requires custom scripts or adapters. Even when standards are used, differences in versions or meanings can cause data loss. Playbooks also need frequent updates to work smoothly across different systems.

2) Source reliability: CTI feeds vary in accuracy, and outdated or incorrect indicators can waste resources. Some researchers suggest rating sources and comparing threat data with local logs to improve reliability [2] [13]. Poisoned feeds, which are deliberately altered threat intelligence, can be misleading. Trust mechanisms, such as approved source lists and data validation, are still being refined. The ETIP study suggests combining OSINT with internal data to make intelligence more reliable [2].

3) Data quality and enrichment: Raw CTI data often in- cludes duplicate or irrelevant information, making enrich- ment important. Adding extra details—like an IP's loca- tion, domain registration, or known vulnerabilities—helps focus on important threats. Scoring models and graph analytics are being explored to make threat intelligence more accurate and reduce false positives.

4) Processing speed: CTI needs to be processed and shared fast to stay relevant. Delays can weaken security re- sponse. A Ponemon study in 2017 found that compa- nies took an average of 206 days to detect a breach [15]. Automation aims to bring this down to hours, but complex analysis and human approvals can slow things down. Stream processing and incremental updates are being tested to make CTI faster.

5) Scalability: Millions of Indicators of Compromise (IoCs) are created every day, so systems need to handle large amounts of data efficiently. Traditional SOAR platforms often struggle as playbooks and datasets expand [1]. Running multiple enrichment tasks at the same time and using big-data tools like Kafka and Spark can help manage the load, but balancing resources properly is still tricky.

6) Trust and governance: Blocking an IP or domain based on CTI can lead to false positives and liability concerns. Many workflows include manual approvals to prevent mistakes. Sharing CTI across organizations also comes with legal and privacy challenges, such as GDPR- compliant data tagging in STIX, which security tools must support.

These challenges show that automation alone isn't enough. A strong CTI system requires careful planning, human over- sight, and ongoing validation to remain accurate and effective.

VII. COMPARISON OF MAJOR PLATFORMS AND FRAMEWORKS

Table 1 provides a comparison of major commercial SOAR platforms and open-source CTI frameworks based on impor- tant features. The Input/Output formats section outlines the types of data each platform can process and share, such as JSON, STIX, CSV, or API-based integrations. Integrations highlight built-in connectors to security tools like SIEMs, EDRs, and external threat intelligence feeds. Response ca- pabilities focus on automation features, including playbooks, case management, and API-based controls. Lastly, AI/ML sup- port indicates whether the platform includes machine learning features like threat scoring or anomaly detection.

TABLE I
COMPARISON OF COMMERCIAL SOAR PLATFORMS AND OPEN-SOURCE CTI FRAMEWORKS

Platform / Framework	Type	Formats (I/O)	Integrations & Connectors	Response / Capabilities	AI/ML Support
Splunk SOAR	SOAR	JSON, REST APIs, STIX (via apps)	300+ apps (SIEM, EDR, firewall, cloud), Splunk ES, threat intel feeds	Graphical playbooks, case management, dashboards	ML-based risk scoring
Cortex XSOAR	SOAR	JSON, REST, STIX (via plug-ins)	900+ automation packs (EDR, email, cloud, DB), Unit42 intel	Visual playbooks, "war room" collaboration, autodocumentation	ML-assisted triage (phishing alerts)
IBM QRadar SOAR	SOAR	JSON, REST, Python scripts	300+ integrations (via IBM App Exchange)	Dynamic drag-drop playbooks, case/ticketing, reporting	Not emphasized (rule- based logic)
TheHive	CTI/SIRP	JSON (its API), MISP format	MISP sync, Cortex analyz- ers, SIEM alerts (via The-Hive4py)	Case management, task tem- plates, multi-analyst collabo- ration	No built-in ML (relies on analyzer tools)
MISP	CTI	JSON, STIX, CSV, Ope- nIOC	TIP feeds, other MISP in- stances, IDS/IPS (via ex- ports)	Sharing platform: store IOCs, taxonomies, automated exports to SIEM/IDS	No (focus on sharing and correlation)
OpenCTI	CTI	JSON/GraphQL (STIX2 schema)	Threat feeds (FS-ISAC, CIRCL), intel APIs, NLP modules	Knowledge graph, reports & cases, detection-as-code feeds	Plans NLP/AI compo- nents (in development)
Shuffle	SOAR	JSON, REST, webhooks	Apps via OpenAPI (e.g., Slack, email, VirusTotal, MISP)	Low-code visual workflows, webhook triggers	No built-in ML (pipelines can call ML APIs)
Other commercial: FortiSO	*AR, Siemplify*	*, D3, etc., with similar capa*	*bilities*		

