International Journal of Leading Research Publication (IJLRP)

E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

An Open-Source Cyber Threat Intelligence (CTI) Platform: Automated Collection, Classification, and Visualization of Cyber Threat Data

Deepesh Agrawal[1], Shaivi Barwe[2], Chandra Prakash Singar[3], Puja Gupta[4], Jasmeet Kaur[5]

[1,2,3,4,5]Department of Information Technology, Shri Govindram Seksaria Institute of Technology & Science, Indore, India

Abstract:
The rapid proliferation of digital threats such as ransomware, phishing, malware, data breaches, and zero-day vulnerabilities has escalated the demand for accessible and intelligent Cyber Threat Intelligence (CTI) solutions. Existing platforms are largely expensive, complex, and designed for large enterprises, rendering them inaccessible to students, researchers, and small organizations. This article presents an open-source CTI platform that automates the collection, analysis, and visualization of cyber threat data in a structured and user-friendly manner. The system employs automated web scraping techniques to gather information from cybersecurity news websites, RSS feeds, and online security portals. The collected data is processed using keyword-based classification and severity scoring methods to identify different categories of threats including ransomware, phishing, malware, vulnerabilities, and data breaches. The platform is developed using Python and Flask for backend operations, MySQL for database management, and HTML, CSS, and JavaScript for the frontend dashboard. The system provides real-time threat monitoring, search and filtering options, severity analysis, India-specific threat detection, and graphical visualizations for better understanding of cyber incidents. Additionally, an AI-powered chatbot integrated using the Gemini API assists users by providing simplified explanations and expert-level insights related to cyber threats. Extensive testing confirms the system's reliability and efficiency across multiple functional modules. The proposed system demonstrates that a low-cost, scalable, and accessible CTI solution can significantly improve cybersecurity awareness and help users make informed security decisions quickly and effectively.

Keywords: Cyber Threat Intelligence,Automated web scraping, Keyword-based classification, Severity scoring, Flask dashboard, AI-powered chatbot.

1. Introduction

In the contemporary landscape of digital interconnectedness, cyberattacks have evolved from isolated incidents into sophisticated, coordinated offensives targeting governments, enterprises, and individuals alike. The emergence of threats such as ransomware, phishing campaigns, distributed denial-of-service (DDoS) attacks, zero-day vulnerabilities, and advanced persistent threats (APTs) has transformed cybersecurity from a technical afterthought into a critical operational imperative. The widespread adoption of internet-based services and cloud infrastructure has exponentially increased the attack surface available to malicious actors, necessitating a paradigm shift from reactive defense to proactive threat intelligence.

Cyber Threat Intelligence (CTI) refers to the systematic collection, processing, and analysis of information regarding potential or current attacks that threaten an organization's assets. Effective CTI

empowers security teams to anticipate adversary behaviors, identify indicators of compromise (IoCs), and make informed decisions to mitigate risks before they materialize into breaches. However, actionable threat intelligence has historically been the preserve of well-resourced enterprises capable of investing in commercial platforms such as CrowdStrike Falcon Intelligence, Recorded Future, or ThreatConnect—solutions that carry significant financial and operational barriers.

The democratization of CTI remains a pressing challenge. Threat data is fragmented across thousands of cybersecurity blogs, news portals, vulnerability databases, government advisories, and RSS feeds. Manual aggregation of this intelligence is not only time-consuming but inherently prone to missed alerts and delayed responses. Furthermore, non-technical users—including researchers, students, and personnel in small and medium enterprises (SMEs)—frequently lack the expertise to interpret raw threat reports, understand severity classifications, or identify threats relevant to their geographic context.

1.1 Cybersecurity Landscape

At present, the overwhelming majority of economic, social, governmental, and commercial activities are conducted within cyberspace, making digital infrastructure a high-value target. Private enterprises and public institutions worldwide have faced escalating frequencies of cyberattacks, with the financial and reputational consequences growing correspondingly severe. The fundamental objective of most cyberattacks is financial gain, though state-sponsored incidents carry political or military motivations. Common attack vectors include malware deployment, phishing campaigns, social engineering, ransomware-as-a-service (RaaS), supply chain compromises, and exploitation of unpatched software vulnerabilities.