The table highlights key differences between commercial SOAR platforms and open-source CTI tools. Commercial SOARs focus on broad integration with security tools and automation through playbooks, often incorporating machine learning for threat detection and response. In contrast, open- source CTI solutions like MISP and OpenCTI prioritize data sharing and knowledge management, serving as central repos- itories for CTI rather than full orchestration platforms. Not all products fully support STIX/TAXII; many still depend on pro- prietary or JSON-based APIs for data exchange. While future frameworks may aim to unify these capabilities, organizations today typically use a combination of SOAR and CTI tools to ensure coverage across all phases of the threat intelligence lifecycle [2] [9].

VIII. FUTURE DIRECTIONS AND OPEN CHALLENGES

The field of CTI automation is evolving quickly, with several exciting developments shaping its future:

1) AI-Driven Orchestration: One major area of research is AI-driven orchestration. Security experts are exploring how advanced AI models, including large language mod- els (LLMs), can create dynamic playbooks and make se- curity decisions on their own [1] [14]. Future SOAR plat- forms might be able to suggest and even execute multi- step security actions without manual scripting. However, this brings concerns about trust and transparency. Human oversight is still crucial, and researchers are working on intent-based automation where analysts define their goals rather than coding specific steps.

2) Enhanced Data Fusion: Another promising direction is enhanced data fusion. Combining CTI with other data sources, such as network activity, endpoint security logs, and darknet intelligence, can provide richer insights into cyber threats. Techniques like graph-based data linking and federated learning may help detect advanced threats more effectively. As standards evolve, future versions like STIX 3.x may introduce universal data formats that allow better consistency across different platforms.

3) Trustworthy Intelligence: Ensuring trustworthy intelli- gence is becoming a priority as cybercriminals increas- ingly spread false information. Automated security tools need ways to verify the authenticity of CTI. Technologies such as blockchain and secure logging could help validate threat data, while AI may assist in identifying low-quality or misleading feeds. Research is also focusing on defining CTI provenance metadata, which tracks the origin of threat intelligence for better reliability.

4) Real-Time Pipelines: Real-time pipelines are being de- veloped to reduce delays in processing threat intelligence. Event-driven architectures, including serverless comput- ing and live-streamed data processing, can help make CTI more immediate and actionable. Incorporating automated CI/CD workflows in CTI pipelines ensures that security updates don't introduce errors or delays. Researchers are also working on self-healing systems that can detect and fix broken integrations automatically, improving overall system reliability.

5) Usability and Human Factors: Usability and human factors remain an important focus. Security automation must fit naturally into analyst workflows by offering clear visualizations of automated processes, audit logs for accountability, and intuitive ways for analysts to interact with automated intelligence. Developing a struc- tured guide for CTI automation that classifies tasks from manual to fully automated could help shape future tools.