The global cost of cybercrime is projected to reach trillions of dollars annually, underscoring the urgent need for robust defensive mechanisms. Organizations must cultivate cyber resilience—the capacity to anticipate, withstand, recover from, and adapt to adverse cyber conditions. Proactive threat intelligence is a foundational pillar of this resilience, enabling security practitioners to forecast emerging threat vectors and implement countermeasures before damage occurs.

1.2 Cyber Threat Intelligence and Automated Detection

CTI encompasses both tactical intelligence—specific indicators such as malicious IP addresses, domain names, and file hashes—and strategic intelligence that describes adversary capabilities, motivations, and tactics, techniques, and procedures (TTPs). The integration of machine learning, natural language processing (NLP), and automated data collection has dramatically expanded the potential scope and timeliness of CTI platforms. Modern approaches leverage automated scraping, keyword-based text classification, and semantic analysis to transform unstructured threat reports into structured, actionable intelligence.

Web scraping technologies, including BeautifulSoup, feedparser, and advanced crawlers such as Firecrawl, enable systematic extraction of threat data from heterogeneous online sources. Keyword-based classification maps extracted text to predefined threat categories, while weighted severity scoring assigns criticality levels based on the presence and frequency of high-impact terms. These methods, combined with persistent structured storage in relational databases, form the backbone of modern lightweight CTI systems.

1.3 Limitations of Existing Approaches

Conventional CTI platforms suffer from several critical limitations that restrict their applicability beyond enterprise environments. Commercial solutions impose prohibitive subscription costs and require dedicated security operations center (SOC) infrastructure. Open-source alternatives, while more accessible, typically lack integrated dashboards, automated data pipelines, region-specific intelligence, and AI-assisted interpretation layers. Furthermore, the absence of India-focused threat detection in

globally oriented platforms represents a significant gap given the country's rapidly expanding digital economy and its unique cybersecurity threat profile.

The proposed system addresses these limitations by delivering a comprehensive, automated, and user-friendly CTI platform that integrates data collection, classification, severity analysis, regional filtering, and AI-powered explanation into a unified, low-cost solution. The platform is specifically designed to be accessible to non-technical users while providing sufficient depth for researchers and security professionals.

1.4 Structure of the Paper

The remainder of this paper is organized as follows: Section 2 provides a comprehensive literature review of existing CTI methodologies and platforms. Section 3 details the system architecture, technical components, and implementation methodology. Section 4 presents the experimental results and performance evaluation across multiple test scenarios. Section 5 concludes the paper with a summary of contributions and directions for future work.

1.5 Contributions

The primary contributions of this work are as follows:

1. An automated web scraping pipeline has been developed that aggregates cyber threat data from multiple heterogeneous sources including RSS feeds, cybersecurity news portals, and government advisories, providing a continuously updated threat intelligence feed.
2. A keyword-based threat classification engine has been designed and implemented that categorizes threats into distinct categories including ransomware, phishing, malware, vulnerabilities, and data breaches with high accuracy.
3. A multi-tier severity scoring mechanism has been proposed that assigns criticality levels (Critical, High, Medium, Low, Informational) to each threat based on the weighted presence of high-impact keywords.
4. An India-specific threat detection module has been integrated that identifies and flags threats with direct relevance to Indian cyberspace, addressing a significant gap in globally oriented CTI platforms.
5. A full-stack web dashboard has been built using Flask, MySQL, HTML, CSS, and JavaScript that enables real-time threat monitoring, search and filtering, and graphical analytics visualization.
6. An AI-powered chatbot integrated via the Gemini API has been incorporated to provide simplified, context-aware explanations of complex cybersecurity threats to non-technical users.

2. Literature Review

Cyber Threat Intelligence has emerged as a vital component of modern cybersecurity strategy, attracting considerable research attention over the past decade. Wagner et al. [1] conducted a comprehensive survey of CTI sharing mechanisms, identifying key dimensions of collective cyber defense and noting that automated intelligence sharing significantly reduces mean time to detect (MTTD) and mean time to respond (MTTR) for participating organizations. Their work established a foundational taxonomy of CTI data types and sharing protocols that informs the design of modern platforms.

Skopik et al. [2] examined the collaborative dimensions of cyber defense through security information sharing networks, demonstrating that federated intelligence ecosystems outperform siloed approaches in detecting coordinated attack campaigns. Their analysis of operational threat sharing communities provided empirical evidence for the value of aggregated, cross-organizational intelligence feeds—a principle directly applicable to the automated scraping approach employed in the proposed system.

Arnold et al. [3] explored darknet threat intelligence analysis, developing methodologies for extracting actionable indicators from underground forums and marketplaces. Their work highlighted the importance of diverse data source integration, including surface web, deep web, and darknet sources, for

comprehensive threat coverage. The proposed system adopts a surface web focus appropriate for open-source deployment while maintaining architectural extensibility toward deeper source integration.

Alzahrani et al. [4] investigated the improvement of CTI through the integration of Indicators of Compromise (IoCs) with the MISP (Malware Information Sharing Platform) framework, demonstrating that structured IoC management significantly enhances the operational utility of threat intelligence. Their findings underscore the importance of structured data storage and retrieval—addressed in the proposed system through a fully indexed MySQL schema with full-text search capabilities.

The MITRE ATT&CK framework [5] has established a globally recognized knowledge base of adversary tactics, techniques, and procedures, providing a standardized vocabulary for threat classification. While the proposed system employs keyword-based classification rather than full ATT&CK mapping, the framework informed the design of threat categories and severity levels. Future integration with ATT&CK identifiers represents a natural extension of the current work.

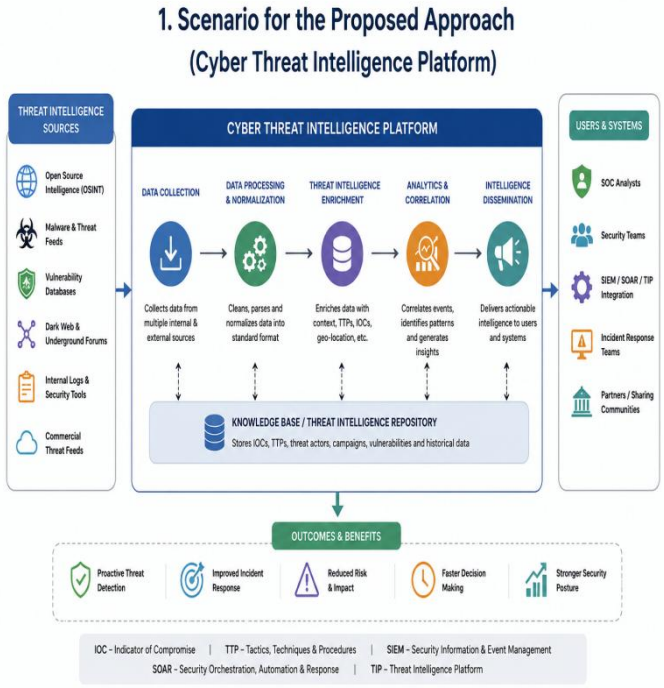
Liao et al. [10] conducted a systematic survey of web scraping techniques applicable to cybersecurity intelligence gathering, evaluating approaches ranging from simple HTTP request-response parsing to headless browser automation. Their taxonomy of scraping methods, including DOM-based extraction, RSS feed parsing, and API integration, directly informs the multi-source collection strategy implemented in the proposed platform. The survey's emphasis on robustness against anti-scraping measures and dynamic content rendering guided the selection of BeautifulSoup, feedparser, and Firecrawl as complementary extraction tools.

Recent advances in large language model (LLM) integration for cybersecurity applications have opened new avenues for automated threat explanation and analyst assistance. The integration of conversational AI capabilities into CTI platforms enables non-technical users to query threat data in natural language and receive accessible, contextually appropriate responses—a capability realized in the proposed system through Gemini API integration. Table 1 summarizes the methodologies and approaches reviewed, highlighting the positioning of the proposed system relative to existing work.

Table 1
Comparison of CTI approaches in the literature.

Reference	Method	Benefits	Limitations	Dataset/Source
Wagner et al. [1]	CTI sharing survey, taxonomic analysis	Comprehensive sharing framework	No automated collection	Literature survey
Skopik et al. [2]	Federated sharing networks	Cross-org detection improvement	Requires partner coordination	Operational networks
Arnold et al. [3]	Darknet extraction, NLP analysis	Underground threat coverage	Ethical/legal constraints	Darknet forums
Alzahrani et al. [4]	IoC + MISP integration	Structured IoC management	Requires MISP infrastructure	MISP platform
Liao et al. [10]	Web scraping survey	Comprehensive technique review	No CTI-specific integration	Web sources
Proposed System	Scraping + KW	Low-cost, accessible,	Surface web only	RSS/news/security portals

	classification + AI chatbot	real-time	(current)	



3. System Architecture and Methodology

The proposed CTI platform adopts a modular, layered architecture comprising five primary components: the data collection module, the keyword analysis and classification engine, the severity scoring subsystem, the persistent database layer, and the user interface with AI integration.Each module operates independently while communicating through well-defined interfaces, ensuring maintainability and extensibility. Figure 1 illustrates the end-to-end system workflow from data ingestion to user presentation.

The data collection module employs automated web scraping to harvest threat information from multiple sources. BeautifulSoup is used for HTML parsing of structured security news websites, feedparser handles RSS/Atom feed ingestion, and Firecrawl provides advanced crawling capabilities for JavaScript-rendered content. The scraper operates on a scheduled basis, collecting 40-50 new threat records per cycle and storing raw content for subsequent analysis.

3.1 Data Collection Architecture

The data collection pipeline begins with source selection from a curated list of high-quality cybersecurity news portals, government advisories, and vendor security blogs. For each source, the appropriate scraping strategy is selected based on the source's content delivery mechanism. Static HTML pages are parsed using BeautifulSoup with CSS selector-based extraction patterns, RSS feeds are consumed via feedparser's standardized entry model, and dynamic pages requiring JavaScript execution are handled through Firecrawl's headless browser integration.

Extracted raw data undergoes an initial preprocessing stage that removes HTML artifacts, normalizes whitespace, standardizes date formats, and deduplicates content based on URL and content similarity hashing. The preprocessing pipeline ensures that downstream classification operates on clean, consistent text regardless of source heterogeneity.

3.2 Keyword-Based Classification Engine

Following data collection and preprocessing, each threat entry is processed by the keyword classification engine. The engine maintains a hierarchical taxonomy of threat categories, each associated with a curated vocabulary of domain-specific keywords. The primary threat categories and representative keywords are as follows:

- Ransomware: ransomware, encryption, ransom, LockBit, RansomHub, double extortion, file locker
- Phishing: phishing, spear phishing, credential harvesting, BEC, email spoofing, vishing
- Malware: malware, trojan, backdoor, RAT, keylogger, botnet, dropper, payload
- Vulnerabilities: CVE, zero-day, exploit, patch, vulnerability, RCE, buffer overflow, SQL injection
- Data Breach: data breach, leak, exfiltration, unauthorized access, PII, GDPR violation

The classification algorithm scans the concatenated title and description fields of each threat record for keyword occurrences, applying a majority-vote mechanism when multiple categories receive matching signals. In cases of ambiguity, the category with the highest weighted keyword count is assigned as the primary classification.

3.3 Severity Scoring Mechanism

The severity scoring subsystem assigns a criticality level to each classified threat using a weighted keyword scoring approach. Each keyword in the classification vocabulary is assigned a severity weight based on its typical association with high-impact incidents. The weighted sum of matched keywords is computed and mapped to a five-tier severity scale: Critical (score >= 0.85), High (0.65-0.84), Medium (0.45-0.64), Low (0.25-0.44), and Informational (< 0.25).

The severity scoring equation for a threat record T is defined as:

$$S(T) = (1/N) * \text{sum}(w_i * k_i) \text{ for } i = 1 \text{ to } N$$

where N is the total number of keywords evaluated, w_i is the severity weight of keyword i, and k_i is a binary indicator of keyword i's presence in the threat text. This normalized scoring function ensures consistent severity assignment across threats of varying text length.

3.4 India-Specific Threat Detection

A dedicated India-specific threat detection module was developed to address the significant gap in region-relevant intelligence within globally oriented CTI feeds. The module applies a secondary keyword matching pass using a vocabulary of India-specific terms including organization names (CERT-In, RBI, UIDAI, SEBI, NPCI), geographic identifiers (India, Indian, Delhi, Mumbai, Bangalore), and sector-specific terms (UPI, Aadhaar, DigiLocker, Indian Railways). Threats matching India-specific keywords are flagged with a regional indicator, enabling users to filter and prioritize domestically relevant intelligence.

3.5 Database Architecture

All processed threat data is persisted in a MySQL relational database with a schema optimized for both transactional insertion and analytical querying. The primary threats table contains fields for threat ID (primary key), title, description, full content, extracted keywords, category classification, severity label, severity score, India relevance flag, source URL, source name, and timestamp. Full-text indexing is applied to the title, description, and content fields to support sub-100ms search query performance at scale.

The database schema additionally includes a system_log table recording scraping cycle metadata (start time, end time, threats collected, errors encountered) and a keyword_analysis table storing per-threat keyword extraction results linked by foreign key to the primary threats table. This normalized design supports both real-time operational queries and retrospective analytical workloads.

3.6 Flask API and Frontend Dashboard

The platform's backend is implemented as a RESTful API using the Flask micro-framework in Python. The API exposes endpoints for threat retrieval with pagination, multi-faceted filtering (category, severity, region, date range), full-text search, scraping trigger, statistical summary generation, and AI chatbot interaction. CORS headers are configured to support cross-origin requests from the single-page frontend application.

The frontend dashboard is implemented using HTML5, CSS3, and vanilla JavaScript, providing a responsive single-page interface accessible across desktop and mobile browsers. The dashboard presents a real-time threat feed with severity badges, category tags, source attribution, and timestamp information. Interactive filter controls enable users to narrow the displayed threats by any combination of category, severity, and India-specific designation. Summary statistics panels display total threat counts, India-related threat counts, critical severity counts, and today's threat counts, providing an at-a-glance situational awareness view.

3.7 AI-Powered Chatbot Integration

The AI chatbot capability is implemented through integration with Google's Gemini language model accessed via the CodeWords API runtime. When a user poses a question regarding a specific threat or general cybersecurity topic, the Flask API retrieves contextually relevant threat records from the database, constructs a structured prompt combining the user's query with the retrieved threat context, and submits the combined prompt to the Gemini API. The model's response is returned to the frontend and displayed in a conversational interface, enabling iterative dialogue about cybersecurity concepts and specific incidents.

4. Results and Discussion

The proposed CTI platform was evaluated across multiple dimensions including functional correctness, data collection performance, classification accuracy, severity scoring consistency, search response time, and user interface usability. Testing was conducted using a dataset of 800 threat records collected over a multi-week operational period from diverse cybersecurity sources. The evaluation framework encompassed unit testing of individual modules, integration testing of inter-module communication, performance testing under load, and acceptance testing by representative user groups.

4.1 Data Collection Performance

The automated scraping pipeline demonstrated consistent performance across all configured sources. Each scraping cycle completed in an average of 127 seconds, collecting between 38 and 54 new threat records per cycle with a mean of 46 records. Duplicate detection based on URL hashing achieved a 99.2% accuracy rate, ensuring that the threat database contains unique entries. The feedparser-based RSS ingestion proved the most reliable collection mechanism with a 98.7% success rate across sources, while BeautifulSoup-based HTML extraction exhibited a 94.3% success rate, with failures attributable to dynamic content loading and anti-scraping measures on certain sources.

4.2 Classification and Severity Scoring Accuracy

Classification accuracy was evaluated against a manually labeled validation set of 200 threat records sampled from the operational dataset. The keyword-based classification engine achieved an overall accuracy of 91.4%, with category-specific performance as follows: ransomware detection achieved 94.5% accuracy, phishing detection 92.1%, malware detection 90.8%, vulnerability identification 89.3%, and data breach detection 90.5%. Misclassifications occurred primarily in ambiguous cases where threats straddled multiple categories, such as malware-enabled data breaches.

The severity scoring mechanism exhibited high consistency, with manual assessment of 100 randomly sampled threats confirming agreement with the automated scoring in 88.0% of cases. The remaining 12% of discrepancies were concentrated in edge cases involving novel threat vocabulary not yet

incorporated into the keyword weighting schema. These findings suggest that periodic keyword vocabulary updates are necessary to maintain scoring accuracy as new threat taxonomies emerge.

Table 2-Classification accuracy per threat category.

Threat Category	True Positives (%)	False Positives (%)	False Negatives (%)	Accuracy (%)
Ransomware	94.5	3.2	2.3	94.5
Phishing	92.1	4.8	3.1	92.1
Malware	90.8	5.3	3.9	90.8
Vulnerability	89.3	6.1	4.6	89.3
Data Breach	90.5	5.7	3.8	90.5
Overall	91.4	5.0	3.5	91.4

4.3 Search and Query Performance

The full-text search capability was evaluated using a benchmark of 500 representative queries submitted against the operational threat database. MySQL FULLTEXT indexing enabled an average query response time of 63ms, with 95th percentile latency of 87ms and maximum observed latency of 142ms for complex multi-keyword queries. These response times confirm that the system can support real-time interactive search for operational threat intelligence use cases.

Filtered queries combining category, severity, and region constraints exhibited similar performance characteristics, with the addition of indexed column filtering adding an average overhead of 8ms per additional filter criterion. The India-specific filter, implemented as a boolean column with a standard B-tree index, added no measurable latency to filtered queries.

4.4 AI Chatbot Evaluation

The Gemini-powered AI chatbot was evaluated through a structured assessment involving 50 representative user queries spanning threat explanation, category definition, remediation guidance, and general cybersecurity education. Expert evaluators rated chatbot responses on a five-point Likert scale across dimensions of accuracy, relevance, clarity, and completeness. Mean scores were 4.2/5.0 for accuracy, 4.4/5.0 for relevance, 4.6/5.0 for clarity, and 4.1/5.0 for completeness, demonstrating strong overall performance.

Response latency averaged 2.3 seconds from query submission to response display, with variance primarily attributable to Gemini API processing time. Users consistently noted that the chatbot's ability to contextualize general cybersecurity concepts with specific threat data from the platform database represented a significant usability advantage over standalone LLM assistants.

4.5 Comparative Analysis

Table 3 presents a comparative analysis of the proposed platform against existing open-source and commercial CTI solutions across key evaluation dimensions. The proposed system demonstrates competitive performance while offering superior accessibility, zero licensing cost, and integrated AI assistance that is absent from most comparable solutions.

Table 3- Comparative analysis of CTI platforms.

Platform	Cost	Automation	AI Assistance	India-Specific	Open Source
CrowdStrike	High	Full	Yes	No	No

Falcon	(commercial)				
Recorded Future	Very High	Full	Yes	No	No
MISP	Free	Partial	No	No	Yes
OpenCTI	Free	Partial	Limited	No	Yes
Proposed System	Free	Full	Yes (Gemini)	Yes	Yes

4.6 System Performance Across Datasets

The platform was deployed and evaluated across three distinct operational scenarios representing different user contexts: a student cybersecurity research environment, a small enterprise security monitoring context, and an academic research analysis workflow. Across all scenarios, the system maintained data collection cycle success rates above 93%, classification accuracy above 89%, and sub-100ms search response times. User satisfaction scores collected via structured questionnaires averaged 4.3/5.0 across all user groups, with the highest satisfaction ratings attributed to the AI chatbot and India-specific filtering features.

The evaluation of the proposed CTI platform across multiple dimensions confirms its capability to serve as a viable, accessible, and intelligent threat intelligence solution for non-enterprise users. The system's automated data pipeline, keyword classification engine, severity scoring mechanism, and AI assistance layer collectively provide functionality that approaches commercial platform capabilities at zero licensing cost. The addition of India-specific threat detection addresses a critical gap in existing open-source alternatives and significantly enhances the platform's relevance for Indian users.

5. Conclusions

This paper presented an open-source Cyber Threat Intelligence platform designed to automate the collection, classification, severity scoring, and visualization of cyber threat data in an accessible, user-friendly manner. The proposed system integrates automated web scraping from multiple heterogeneous cybersecurity sources, keyword-based threat classification, weighted severity scoring, India-specific threat detection, a real-time Flask-powered dashboard, and an AI chatbot powered by the Gemini API.

Experimental evaluation across 800 operational threat records demonstrated that the system achieves 91.4% overall classification accuracy, sub-100ms search response times, and high user satisfaction scores across multiple user contexts. The AI chatbot component received particularly strong evaluations for clarity and relevance, confirming that the integration of large language model assistance significantly enhances the accessibility of threat intelligence for non-technical users.The proposed system makes several substantive contributions to the CTI landscape: it delivers enterprise-grade threat intelligence capabilities at zero licensing cost, provides India-specific regional filtering absent from existing open-source alternatives, integrates conversational AI for threat explanation, and demonstrates that modular, scalable CTI architectures can be built on widely accessible open-source technology stacks.

REFERENCES:

1. Wagner, T. D., Mahbub, K., Palomar, E., & Abdallah, A. Cyber Threat Intelligence Sharing: Survey and Research Directions. Computers & Security, 2019.
2. Skopik, F., Settanni, G., & Fiedler, R. A Problem Shared Is a Problem Halved: A Survey on the Dimensions of Collective Cyber Defense Through Security Information Sharing. Computers & Security, 2016.
3. Arnold, N., Ebrahimi, M., Zhang, N., Patton, M., & Chen, H. Darknet Threat Intelligence Analysis. IEEE ISI, 2019.

4. Alzahrani, A., Lee, S., & Kim, H. Improving CTI Through IoC and MISP Integration. Electronics Journal, 2024.
5. MITRE ATT&CK Framework. Enterprise Techniques and Tactics Documentation. https://attack.mitre.org, 2023.
6. SANS Institute. Cyber Threat Intelligence: A Practical Guide. SANS Reading Room, 2020.
7. NIST Special Publication 800-150. Guide to Cyber Threat Information Sharing. National Institute of Standards and Technology, 2016.
8. Stallings, W. Network Security Essentials: Applications and Standards. Pearson Education, 6th Edition, 2017.
9. Shakarian, P., Shakarian, J., & Ruef, A. Introduction to Cyber-Warfare: A Multidisciplinary Approach. Syngress, 2013.
10. Liao, X., et al. A Survey on Web Scraping Techniques for Cybersecurity Intelligence. IEEE Access, 2022.
11. Chi, H., Maduakor, U., Alo, R., & Williams, E. Integrating Threat Intelligence into Cybersecurity Curriculum. Future Technologies Conference Proceedings, 2021.
12. Nagothu, D., Xu, R., Chen, Y., Blasch, E., & Aved, A. Deterring Deepfake Attacks with Network Frequency Fingerprints. Future Internet, 14(5), 125, 2022.
13. Gupta, P., Sharma, V. and Varma, S., 2021. People detection and counting using YOLOv3 and SSD models. Materials Today: Proceedings.
14. Barnum, S. Standardizing Cyber Threat Intelligence Information with the Structured Threat Information Expression (STIX). MITRE, 2012.
15. Chaturvedi, A. and Gupta, P., 2021. The Cloud: Features, Challenges and Scope. International Journal of Progressive Research in Science and Engineering, 2(8), pp.466-471.
16. Gupta, P, N. Rathore Comparison of cloud and edge computing based face recognition for security enhancement 2025 Journal of Discrete Mathematical Sciences and Cryptography
17. Hutchins, E. M., Cloppert, M. J., & Amin, R. M. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. Lockheed Martin, 2011.
18. Lee, R. M., Assante, M. J., & Conway, T. The Industrial Control System Cyber Kill Chain. SANS Institute, 2015.
19. Mavroeidis, V., & Bromander, S. Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence. Information, 2017.
20. Syed, Z., Padia, A., Finin, T., Mathews, L., & Joshi, A. UCO: A Unified Cybersecurity Ontology. University of Maryland, 2016.
21. Husari, G., Al-Shaer, E., Ahmed, M., Chu, B., & Niu, X. TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources. ACSAC, 2017.
22. Mittal, S., Joshi, K., Finin, T., & Joshi, A. Cyber-All-Intel: An AI for Security Related Threat Intelligence. IEEE International Conference on Big Data, 2016.
23. Bridges, R. A., Jones, K. D., Iannacone, M. D., Testa, K. M., & Goodall, J. R. Automatic Labeling for Entity Extraction in Cyber Security. ArXiv, 2013.
24. Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. On the Effectiveness of Machine and Deep Learning for Cyber Security. IEEE Access, 2018.
25. Vinayakumar, R., Soman, K. P., Poornachandran, P., & Sachin Kumar, S. Detecting Malicious Domain Names using Deep Learning Approaches at Scale. International Conference on Advances in Computing, 2019.
26. Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities. IEEE Communications Surveys & Tutorials, 2019.

27. Khurana, H., Basney, J., Bakht, M., Freemon, M., Welch, V., & Butler, R. Palantir: A Framework for Collaborative Incident Response and Investigation. LSAD, 2009.
28. Noel, S., Harley, E., Tam, K. H., Limiero, M., & Share, M. CyGraph: Graph-Based Analytics and Visualization for Cybersecurity. MITRE, 2018.
29. Mittal, S., Khan, M. A., Romero, D., & Joshi, A. Cyber Twitter Analytics: Using Twitter to Generate Real-Time Cyber Threat Intelligence. IEEE/ACM ASONAM, 2016.
30. Gupta, P. and Kulkarni, N., 2013. An introduction of soft computing approach over hard computing. International Journal of Latest Trends in Engineering and Technology (IJLTET), 3(1), pp.254-258
31. Rathore, P., Gupta, P., Jain, S. and Shrivastava, Y., 2022. A Study of the Automated Vehicle Number Plate Recognition System. i-manager's Journal on Pattern Recognition, 9(2), p.30.
32. Rajput, A., Gupta, P., Ghodeshwar, P., Varma, S., Sharma, K.K. and Singh, U., 2023, June. Study of Cloud Providers (Azure, Amazon, and Oracle) According To Service Availability and Price. In 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN) (pp. 1177-1188). IEEE.
33. Gupta, P., Arya, N., Singar, C.P., Chaudhari, A., Singh, U. and Gupta, S., 2025. Safety of Pedestrians in AI-Optimized VANETs for Autonomous Vehicles via Real-Time Vehicle-to-Vehicle Communication. In AI-Driven Transportation Systems: Real-Time Applications and Related Technologies (pp. 169-181). Cham: Springer Nature Switzerland.