-

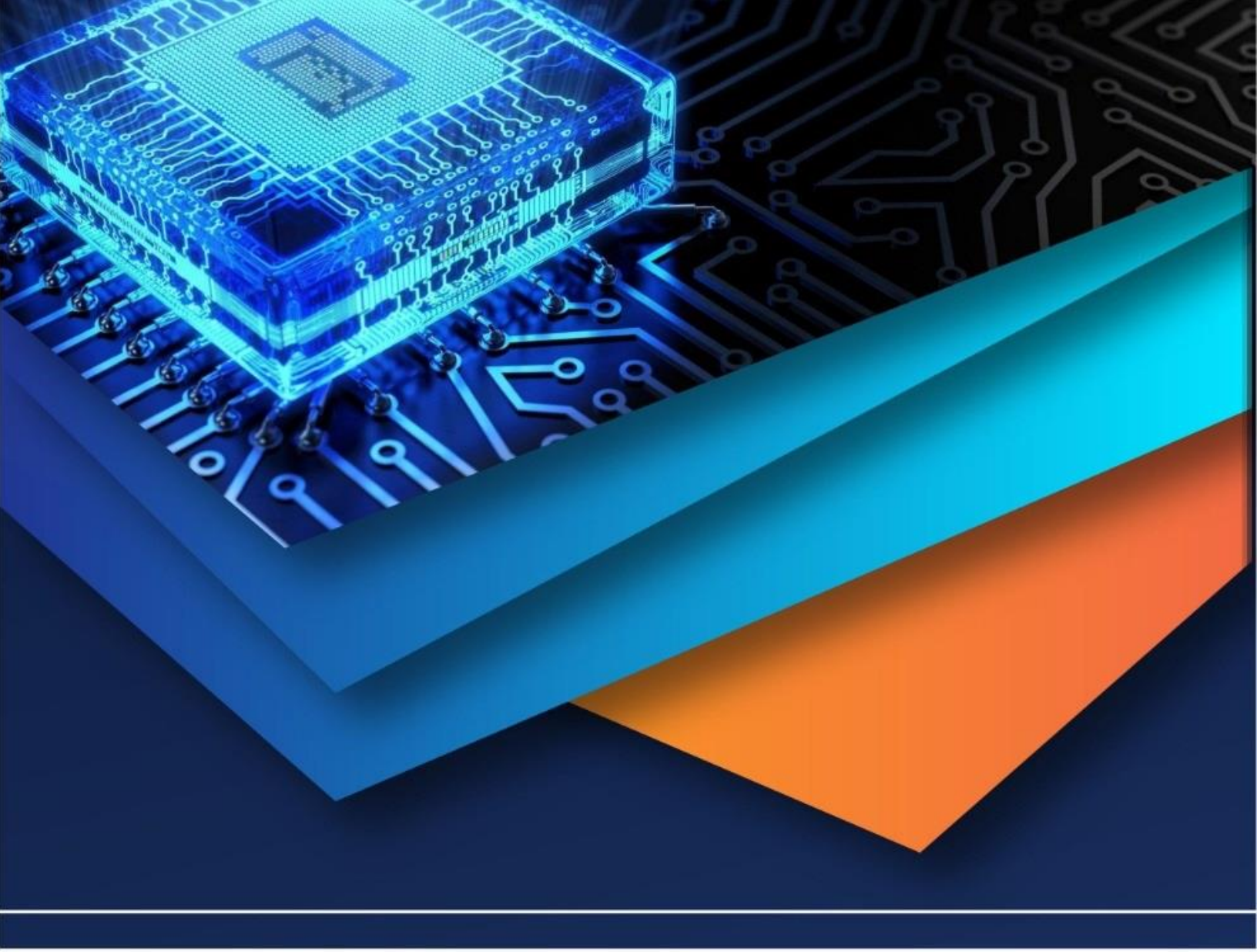
IX. CONCLUSION

While automation and orchestration are key to scaling CTI operations, complete autonomy is still a long way off. Future advancements will need to balance cutting-edge AI technology with practical security operations, ensuring that automated threat intelligence remains accurate, fast, and trustworthy.

REFERENCES

[1] Pantelopoulos et al., "Toward Robust Security Orchestration and Auto- mated Response in SOCs with a HyperAutomation Approach," *Infor- mation*, vol. 16, no. 5, 2025. (MDPI)
[2] G. Gonza´lez-Granadillo et al., "ETIP: An Enriched Threat Intelligence Platform for improving OSINT correlation, analysis, visualization and sharing capabilities," Journal of Information Security and Applications, vol. 58, 2021, article 102715.
[3] Splunk Inc., "Splunk Security Orchestration, Automation and Response (SOAR)," product brief, 2024.
[4] Palo Alto Networks, "Cortex XSOAR: Automate your manual work- flows," product page, 2024.
[5] IBM, "QRadar SOAR Platform – Features," IBM product documenta- tion, 2024.
[6] F. Lamanna and V. Pelissier, "MISP – The Malware Information Sharing Platform & Threat Sharing," MISP Project (website), 2025.
[7] Fre´de´ric By et al., "TheHive: A scalable open-source Security Incident Response Platform," TheHive Project (GitHub README), 2025.
[8] Shuffle Project, "Shuffle: Open source automation platform," GitHub README, 2025.
[9] Filigran (OpenCTI) Blog, "Introducing threat intelligence automation and playbooks in OpenCTI," Oct. 2023.
[10] OASIS CTI Technical Committee, "STIX™ Version 2.1," standard specification, 2021. See also "Introduction to STIX" (OASIS).
[11] OASIS, "TAXII™ Version 2.1," standard specification, 2021.

[12] AWS, “Cyber threat intelligence architecture on AWS,” AWS Prescrip- tive Guidance, 2023.
[13] S. Spyros et al., “AI-based Holistic Framework for Cyber Threat Intelligence Management,” IEEE Access, Jan. 2025.
[14] A. Lampis and M. Dekker, “Towards an AI-Enhanced Cyber Threat Intelligence Processing Pipeline,” Electronics, vol. 13, May 2024.
[15] X. Shu, “Threat intelligence computing for efficient cyber threat hunt- ing,” Tech Xplore, Oct. 17, 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY