applied sciences

MDPI

Review

Redefining Cyber Threat Intelligence with Artificial Intelligence: From Data Processing to Predictive Insights and Human–AI Collaboration

Mateo Barrios-González [1], Javier Manuel Aguiar-Pérez [2], María Ángeles Pérez-Juárez [2,*] and Enrique Castañeda-de-Benito [3]

[1] Centro de Estudios Superiores y Técnicos de Empresa (CESTE), Paseo de los Infantes de España 3, 50012 Zaragoza, Spain; mbarrios@ceste.com
[2] Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, Universidad de Valladolid, ETSI Telecomunicación, Paseo de Belén 15, 47011 Valladolid, Spain; javier.aguiar@uva.es
[3] Mando Conjunto de Ciberespacio (MCCE), Escuela Militar de Ciberoperaciones (EMCO), Base de Retamares, Crta. de Boadilla del Monte, km. 3.4, Pozuelo de Alarcón, 28223 Madrid, Spain; ecasde@et.mde.es
* Correspondence: mperez@uva.es; Tel.: +34-983-423660

Abstract

The increasing complexity and scale of cyber threats have pushed Cyber Threat Intelligence (CTI) beyond the capabilities of traditional rule-based systems. This article explores how Artificial Intelligence (AI), particularly Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), and graph-based analytics, is reshaping the CTI landscape. By automating threat data processing, enhancing attribution, and enabling predictive capabilities, AI is transforming CTI into a proactive and scalable discipline. By analysing CTI architectures, real-world use cases, platform comparisons, and current limitations, this study highlights the emerging opportunities and challenges at the intersection of cybersecurity and AI. This analysis concludes that the future of CTI lies in hybrid systems that seamlessly combine human expertise with intelligent automation.

Keywords: deep learning; indicators of compromise; machine learning; natural language processing; predictive cybersecurity; threat attribution; threat intelligence platforms

1. Introduction

In an increasingly interconnected world, Cyber Threat Intelligence (CTI) has become a central component of modern cybersecurity strategies. CTI refers to the collection, analysis, and dissemination of information about current and potential cyber threats, including threat actors, their Tactics, Techniques, and Procedures (TTPs), and Indicators of Compromise (IoCs) associated with malicious activities. Its relevance stems from the growing complexity and frequency of cyberattacks, which increasingly require organisations to adopt intelligence-informed and proactive approaches to risk mitigation and defence planning. Rather than replacing traditional security controls, CTI complements them by providing contextual and anticipatory insights that support more informed decision-making.

By enabling organisations to move beyond purely reactive responses, CTI supports earlier detection, prioritisation, and mitigation of cyber threats, thereby reducing the potential impact of incidents and contributing to overall organisational resilience. In practice, CTI is commonly leveraged to inform a range of defensive activities, including threat prioritisation, Incident Response (IR), and strategic planning. Key motivations for adopting CTI include the following:

Academic Editor: Luis Javier García Villalba

Received: 28 November 2025
Revised: 28 January 2026
Accepted: 4 February 2026
Published: 6 February 2026

- Proactive defence against cyber threats: Anticipating potential attacks allows organisations to identify relevant threats before they materialise. By analysing adversaries' motives, tactics, and capabilities, CTI can support risk-informed forecasting and defensive preparation, as well as the prioritisation of resources toward higher-impact threats.
- Reducing the impact of cyber incidents: early identification of threats enables faster and more targeted responses, limiting the time available for attackers to cause harm and reducing potential operational and reputational damage.
- Enhancing IR: CTI provides contextual information that helps understand the actors, objectives, and techniques involved in an incident. Insights derived from prior analyses can also be used to refine response workflows and playbooks, improving consistency and effectiveness over time.
- Aligning defences with adversarial tactics: ongoing CTI analysis helps identify emerging tools, techniques, and attack patterns, enabling defensive measures to evolve in response to observed adversary behaviour.
- Supporting decision-making: CTI informs strategic and tactical decisions by highlighting risks relevant to an organisation's sector, scale, and infrastructure, supporting more transparent and defensible resource allocation.
- Protecting sensitive data and critical assets: by identifying exploitable weaknesses in systems and processes, CTI contributes to risk reduction strategies aimed at safeguarding sensitive information and necessary services.
- Enhancing threat intelligence sharing: CTI facilitates participation in intelligence-sharing initiatives, including open-source platforms such as OpenCTI (Open Cyber Threat Intelligence). The exchange of analysed intelligence can strengthen collective defence against shared threats.
- Supporting compliance and risk management: Regulatory frameworks, including GDPR (General Data Protection Regulation), increasingly require organisations to assess and manage cyber risks. CTI can support these obligations by providing structured threat analysis as part of broader risk management processes.
- Preventing operational disruption: CTI supports business continuity planning by identifying threats that could disrupt critical operations and informing mitigation strategies in advance.
- Strengthening organisational resilience: by integrating technical, procedural, and human factors, CTI encourages a holistic view of cybersecurity and promotes preparedness across organisational roles.

As the volume, velocity, and variety of cyber threat data continue to increase, traditional CTI approaches—often reliant on manual analysis and rule-based systems—face scalability and timeliness limitations. The diversity of data sources, the use of obfuscation techniques by attackers, and the demand for near-real-time intelligence have motivated the exploration of more adaptive and automated approaches. Within this context, Artificial Intelligence (AI) has emerged as a key enabling technology rather than a standalone solution.

The application of AI in CTI has attracted growing attention as a means to enhance threat detection, analysis, and forecasting capabilities [1–4]. Advances in Machine Learning (ML) [5–10] and Natural Language Processing (NLP) [11–14] have influenced how threat intelligence is collected, enriched, and operationalised. ML techniques enable the processing of large and heterogeneous datasets, support pattern recognition, and facilitate data-driven prioritisation and prediction tasks. In this work, the term "predictive" denotes probabilistic, time-aware forecasting capabilities based on historical patterns and behavioural signals, rather than descriptive statistics or correlational analysis. NLP methods allow the extraction of relevant information from unstructured sources, such as Open-Source INTelligence (OSINT). At the same time, graph-based learning techniques support the

correlation and analysis of complex attack campaigns [15]. Collectively, these approaches enable CTI workflows that are more scalable and responsive, while also introducing new technical and operational challenges.

The integration of AI into CTI, therefore, represents not only a technological enhancement but also a shift in operational practice. Several factors contribute to this convergence, including:

- The exponential growth of security-relevant data from logs, threat feeds, and external sources.
- The increasing sophistication of Advanced Persistent Threats (APTs).
- The widespread adoption of CTI platforms, including open-source solutions such as MISP (Malware Information Sharing Platform and Threat Sharing) and OpenCTI, as well as commercial Threat Intelligence Platforms (TIPs).
- The maturation of AI techniques capable of operating under noisy, incomplete, and adversarial conditions.

A growing body of literature reflects these developments. Reference [16] examines CTI sharing practices and the challenges associated with automation, while [17] discusses limitations in existing sharing frameworks and explores the potential of Deep Learning (DL) to improve collaboration. The survey in [18] provides a comparative analysis of DL-based intrusion–detection approaches and datasets. Other studies focus on cyberattack detection in Internet of Things (IoT) environments using DL [19,20] and Federated Learning (FL) [21]. Reference [22] analyses the role of AI in automated cybersecurity response and decision-making, while [23] presents an AI-enhanced CTI processing pipeline and emphasises the importance of human–AI synergy. Reference [24] employs ML techniques and the UNSW-NB15 dataset [25] to train and evaluate a hybrid model combining Support Vector Machines (SVM), random forest, and Multilayer Perceptron (MLP). Meanwhile, Ref. [26] explores the use of agentic AI to select SOC functions automatically. Additionally, the European Union Agency for Cybersecurity (ENISA) provides ongoing assessments of emerging threats through its annual Threat Landscape reports [27].

Taken together, these studies indicate that AI-enhanced CTI is transitioning from experimental exploration to practical adoption, while raising important questions about evaluation, trust, and governance.

This study examines opportunities and challenges at the intersection of cybersecurity and AI by analysing CTI architectures, representative use cases, and CTI platform capabilities. It follows a structured qualitative review approach, focusing on representative architectures, AI techniques, and platforms rather than exhaustive coverage. In contrast to surveys centred on individual AI techniques or isolated CTI tasks, this work adopts a holistic, AI-centric perspective spanning the entire CTI lifecycle. This study is guided by the following research question: how is AI reshaping the architecture, capabilities, and operational workflows of CTI systems throughout the CTI lifecycle? The main contributions of this paper are as follows:

- The definition of a unified AI-enhanced CTI architectural model integrating data collection, enrichment, correlation, prioritisation, and operationalisation within a single conceptual framework.
- A systematic mapping of AI techniques to CTI functions, contrasting traditional and AI-driven workflows across multiple operational dimensions.
- A structured comparison of open-source and commercial CTI platforms, considering functional requirements, non-functional properties, and AI capabilities.
- An analysis of current limitations and future research directions, including autonomous CTI agents, federated intelligence, and human–AI collaborative systems.

Together, these contributions aim to provide a reference framework for researchers, practitioners, and decision-makers involved in the design, evaluation, and adoption of AI-enhanced CTI systems.

The remainder of the paper is as follows: Section 2 describes the architecture of a modern CTI system. Section 3 examines the role of AI in CTI and associated use cases. Section 4 presents a comparative qualitative analysis of CTI platforms with AI capabilities. Section 5 discusses challenges and limitations of AI-driven CTI. Section 6 outlines future research directions, and Section 7 concludes the paper.

2. Unified AI-Enhanced CTI Conceptual Architecture

2.1. Conceptual Overview and Design Rationale

This review adopts a structured qualitative approach, focusing on representative architectures, AI techniques, and CTI platforms rather than exhaustive enumeration, to synthesise how Artificial Intelligence supports the CTI lifecycle at a system level. Accordingly, this work should be interpreted as a structured qualitative synthesis rather than a systematic or algorithm-centric review, prioritising conceptual integration, architectural coherence, and functional abstraction over low-level implementation detail.

In this context, the paper's primary contribution is to unify AI-enabled CTI capabilities across the full intelligence lifecycle under a single architectural perspective. As such, the analysis prioritises architectural and system-level abstraction over algorithmic depth, emphasising how different classes of AI techniques are positioned, combined, and operationalised within CTI pipelines, rather than proposing or optimising specific models.

Building on this layered perspective, the proposed conceptual model integrates AI capabilities directly within each stage of the CTI pipeline, rather than treating AI as an external or auxiliary component. AI functions are therefore modelled as cross-cutting enablers embedded throughout the lifecycle, supporting data ingestion, enrichment, analysis, prioritisation, and operational decision support, while accommodating both automated and analyst-driven workflows.

Human expertise is incorporated across all architectural layers through Human-in-the-Loop (HITL) mechanisms, ensuring that automated intelligence generation remains interpretable, accountable, and operationally relevant. This design choice reflects current evidence that AI-driven CTI systems achieve greater effectiveness and trustworthiness when automation is combined with expert validation, contextual judgement, and continuous oversight.

2.2. Core Architectural Layers of AI-Enhanced CTI

In the proposed model, AI techniques support five tightly coupled architectural layers, each corresponding to a distinct stage of the CTI lifecycle:

- Data acquisition and ingestion, where NLP techniques and automated crawlers extract structured intelligence from heterogeneous sources, including OSINT, technical threat feeds, and internal telemetry.
- Normalisation and enrichment, where ML-based classifiers and entity-resolution models enhance raw indicators with contextual information. At the ingestion and enrichment layers, recent report-to-STIX extraction frameworks (e.g., AZERG [28]) demonstrate how Large Language Models (LLM)-based pipelines can operationalise the generation of structured CTI artefacts by extracting STIX entities and relationships from unstructured reports.
- Analysis and correlation, where graph-based learning and pattern recognition techniques infer relationships between indicators, campaigns, and threat actors.

- Threat scoring and prioritisation, where probabilistic models and anomaly-detection techniques assess relevance, urgency, and operational risk.
- Operationalisation and decision support, where AI-assisted analytics and human–AI interfaces support IR, reporting, and strategic decision-making.

2.3. Unified AI-Enhanced CTI Reference Architecture

Figure 1 illustrates the unified AI-enhanced CTI reference architecture, integrating AI capabilities across the full CTI lifecycle. AI techniques support data ingestion, enrichment, analysis and correlation, threat prioritisation, and operational decision support, while human–AI collaboration spans all architectural layers. The model emphasises a transition from static, indicator-centric CTI pipelines toward adaptive, intelligence-driven systems. This unified architecture serves as a methodological and architectural framework for the design, analysis, comparison, and practical implementation of AI-enhanced CTI systems.

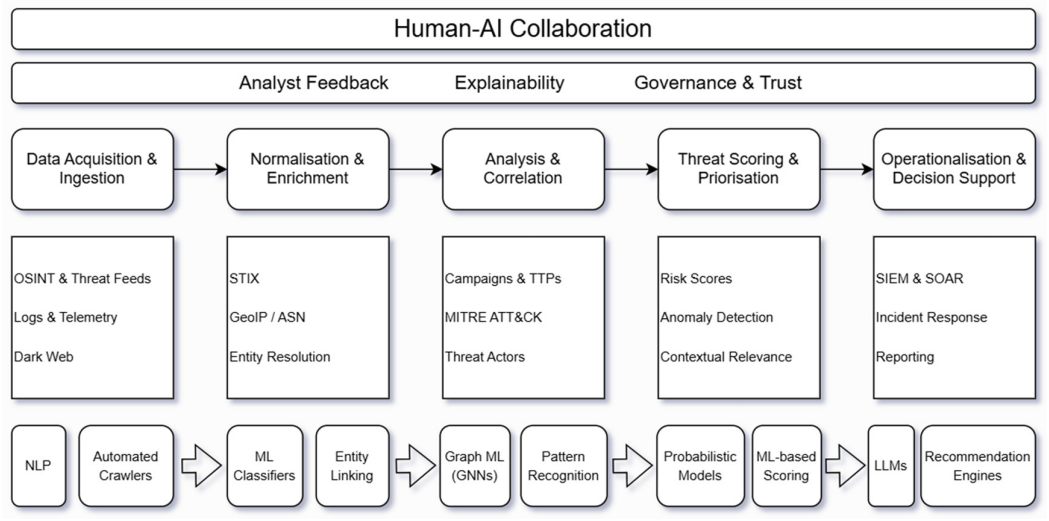

Figure 1. Unified AI-enhanced CTI conceptual architecture.

2.4. Functional Capabilities of CTI Systems

A CTI system is a specialised system designed to collect, process, analyse, enrich, correlate, and disseminate threat data to support the detection, response, and mitigation of cyber threats. It serves as a central coordination layer for managing intelligence on malicious activities, actors, tools, and tactics. CTI systems support security roles such as analysts, incident responders, and threat hunters by providing structured intelligence products and contextual insights, rather than raw indicators alone.

From an operational perspective, CTI systems typically provide the following benefits:

- Enhanced threat detection by incorporating up-to-date intelligence to identify emerging threats.
- Improved IR, by contextualising security events and supporting faster triage and investigation.
- Centralised threat management, enabling correlation across heterogeneous data sources.
- Risk-informed defence planning, through mapping adversarial TTPs and potential attack paths.
- Cross-team collaboration, facilitating shared situational awareness among security stakeholders.
- Compliance and reporting support, by structuring intelligence in line with regulatory and governance requirements.

To deliver these capabilities, CTI systems must support a set of core functional components:

- Threat data collection, aggregating information from OSINT, commercial feeds, ISACs (Information Sharing and Analysis Centres), and internal sources, and supporting structured exchange through standards such as STIX/TAXII (Structured Threat Information Expression/Trusted Automated Exchange of Indicator Information) [29,30].
- Threat intelligence processing, transforming raw data into actionable intelligence through normalisation, de-duplication, and enrichment of IoCs.
- Analysis and correlation, linking intelligence to threat actors, APT groups, campaigns, and TTPs using frameworks such as MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge).
- Threat visualisation, providing graphical representations and dashboards to support situational awareness and trend analysis.
- Operationalisation integration, enabling intelligence-driven action through integration with SIEM (Security Information and Event Management), SOAR (Security Orchestration, Automation, and Response), firewalls, and vulnerability management tools.
- Threat scoring and prioritisation, ranking threats according to severity, relevance, and potential impact.
- Collaboration and information sharing, supporting controlled dissemination using protocols such as the Traffic Light Protocol (TLP).

2.5. Reference Architecture and Deployment Model

Modern CTI architectures have evolved from rigid, linear processing pipelines to more dynamic, adaptive systems. Based on a synthesis of existing CTI platforms and recent AI-driven advances, this work adopts a layered reference architecture for AI-enhanced CTI systems, summarised in Table 1.

Table 1. Architecture of a CTI system with AI integration.

Layer	Key Elements
Data collection layer	Data ingestion (OSINT/Feeds/Logs/Sensors), NLP-based extraction, automated crawlers
Data normalisation and enrichment layer	STIX 2.1, GeoIP, ASN, DNS, entity resolution and AI-based enrichment
Analysis and correlation layer	MITRE ATT&CK, clustering, graph-based analysis (e.g., GNNs)
Threat scoring and prioritisation layer	Threat scores, anomaly detection, probabilistic and ML-based scoring
Integration and action layer	SIEM, SOAR, EDR/XDR, CTI platforms, LLM-assisted decision support

The data collection layer ingests structured and unstructured threat data from diverse sources, including structured feeds (e.g., STIX/TAXII, commercial intelligence feeds such as Anomali ThreatStream [31], Recorded Future [32]), unstructured sources such as blogs and reports, internal telemetry (e.g., firewall logs, IDS/IPS (Intrusion Detection System/Intrusion Prevention System) alerts, SIEM data), and community-driven platforms such as MISP and OTX (Open Threat Exchange). AI-based crawlers and NLP models are increasingly used at this stage to extract relevant IoCs and TTPs from unstructured content.

The data normalisation and enrichment layer converts collected data into a standard schema (typically STIX 2.1) and augments it with contextual attributes such as geolocation, ASN (Autonomous System Number), DNS (Domain Name System) information, historical sightings (e.g., through sources like VirusTotal [33]), and attribution hypotheses (linking IoCs to known threat actors or APT groups). AI techniques support entity recognition, disambiguation (e.g., distinguishing between benign and malicious domains) [34], and enrichment recommendation, particularly in large and heterogeneous datasets.

The analysis and correlation layer represents the analytical core of CTI systems. It supports TTP mapping using MITRE ATT&CK or similar ontologies [35], pattern recognition, clustering of related incidents, and graph-based analysis of relationships between IoCs, malware families, and threat actors. Graph Neural Networks (GNNs) and associated techniques have been applied to uncover latent relationships within large CTI knowledge graphs.

The threat scoring and prioritisation layer assists analysts in triaging intelligence by combining historical observations, contextual relevance, and anomaly detection to rank threats and highlight deviations from expected behaviour.

Finally, the integration and action layer operationalises CTI by connecting intelligence outputs with the broader security ecosystem, including SIEM, SOAR, EDR/XDR (Endpoint Detection and Response/Extended Detection and Response), and CTI management platforms, enabling both automated and analyst-driven response workflows.

CTI architectures may be centralised, decentralised, or distributed [36]. The architecture summarised in Table 1 reflects a centrally orchestrated CTI pipeline, in which intelligence processing is coordinated through a unified analytical core, while data collection remains distributed across multiple sources. This design supports end-to-end coordination among ingestion, analysis, prioritisation, and action, while preserving flexibility at the data-collection layer.

3. The Role of AI in CTI

The AI techniques discussed in this section are analysed in the context of the unified AI-enhanced CTI architecture introduced in Section 2.1, highlighting how different models support specific layers of the CTI pipeline.

Traditional CTI systems have primarily relied on predefined rule sets, manual analysis, and static threat feeds. While effective for specific tasks, these approaches face scalability and timeliness limitations in highly dynamic, adaptive threat environments. AI techniques address some of these limitations by enabling data-driven, adaptive, and scalable analysis, supporting both automation and analyst decision-making. The following subsections examine how different AI techniques contribute to specific CTI functions.

3.1. Machine Learning for Indicator Classification and Enrichment

Manual classification of domains, IP addresses (IPs), Uniform Resource Locators (URLs), and file hashes as benign or malicious are time-consuming and prone to inconsistency. In [37], the authors analyse the evolving nature of IoCs and their implications for predictive and defensive mechanisms. Supervised ML models can be trained on historical datasets (e.g., MISP, VirusTotal [32], PhishTank [38]) to classify previously unseen IoCs based on syntactic features, behavioural attributes, and contextual metadata. Commonly applied algorithms include Random Forests, XGBoost, SVM, and Logistic Regression.

DL approaches, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have been applied to feature learning from raw strings and traffic logs. These models reduce reliance on manual feature engineering but introduce challenges related to interpretability and generalisation. In [39], a CNN-based system was applied to anomaly detection and multilingual threat intelligence analysis across multiple data sources. In [40], a collaborative DL model combining an unsupervised LSTM with a supervised CNN was evaluated on the CTU-13 and IoT-23 datasets, achieving high detection rates with low false-positive rates.

3.2. Natural Language Processing for Threat Intelligence Extraction

A substantial portion of CTI originates from unstructured textual sources, including blogs, forums, dark web marketplaces, and APT reports. NLP techniques enable the auto-

matic extraction of entities (e.g., malware names, Common Vulnerabilities and Exposures (CVEs), domains) and relationships from such sources. Typical use cases include entity recognition, document summarisation for situational awareness, and sentiment or urgency analysis in underground communications.

Commonly used models include Bidirectional Encoder Representations from Transformers (BERT), Robustly Optimised BERT Pretraining Approach (RoBERTa), Generative Pre-trained Transformers (GPT), and custom Named Entity Recognition (NER) pipelines built on frameworks such as spaCy. In [41], NLP-based extraction and classification models achieved high accuracy under controlled experimental conditions.

Recent CTI extraction frameworks have demonstrated end-to-end pipelines for converting unstructured reports into structured threat intelligence. For example, AZERG proposes a task-decomposed LLM-based workflow to extract STIX entities and relationships, supporting semi-automated report generation with expert validation [28]. These approaches automate early CTI lifecycle stages while reinforcing the need for human oversight to ensure correctness and accountability.

3.3. Graph-Based Learning for Correlation and Attribution

Cyberattack campaigns often share infrastructure, tooling, or tactics that are difficult to identify through isolated indicators. Graph-based representations enable the modelling of complex relationships between threat actors, malware, infrastructure, and IoCs. CTI platforms such as OpenCTI already employ graph-based visualisation to support analyst-driven correlation.

Graph-based AI techniques, including GNNs, Knowledge Graph Embeddings (e.g., TransE (Translating Embeddings), ComplEx (Complex Embeddings)), and community detection algorithms (e.g., Louvain, Infomap), have been applied to tasks such as campaign clustering, attribution hypothesis generation, and discovery of latent threat relationships. These techniques support correlation at scale but remain sensitive to data quality and graph completeness.

3.4. Predictive Analytics and Threat Forecasting

AI techniques also support predictive CTI capabilities, enabling a shift from retrospective analysis toward anticipatory intelligence. By analysing historical campaigns, attacker behaviour, and temporal patterns, models can forecast likely future targets, techniques, or indicators. Use cases include predicting phishing campaigns based on domain registration patterns and forecasting trends in vulnerability exploitation using signals such as CVSS (Common Vulnerability Scoring System) scores and discussion volume.

Standard techniques include time-series models (e.g., AutoRegressive Integrated Moving Average (ARIMA), Prophet, LSTM), Reinforcement Learning (RL) for adversarial behaviour modelling, and meta-learning approaches for generalising across attack patterns. As highlighted in [42], forecasting zero-day exploitation remains a particularly challenging problem, underscoring the uncertainty inherent in predictive CTI.

3.5. AI-Assisted Decision Support and Response

The integration of AI with SOAR platforms has enabled automated triage, response orchestration, and analyst augmentation. LLMs can generate incident summaries, recommend response actions, and translate technical findings into formats suitable for decision-makers [43]. Recent surveys, such as [44], analyse the application of LLMs across intrusion detection, malware analysis, and phishing detection.

However, LLM deployment in CTI introduces new risks, including prompt injection, data poisoning, insecure output handling, and adversarial manipulation. Empirical evaluations in [45] demonstrate promising results for AI-assisted CTI extraction and report

generation, but also reveal variability across models. These findings motivate hybrid approaches that combine automated extraction with validation and ensemble strategies to approach human-level accuracy.

3.6. Operational Definition of Predictive CTI

In operational terms, predictive CTI refers to a CTI system's ability to generate time-aware, forward-looking intelligence outputs that anticipate the likelihood, evolution, or impact of cyber threats before they materialise. In this context, predictive CTI does not imply deterministic prediction of specific future attacks, but rather probabilistic, time-aware anticipation based on historical patterns, behavioural signals, and contextual indicators, subject to uncertainty and analyst validation. Such systems typically ingest temporally ordered data (e.g., incident reports, campaign timelines, telemetry, and external intelligence feeds) and perform tasks including threat occurrence forecasting, campaign evolution prediction, and risk-based prioritisation. Descriptive and correlational analyses are treated as enabling components within the CTI pipeline, but are not, by themselves, interpreted as predictive unless validated in a time-aware, forward-looking setting.

The outputs of predictive CTI extend beyond alerts to include ranked threat assessments, confidence-scored forecasts, and recommended pre-emptive actions that inform defensive planning and resource allocation. From an evaluation perspective, predictive CTI should be assessed using metrics that capture timeliness, lead time, prioritisation quality, and operational relevance, rather than relying solely on static accuracy.

Table 2 summarises several areas of work in CTI and how they can be approached differently, depending on whether AI is used. This structured comparison explicitly contrasts traditional CTI processes with AI-enhanced approaches across the full intelligence lifecycle.

Table 2. Traditional approach to CTI vs. AI-enhanced approach to CTI.

Area	Traditional CTI Approach	AI-Enhanced CTI Capability
IoC Classification	Manual review, analyst-driven tagging, static rule sets	Supervised ML and DL models supporting automated classification and anomaly detection
OSINT parsing	Manual reading, keyword matching, and regular expressions	NLP pipelines and LLM-based extraction for entity recognition and relation identification
Campaign correlation	Analyst-driven graphing and manual hypothesis building	Graph-based learning and community detection to support scalable correlation
Threat attribution	Expert-driven analysis based on TTP similarity and infrastructure reuse	Probabilistic correlation and graph-based learning to support attribution analysis
Threat forecasting	Retrospective analysis and trend extrapolation	Time-series modelling and behaviour-based forecasting to support probabilistic, anticipatory analysis
Alert prioritization	Static scoring rules and severity thresholds	Context-aware ML-based scoring supporting dynamic prioritisation
Response automation	Manual execution or rule-based playbooks	AI-assisted playbook recommendation and selective automation in SOAR platforms
Dark web monitoring	Manual crawling and ad hoc monitoring	Automated collection and content classification using AI-driven agents
Report generation	Analyst-written technical and executive summaries	LLM-assisted drafting of technical and executive intelligence reports
Decision support	Periodic static reports	Interactive decision-support tools and natural language interfaces

The integration of AI into CTI has moved far beyond academic research and is actively transforming operational environments. Table 3 summarises several concrete CTI use cases and how they can be approached differently, depending on whether AI is used. Across the use cases presented, AI can significantly enhance CTI capabilities by improving scalability, timeliness, and analytical depth. In many operational contexts, the use of AI supports a transition from predominantly reactive defence mechanisms toward more proactive, adaptive, and predictive cyber defence [46]. Beyond improving speed and scalability, AI models also uncover hidden patterns, adversarial strategies, and early warning signals that are extremely difficult to detect manually at scale.

Table 3. Concrete CTI use cases with and without AI.

Use Case	Without AI	With AI
Malicious URL Detection	Reliance on manually curated blocklists and rule-based filtering, which require frequent updates and may miss novel or obfuscated URLs.	ML models (e.g., Bidirectional Long Short-Term Memory (BiLSTM), CNNs) trained on large labelled datasets to support detection based on lexical, structural, and contextual features.
Extraction of IoCs from OSINT and threat reports	Manual review of reports and forum posts to identify IPs, domains, CVEs, and malware names, resulting in time-consuming and inconsistent extraction.	NLP models (e.g., BERT, GPT) and LLM-based pipelines to support automated extraction and structuring of threat intelligence from unstructured text.
Threat actor campaign attribution	Attribution based primarily on expert judgement and manual correlation of TTPs and infrastructure.	Graph-based learning techniques to support clustering of campaigns and generation of attribution hypotheses.
Threat forecasting	Retrospective analysis and heuristic trend extrapolation based on historical incidents.	Time series and behaviour-based models (e.g., LSTM, Prophet) to support forecasting of potential future threat activity.
Triage and automated response	Manual alert triage and analyst-driven execution of response actions.	AI-assisted risk scoring, contextual correlation, and selective automation through SOAR platforms.

4. CTI Platforms and Tools with AI Capabilities

4.1. Platform Evaluation Methodology

The comparison of CTI platforms in this study follows a criteria-based qualitative evaluation methodology designed to be transparent and reproducible. Rather than relying on performance metrics tied to specific deployments or proprietary datasets, platforms are assessed against a standard set of evaluation dimensions that reflect their functional scope, operational characteristics, and AI-related capabilities.

The evaluation criteria are grouped into three categories:

- Functional requirements, covering threat intelligence ingestion, organisation, enrichment, analysis, visualisation, integration, and IR support.
- Non-functional requirements, including scalability, usability, reliability, security, cost, and extensibility.
- AI-related capabilities, such as support for NLP, graph-based analysis, automated enrichment, predictive analytics, and integration of custom ML models.

Each platform is evaluated against these criteria using publicly available documentation, such as open-source feature sets or vendor information. This approach enables

consistent, repeatable comparisons while acknowledging that quantitative benchmarking would require controlled experimental environments beyond the scope of this work.

The CTI platforms analysed in this section can be interpreted as different implementations of the unified AI-enhanced CTI architecture introduced in Section 2.1, each emphasising specific layers or capabilities depending on design goals and target use cases.

Platforms were selected to represent a diverse set of widely referenced open-source and commercial CTI solutions. Inclusion criteria comprised documented support for CTI standards (e.g., STIX/TAXII), relevance to AI-enhanced intelligence workflows, and sufficient public technical documentation. Platforms lacking adequate documentation or relevance to AI-enabled CTI were excluded from the analysis.

4.2. Functional and Non-Functional Evaluation Criteria

When analysing CTI platforms, both functional and non-functional criteria are essential to assess alignment with organisational requirements, operational workflows, and technical constraints.

Functional Evaluation Criteria

The main functional criteria considered in this study include:

- Threat intelligence collection and ingestion: support for importing threat data from multiple sources, including open-source feeds, commercial feeds, ISACs, and internal reports; compatibility with STIX/TAXII; and integration with external sources such as MISP, VirusTotal [33], and public APIs (Application Programming Interface).
- Threat data organisation and management: normalisation and de-duplication of redundant data; support for IoCs (e.g., IPs, hashes, domains) and higher-level entities (e.g., actors, TTPs); and contextual linking of entities across campaigns and incidents.
- Threat intelligence analysis and enrichment: automated enrichment (e.g., geolocation, reputation scoring), correlation with frameworks such as MITRE ATT&CK, and historical analysis to identify recurring patterns.
- Visualisation and reporting: graph-based representations of relationships, dashboards for Key Performance Indicators (KPIs) and trends, and configurable reporting for stakeholders with varying technical expertise.
- Integration with other systems: connectivity with SIEM, SOAR, firewalls, endpoint protection platforms, and IR tools such as The Hive [47], supported by APIs and bidirectional data exchange.
- Collaboration and information sharing: support for intelligence sharing within trusted communities, implementation of the TLP, and Role-based Access Control (RBAC).
- IR and mitigation support: linkage between threat intelligence and active incidents, support for response workflows, and integration with automated or semi-automated playbooks.
- Interoperability and standards compliance: adherence to STIX/TAXII, API-based interoperability, and alignment with regulatory and industry frameworks such as GDPR and MITRE ATT&CK.

Non-functional Evaluation Criteria

The non-functional criteria considered include:

- Performance and scalability: capacity to handle large volumes of threat data, query latency, and scalability across distributed or cloud-based deployments.
- Usability: interface design, accessibility for non-technical users, learning curve, and availability of documentation and training resources.
- Reliability and availability: platform stability, high-availability options, and support for backup and disaster recovery.
- Security: Encryption for data at rest and in transit, RBAC, and audit logging.

- Cost and licensing: licensing models, total cost of ownership, and availability of open-source or hybrid options.
- Customisation and flexibility: support for custom workflows, dashboards, enrichment sources, and extensibility through APIs or scripting.
- Support and community: vendor support quality (e.g., SLAs (Service Level Agreement)) or, for open-source platforms, the strength and activity of the supporting community.
- Innovation and roadmap: update frequency, adoption of emerging techniques, and transparency regarding future development.

4.3. Comparative Analysis of Open-Source and Commercial CTI Platforms

A range of CTI platforms integrate AI to varying degrees in support of ingestion, enrichment, analysis, and response. Tables 4 and 5 present a structured comparison of seven representative platforms—MISP [48], OpenCTI [49], YETI (Your Everyday Threat Intelligence) [50], Recorded Future [32], Anomali ThreatStream [31], ThreatQuotient (ThreatQ) [51], and the integrated IBM X-Force [52], QRadar [53], and Watson [54] stack—based on functional and non-functional criteria.

Table 4. Comparison of functional requirements of key CTI platforms.

Functional Requirement	MISP	OpenCTI	YETI	Recorded Future	Anomali ThreatStream	ThreatQ	IBM X-Force + QRadar + Watson
Threat intelligence collection and ingestion	Manual input, APIs, STIX/TAXII, feeds, external enrichers	Native connectors, feeds, APIs, STIX2.1, TAXII	Manual input and basic REST API	Automated ingestion from open web, dark web, and technical sources	Aggregation from multiple feeds, STIX/TAXII, API support	Ingestion of structured and unstructured data via APIs, feeds, and manual inputs	Collection through IBM threat feeds and integration with QRadar
Threat data organisation and management	Tagging, taxonomy, correlation mechanisms	Graph-based data model with explicit entity relationships	Flat tagging and IoC labelling	Ontologies and knowledge-graph-based organisation	Structured data models and workspace-based organisation	Contextual threat library with entity linking	Structured CTI management across X-Force and QRadar components
Threat intelligence analysis and enrichment	External scripts and enrichers for scoring and contextualisation	Native and extensible enrichment modules, including ML and NLP integrations	Limited native analysis, typically combined with external tools	Integrated ML, NLP, and analytics for enrichment and prioritisation	ML-based enrichment, risk scoring, anomaly detection via enrichers	API-based enrichment and contextual analysis	NLP-driven analysis and enrichment using Watson services
Visualisation and reporting	Tabular views and basic graphs, limited dashboards	Graph, visualisation, timelines, dashboards, and filtering	Minimal User Interface (UI) and basic IoC views	Dashboards, graph exploration, and automated reporting	Dashboards, correlation views and feed monitoring	Customisable dashboards, threat timelines, and correlation views	Dashboards and visualisation through QRadar and Watson interfaces
Integration with other systems	API, STIX/TAXII, script-based third-party integration	Extensive APIs and connectors for SIEM, SOAR, and TIPs	REST API for basic integrations	Integration with SIEM, SOAR, ticketing, and vulnerability platforms	Integration with SIEM, SOAR, and enrichment services	Integration with SIEM, SOAR, and EDR platforms via APIs	Deep integration with IBM SIEM, SOAR, EDR, and external systems
Collaboration and information sharing	Community-driven sharing, tagging, TAXII servers	Multi-user collaboration with granular roles and organisational views	Single-user or lightweight collaboration	Collaboration primarily oriented toward intelligence consumption	Role-based sharing within workspaces and communities	Team collaboration through role-based workflows	Multi-user collaboration across SOC teams with workflow coordination
IR and mitigation support	Not native, relies on integrations with IR tools	Case management and observable tracking	Not designed for IR workflows	Contextual intelligence to support IR, not a full IR platform	Integration with SOAR and ticketing systems	Native support for IR workflows, case management, and evidence handling	Integrated IR and response through IBM SOAR and QRadar
Interoperability and standards compliance	Strong STIX 1.x/2.x and TAXII support, scripting extensibility, GDPR-awareness	Native STIX 2.1, TAXII, MITRE ATT&CK support; audit and schema enforcement	Partial STIX support, limited compliance features	Support for STIX and MITRE ATT&CK, alignment with selected compliance frameworks	STIX/TAXII support and alignment with regulatory requirements	STIX and MITRE ATT&CK alignment, API-based interoperability	Native STIX/TAXII support and alignment with NIST, ISO, and GDPR frameworks

Table 5. Comparison of non-functional requirements of key CTI platforms.

Non-Functional Requirement	MISP	OpenCTI	YETI	Recorded Future	Anomali ThreatStream	ThreatQ	IBM X-Force + QRadar + Watson
Performance and scalability	Suitable for small to mid-sized deployments, scalability depends on deployment and tuning	Modular and scalable architecture using workers and connectors, suitable for large data volumes	Limited scalability, primarily suited for small-scale or laboratory use	Cloud-based infrastructure designed to handle large volumes of threat data	Scalable for enterprise use, including multi-tenant deployments	Enterprise-oriented scalability, performance depends on architecture and configuration	Designed for large-scale enterprise deployments leveraging IBM infrastructure
Usability	Moderate learning curve, functional web UI	Intuitive UI, increased complexity with large graph datasets	Basic UI with limited interactivity	Dashboards and interfaces designed for operational consumption	Professional interface with a moderate learning curve	Analyst-oriented interface, effective use may require training	Guided workflows and interfaces integrated with Watson services
Reliability and availability	Stable operation dependent on local deployment; high availability possible with configuration	Generally stable with appropriate configuration and redundancy	Limited high-availability features	High availability through cloud-based deployment models	High availability through cloud or hybrid deployment options	High-availability options available for critical deployments	Enterprise-grade availability supported by IBM service guarantees
Security	Depends on configuration, RBAC	RBAC, authentication integration, and audit logging	Minimal built-in security features	Encryption and access controls managed within a secure cloud environment	Security features aligned with enterprise standards (e.g., RBAC, secure APIs)	Strong access control and integration with enterprise security frameworks	Enterprise-grade security aligned with NIST (National Institute of Standards and Technology), ISO and GDPR requirements
Cost and licensing	Open-source, operational costs may apply	Open-source, operational costs may apply	Open-source, operational costs may apply	Commercial licensing	Commercial licensing	Commercial licensing	Commercial licensing
Customisation and flexibility	High flexibility through scripting and modules	High flexibility through connectors, APIs, and extensible data model	Limited flexibility by design	Limited customisation of core platform, configurable outputs and reports	Moderate flexibility through enrichers, templates, and APIs	High flexibility via APIs, workflow customisation, and dashboard configuration	Moderate flexibility within IBM frameworks and predefined models
Support and community	Large and active open-source community with extensive documentation	Active open-source community	Smaller community and limited ongoing development	Vendor support	Vendor support	Vendor support	Global IBM support
Innovation and roadmap	Community-driven development with incremental updates	Active development with frequent feature additions	Limited development activity	Ongoing development focused on analytics and automation features	Ongoing development focused on enrichment and orchestration	Ongoing development focused on automation and integration	Continued innovation through IBM Research

4.4. Discussion and Platform Selection Considerations

When evaluating CTI platforms against functional requirements, the following analysis summarises how each solution supports critical areas, including threat intelligence collection, data organisation, analysis, and IR.

- Threat intelligence collection and ingestion: Recorded Future and Anomali ThreatStream provide broad coverage of external sources and support automated ingestion workflows. Recorded Future includes collection from a wide range of sources, including the dark web, while Anomali ThreatStream focuses on aggregating and orchestrating multiple feeds with enrichment support. These characteristics may be relevant for organisations requiring high-volume, multi-source ingestion.
- Threat data organisation and management: OpenCTI supports a graph-based data model and extensible entity relationships, which can facilitate the structuring and management of large volumes of CTI. Recorded Future also uses a knowledge graph-based organisation to support navigation across entities and their relationships.
- Threat intelligence analysis and enrichment: Recorded Future provides integrated analytics capabilities, including ML and NLP components, to support enrichment and

prioritisation tasks. OpenCTI supports analysis through native features and extensibility, enabling integration with external ML tools and anomaly detection modules.

- Visualisation and reporting: Recorded Future provides dashboards and visual exploration features aimed at communicating intelligence to different stakeholders. Anomali ThreatStream includes dashboards and reporting functions that support monitoring and feed management.
- Integration with other systems: OpenCTI provides extensive API and connector-based integration options, supporting interoperability with SIEM and SOAR ecosystems. The IBM X-Force Threat Intelligence + QRadar + Watson stack provides tight integration within IBM environments, which may be relevant for organisations already operating IBM's SIEM and response tooling.
- Collaboration and information sharing: MISP and OpenCTI both support multi-user collaboration and intelligence sharing. MISP is commonly used in community-driven sharing environments and provides robust support for exchanging IoCs via standards-based mechanisms (e.g., STIX/TAXII). OpenCTI supports role-based access and collaborative workflows across multiple users and organisational contexts.
- IR and mitigation support: The IBM X-Force Threat Intelligence + QRadar + Watson stack integrates CTI with SIEM and SOAR workflows to support triage and response activities. ThreatQ includes IR-oriented workflows (e.g., case management and analyst tasking) to support operational incident handling.

When evaluating platforms against non-functional requirements, the following analysis highlights relevant considerations such as performance, scalability, usability, and operational constraints.

- Performance and scalability: The IBM X-Force Threat Intelligence + QRadar + Watson stack is designed for large-scale enterprise deployments, leveraging QRadar's infrastructure. OpenCTI supports scalability through its modular architecture (e.g., workers and connectors), allowing organisations to expand ingestion and processing capacity as needed.
- Usability: Recorded Future provides dashboards and reporting interfaces aimed at supporting operational consumption of intelligence. ThreatQ also offers analyst-oriented workflows, although effective use may depend on training and organisational maturity.
- Reliability and availability: IBM X-Force Threat Intelligence + QRadar + Watson typically leverages enterprise infrastructure and service guarantees (depending on deployment). Recorded Future also offers high-availability delivery models through cloud-based operation.
- Security: The IBM stack and Anomali ThreatStream include security features aligned with enterprise deployments, such as access control and secure integration options.
- Cost and licensing: Open-source platforms such as MISP and OpenCTI reduce licensing costs but may require additional investment in deployment, integration, and maintenance. Commercial platforms (e.g., Recorded Future, Anomali ThreatStream, IBM X-Force) follow subscription or enterprise licensing models that vary with scale, data volume, and service levels.
- Customisation and flexibility: OpenCTI supports extensibility through connectors and APIs, enabling customised workflows and integrations. ThreatQ also supports workflow tailoring and external integrations through open APIs.
- Support and community: MISP and OpenCTI benefit from active open-source communities and extensive shared resources. Commercial platforms typically provide structured vendor support depending on contractual arrangements.

Finally, Table 6 provides a structured comparison of the same platforms—MISP [48], OpenCTI [49], YETI [50], Recorded Future [32], Anomali ThreatStream [31], ThreatQ [51], and IBM X-Force [52] + QRadar [53] + Watson [54]—with a specific focus on AI-related capabilities (e.g., NLP support, graph-based analysis, STIX 2.1 compliance, and integration of custom ML models). This table complements Tables 4 and 5 by isolating AI-related functionality under the same evaluation framework, while acknowledging that the degree of AI integration and transparency varies substantially across platforms.

Table 6. Comparison of AI features, interoperability, and typical use cases of key CTI platforms.

Platform	Type	AI integration	NLP Support	Graph Analysis	STIX2.1	Custom ML Integration	AI Usage	Representative Use Cases
MISP	Open-source	External	External	No	Partial	External	Typically extended using external enrichment scripts or integrations for scoring, NLP-based report parsing, or IoC analysis	Centralised CTI feed management, IoC sharing, and community-driven correlation
OpenCTI	Open-source	Native and extensible	Yes	Yes	Yes	Native and external	Integration of NLP tools, anomaly detection modules, and graph-based ML (e.g., GNNs) for correlation and attribution support	Strategic and operational CTI, knowledge graph construction, campaign correlation
YETI	Open-source	No	No	No	Partial	Minimal	Commonly used alongside external analytics pipelines rather than as a native AI-enabled platform	Lightweight IoC management, lab environments, front-end for enriched CTI
Recorded Future	Commercial	Proprietary	Advanced	Yes	Partial	Proprietary	NLP-driven extraction, ML-based scoring, trend analysis, and report generation using proprietary models	Threat monitoring, vulnerability prioritisation, geopolitical and strategic CTI
Anomali ThreatStream	Commercial	Moderate (via enrichers)	Basic	Partial	Yes	Via enrichers	ML-based filtering, enrichment, and risk scoring primarily through external or integrated enrichers	Threat feed orchestration, IoC enrichment, anomaly detection
ThreatQ	Commercial	Available via integrations	Limited	Yes	Yes	Via APIs	Integration with ML-based enrichment and analytics tools through open APIs	Threat analysis, workflow-driven CTI, SOC and IR support
IBM X-Force Threat Intelligence + QRadar + Watson	Commercial	High (Watson services)	Advanced	Partial	Yes	Cognitive Services	NLP-based analysis of threat reports, automated summarisation, and analytics integrated with SIEM and SOAR workflows	Enterprise SOC operations, AI-assisted triage, decision support, and response orchestration

Among the open-source platforms, OpenCTI provides a comparatively broad set of native features, including graph-based data modelling, STIX 2.1 support, and integration with NLP tools and ML-based enrichment modules. Its modular architecture, based on connectors and workers, supports extensibility and scalability, which may be relevant for strategic and operational CTI use cases in complex environments.

MISP, by contrast, focuses primarily on intelligence sharing, storage, and correlation of IoCs. While it does not include built-in AI capabilities, it is commonly extended through external scripts and integrations to support enrichment, scoring, or basic NLP-driven processing. Owing to its strong community support and widespread adoption, MISP remains a common choice for collaborative threat intelligence sharing across organisations and research settings.

YETI offers a lightweight alternative, well-suited to small-scale deployments or laboratory environments. It is often used as an interface or front-end for externally enriched intelligence rather than as a standalone analytics platform. Researchers frequently rely on

open-source platforms—particularly MISP [6,55–64] and OpenCTI [65–68], due to their accessibility, extensibility, and active user communities.

On the commercial side, Recorded Future integrates proprietary analytics capabilities, including NLP, ML-based scoring, and trend analysis, to synthesise threat intelligence from a wide range of sources such as the surface web, dark web, and technical feeds. Its use of knowledge-graph-based representations supports tasks such as prioritisation, contextual analysis, and report generation, which may be relevant for strategic and vulnerability-focused CTI.

The IBM integrated solution combining X-Force Threat Intelligence, QRadar, and Watson incorporates cognitive services and NLP-driven analysis within an enterprise SIEM and SOAR ecosystem. This integration supports use cases related to large-scale SOC operations, automated triage, and decision support.

Anomali ThreatStream provides AI-assisted functionality primarily through external enrichers and ML-based filtering mechanisms, with a focus on threat feed aggregation, classification, and orchestration. Its graph and NLP capabilities are more limited than those of platforms with native graph-based CTI models.

ThreatQ emphasises contextualisation and workflow customisation, integrating with ML-based enrichment and analytics tools via open APIs to support threat analysis and incident response workflows.

Overall, Table 6 illustrates a spectrum of CTI platform maturity. Open-source solutions offer flexibility and cost advantages but often require additional technical effort to extend AI-driven functionality. Commercial platforms typically provide more integrated analytics and automation capabilities, at the cost of reduced transparency and vendor dependency.

Given the diversity of available CTI platforms, selecting an appropriate solution requires balancing functional and non-functional requirements with AI-related capabilities. Considerations such as alignment with organisational use cases (e.g., IoC sharing, IR support), ease of adoption, integration with existing infrastructure, and the reliability of vendor or community support should be taken into account. Ultimately, the suitability of a CTI platform is shaped by the organisation's size, technical capacity, operational context, and intelligence objectives.

5. Technical and Operational Limitations of AI-Driven CTI

The limitations discussed in this section can be mapped directly to specific layers of the unified AI-enhanced CTI architecture, underscoring that challenges emerge at data, model, system, and human–AI interaction levels.

Despite the advances enabled by AI in CTI, its integration into operational environments introduces a range of technical, operational, and human-centric limitations. These constraints are not merely implementation issues, but reflect fundamental challenges associated with the nature of threat intelligence data, adversarial behaviour, and the complexity of socio-technical security systems. A rigorous understanding of these limitations is therefore essential to avoid overestimating the maturity, reliability, and autonomy of AI-driven CTI solutions.

To provide a structured and critical analysis, the limitations of AI-enhanced CTI are examined across multiple dimensions: data-related constraints; model-level and algorithmic limitations; vulnerability to adversarial manipulation; system-level and integration constraints; and human–AI interaction and decision-making risks, followed by ethical, regulatory, and governance considerations, as well as practical safeguards for LLM-in-the-loop CTI systems.

5.1. Data-Centric Limitations

AI models deployed in CTI are strongly dependent on the quality, representativeness, and availability of threat intelligence data. In practice, CTI datasets exhibit several structural deficiencies that directly affect the robustness and generalisability of AI-driven analysis.

First, data scarcity and class imbalance are pervasive, particularly in supervised learning settings, where benign indicators significantly outnumber malicious ones. This imbalance can bias classifiers, inflate false-negative rates for rare or emerging threats, and reduce sensitivity to low-frequency attack patterns. Second, CTI data often lacks reliable ground truth, especially for higher-level constructs such as threat actor attribution or campaign identification. As a result, labels are frequently noisy, incomplete, or weakly supervised, introducing uncertainty into both training and evaluation processes.

Beyond labelling challenges, CTI data is characterised by heterogeneous source reliability and systematic bias. Intelligence feeds vary widely in coverage, timeliness, and trustworthiness, while shared or community-driven datasets may reflect reporting bias, regional focus, or uneven visibility across threat actors and sectors. In addition, the temporal validity of CTI artefacts is often limited: indicators, infrastructure, and TTPs can rapidly lose relevance as adversaries adapt their behaviour.

CTI data is therefore highly dynamic. Concept drift—driven by evolving TTPs, malware polymorphism, and infrastructure rotation—can degrade model performance over time if continuous retraining and time-aware validation are not applied. Furthermore, a substantial portion of high-quality CTI data remains proprietary or classified, limiting the availability of open datasets for training, evaluation, and benchmarking. Consequently, many AI models are trained on incomplete or biased representations of the threat landscape, constraining their operational robustness and real-world applicability.

In operational CTI environments, these data quality challenges are typically addressed through architectural and workflow-level mechanisms rather than static dataset curation alone. Common mitigation strategies include source confidence scoring, temporal decay and ageing functions for indicators, cross-source corroboration of intelligence, and analyst-in-the-loop validation to contextualise, confirm, or override automated outputs. These approaches reflect the reality that CTI data quality must be managed continuously as part of the intelligence pipeline, rather than assumed at model training time.

5.2. Model-Level and Algorithmic Limitations

At the model level, many AI techniques applied to CTI—particularly DL architectures such as LSTM networks and GNNs—tend to exhibit limited explainability and interpretability. These models often operate as black boxes, making it difficult for analysts to understand why an IoC is classified as malicious, how relationships between entities are inferred, or how attribution hypotheses are formed. This opacity directly undermines analyst trust and complicates deployment in environments that require accountability, auditability, and regulatory compliance.

In graph-based CTI systems, inferred correlations do not necessarily imply causation. Relationships between IoCs, malware families, and threat actors are typically probabilistic and similarity-based rather than causally grounded. As CTI graphs scale in size and complexity, error propagation becomes a significant concern, as local inaccuracies or noisy edges may be amplified through graph embeddings, leading to misleading confidence in attribution or campaign linkage. While Explainable Artificial Intelligence (XAI) techniques such as LIME (Local Interpretable Model-agnostic Explanations) [69] and SHAP (SHapley Additive exPlanations) [70] provide partial insight into model behaviour, their applicability to large-scale, continuously evolving CTI pipelines remains limited, and their outputs are often difficult for non-specialist analysts to interpret.

5.3. Vulnerability to Adversarial Manipulation

AI-enhanced CTI systems are themselves high-value targets for adversarial manipulation, as attackers may actively seek to poison data sources, evade detection and classification models, or exploit automated intelligence pipelines. As a result, AI-driven CTI pipelines introduce new attack surfaces that adversaries can intentionally exploit.

Adversarial threats include data poisoning, where misleading or fabricated indicators are injected into shared intelligence feeds to distort downstream analysis, and evasion attacks, where inputs are deliberately crafted to bypass detection mechanisms (e.g., adversarial domain names, polymorphic malware, or obfuscated command-and-control infrastructure). In NLP-based CTI extraction, adversaries may also manipulate language in threat reports, forums, or underground communications to conceal malicious intent while remaining syntactically and semantically plausible.

This adversarial dynamic effectively transforms AI-driven CTI into a continuous arms race [71]. Models trained on historical data may degrade or fail when confronted with adaptive adversarial behaviour that exploits model assumptions, feature representations, or automation dependencies. Although robust training techniques such as adversarial learning and data augmentation exist, they are computationally expensive, require continuous updating, and remain difficult to operationalise at scale within production CTI environments.

5.4. System-Level and Integration Constraints

Operational CTI environments are typically heterogeneous, combining modern AI-enabled platforms with legacy SIEM, EDR, and IR systems. Integration challenges frequently arise from data format inconsistencies, limited API support in legacy tools, and strict regulatory or compliance requirements in sectors such as finance, healthcare, and defence.

In addition, AI-driven CTI systems often impose substantial computational and infrastructural demands. Real-time or near-real-time intelligence processing may conflict with the latency introduced by complex ML pipelines, while resource-intensive models may be impractical for smaller SOCs. As a result, adopting AI-enhanced CTI requires careful consideration of operational constraints, costs, and long-term maintainability.

5.5. Human–AI Interaction and Decision-Making Risks

AI in CTI is intended to augment, rather than replace, human analysts [72]. However, insufficient attention to human–AI interaction can introduce new risks. Opaque threat scores and automated recommendations may foster automation bias if analysts over-trust AI outputs without adequate scrutiny. Conversely, excessive false positives generated by immature or poorly calibrated models can contribute to analyst fatigue and erode confidence in AI-assisted workflows.

Effective CTI systems should therefore incorporate HITL mechanisms that enable analysts to review, contextualise, and override AI-driven outputs [73]. In operational practice, human–AI collaboration in CTI is typically realised through interaction patterns such as analyst-in-the-loop validation, feedback-driven model refinement, confidence-aware recommendations, and selective automation across different stages of the CTI workflow.

Evaluation of human–AI collaboration in this context extends beyond isolated model-centric accuracy metrics and instead focuses on dimensions such as analyst trust, system usability, decision quality, workload reduction, and the mitigation of automation bias and error propagation. Reciprocal learning approaches, such as the Reciprocal Human–Machine Learning (RHML) model [74], illustrate how iterative interaction between human expertise and ML can improve both analytical performance and analyst understanding.

Without such mechanisms, AI-enhanced CTI risks either underutilisation—where analysts disregard automated outputs—or inappropriate reliance, where AI recommendations are accepted uncritically, both of which can undermine operational effectiveness and decision-making quality in real-world settings.

5.6. Ethical, Regulatory, and Governance Challenges

AI-driven CTI also raises ethical and regulatory concerns. The ingestion of OSINT and dark web content may inadvertently collect Personally Identifiable Information (PII) or sensitive geopolitical data, creating privacy and compliance risks. Automated threat scoring and attribution, particularly in public-sector or multinational contexts, may have legal and ethical implications when decisions are based on opaque or biased models.

Emerging regulatory frameworks, such as the EU AI Act [75], are expected to impose additional requirements on the deployment of high-risk AI systems in cybersecurity. Consequently, governance mechanisms, transparency obligations, and accountability frameworks should be considered integral components of AI-enhanced CTI architectures rather than secondary considerations.

5.7. Practical Safeguards for LLM-in-the-Loop CTI

To move beyond risk acknowledgement toward operational deployment, LLM-assisted CTI systems should incorporate safeguards at multiple levels. First, generated outputs should be evidence-grounded using retrieval-augmented mechanisms and explicit citations of supporting sources to reduce intelligence driven by hallucination. Second, when producing structured CTI artefacts, systems should enforce constrained outputs and apply schema validation (e.g., STIX compliance checks) before ingestion into downstream workflows.

Third, decision-making processes should be uncertainty-aware, incorporating calibrated thresholds, abstention policies, and analyst escalation for high-impact actions. Fourth, pipelines should include adversarial hardening measures such as prompt-injection detection, input sanitisation, and poisoning-resistant data governance for external feeds. Finally, deployment should be supported by HITL verification, audit logging, model and version tracking, continuous monitoring, and periodic red-teaming to ensure accountability and safe evolution over time.

The limitations of AI-driven CTI are therefore multifaceted and closely intertwined with data quality, model design, adversarial dynamics, system integration, and human decision-making. While AI provides substantial benefits in scalability and analytical capability, these constraints underscore the need for cautious deployment, continuous validation, and hybrid intelligence models that combine automation with human expertise.

Table 7 summarises the key technical, operational, and governance challenges discussed in this section, together with representative mitigation strategies.

Table 7. Key technical, operational, and governance challenges in AI-driven CTI.

Challenge	Impact on CTI	Representative Mitigation Strategies
Data scarcity and noise	Reduced model accuracy, bias towards frequent classes, limited detection of emerging threats	Use of hybrid datasets, active learning, semi-supervised approaches, continuous data curation
Label uncertainty and weak ground truth	Unreliable attribution and campaign modelling	Analyst validation, probabilistic labelling, confidence-aware outputs
Concept drift	Performance degradation over time due to evolving TTPs	Continuous monitoring, periodic retraining, drift detection mechanisms

Table 7. *Cont.*

Challenge	Impact on CTI	Representative Mitigation Strategies
Lack of explainability	Reduced analyst trust and limited auditability	Use of XAI techniques, interpretable models where feasible, and explanatory metadata
Error propagation in graph-based models	Amplified inaccuracies in correlation and attribution	Graph validation checks, uncertainty propagation, conservative confidence thresholds
Adversarial manipulation	Evasion of detection and poisoning of intelligence pipelines	Adversarial training, input sanitisation, feed trust management, anomaly detection
Infrastructure and integration constraints	Deployment friction and operational delays	Modular, API-first architectures, staged integration, compatibility testing
Computational and resource overhead	Limited feasibility for smaller SOCs	Model optimisation, tiered deployment, selective automation
Human–AI collaboration gaps	Automation bias, analyst fatigue, under- or over-reliance on AI outputs	HITL workflows, calibrated alerting, analyst feedback loops
Ethical and regulatory risks	Compliance violations and reputational damage	Data governance policies, transparency measures, regulatory alignment (e.g., GDPR, EU AI Act)
LLM-specific risks (hallucination, prompt injection)	Generation of unreliable or misleading intelligence	Retrieval-augmented generation, constrained outputs, schema validation, red-teaming

6. Opportunities and Future Directions in AI-Enhanced CTI

As AI technologies continue to evolve, CTI is gradually shifting from predominantly reactive, analyst-driven processes toward more automated, anticipatory, and augmented workflows. Rather than representing a complete paradigm replacement, these developments indicate a progressive reconfiguration of how CTI is produced, evaluated, and operationalised. The most relevant opportunities for AI-enhanced CTI are outlined below.

- Predictive threat intelligence: One of the most actively explored research directions is the transition from retrospective analysis toward predictive CTI [76–78]. By leveraging historical threat data, behavioural patterns, and contextual signals (e.g., geopolitical developments or vulnerability disclosures), AI models can support forecasting tasks such as the potential emergence of attack campaigns, likely attack vectors or TTPs, exploitation trends associated with newly disclosed vulnerabilities, and sectoral or geographical targeting patterns. For example, combining domain registration data, underground forum activity, and CVE information may support early-warning assessments for phishing or ransomware campaigns. From an evaluation perspective, predictive CTI systems should be assessed using metrics that capture timeliness, prioritisation quality, and operational relevance, rather than relying solely on static classification accuracy. Time-aware forecasting measures, ranking-based metrics, and lead-time analysis are particularly relevant for early-warning intelligence. However, the limited availability of shared, longitudinal CTI benchmarks remains a significant challenge, underscoring the need for context-aware, organisation-specific evaluation approaches.
- Autonomous and semi-autonomous threat intelligence agents: Emerging research explores the use of autonomous or semi-autonomous CTI agents that continuously monitor open sources, underground forums, and dark web marketplaces. These agents may support tasks such as information extraction, classification, contextualisation, and enrichment using LLMs, and can trigger downstream workflows under predefined constraints. In practice, such agents are more likely to operate under human supervision, functioning as continuously running analytical assistants rather than fully autonomous decision-makers.
- Multidomain intelligence fusion (CTI, SIGINT, HUMINT, and CEMA): Future CTI systems are increasingly expected to operate within multidomain intelligence en-

vironments. AI can support the fusion of CTI with Signals Intelligence (SIGINT), Human Intelligence (HUMINT), and Cyber Electromagnetic Activities (CEMA). For example, intercepted communication patterns may be correlated with CTI indicators (e.g., Command and Control (C2) infrastructure). In contrast, textual analysis of adversary communications or social media activity can complement traditional CTI sources. In defence and critical infrastructure contexts, this fusion may extend to cyber–physical environments. Recent work in cyber–physical security highlights the role of digital twins and edge intelligence as predictive and decision-support mechanisms, enabling continuous monitoring and adaptive response [79,80]. Integrating CTI with such cyber–physical intelligence layers represents a promising research direction, particularly for healthcare, industrial systems, and critical infrastructure.

- Sector-specific and personalised CTI: AI enables the tailoring of CTI outputs to sector-specific threat models and operational contexts. In healthcare, emphasis is placed on medical IoT vulnerabilities and ransomware targeting hospitals; in finance, on phishing, credential abuse, and fraud-related campaigns [81]; and in industrial and ICS (Industrial Control Systems) environments, on Operational Technology (OT) threats, supply chain compromise, and hardware-level attacks. AI-driven personalisation can support adaptive dashboards and prioritisation mechanisms that focus analyst attention on threats most relevant to a given sector or organisational context. Prior studies suggest that interactive, context-aware dashboards may reduce response times by improving situational awareness and the interpretability of analytical outputs [82].
- Collaborative AI and federated threat learning: Privacy-preserving collaboration represents a key opportunity for large-scale CTI sharing. Techniques such as FL, differential privacy, and secure multiparty computation enable organisations to learn from shared threat patterns without exposing raw data. Community-based AI agents trained on anonymised intelligence can support collective learning while preserving data sovereignty and confidentiality. These approaches offer a pathway toward scalable, cross-organisational CTI collaboration under regulatory and privacy constraints [83].
- Evaluation, benchmarks, and operational validation: Recent benchmarking efforts provide initial reference points for evaluating AI-driven CTI under controlled conditions. CTIBench focuses on assessing LLMs on CTI-specific extraction and reasoning tasks [84], while CTIArena extends evaluation to heterogeneous, multi-source CTI scenarios requiring knowledge-augmented reasoning [85]. Together, these benchmarks illustrate both the potential and current limitations of LLMs in CTI contexts. Nonetheless, comprehensive and longitudinal benchmarks remain an open research need, particularly to evaluate end-to-end CTI pipelines (extraction, correlation, prioritisation, operationalisation) and to measure robustness, timeliness, and analyst-facing utility in realistic operational settings. While this study does not include experimental benchmarking, future work could complement architectural analyses with reproducible, task-based evaluations of specific AI capabilities (e.g., IR summarisation) across open models and CTI benchmarks.
- Human–AI collaboration and LLM adoption in SOCs: Recent research increasingly emphasises human–AI collaboration as a critical factor in the effective deployment of AI-driven CTI and cybersecurity systems. Conceptual models of trusted autonomy highlight the importance of calibrated automation levels, explicit human oversight, and accountability mechanisms within Security Operations Centres (SOCs). Empirical studies of LLM use in operational SOC environments indicate that these models are predominantly used as decision-support and sensemaking tools rather than as fully autonomous agents, reinforcing the relevance of HTIL architectures for CTI and IR workflows [86,87]. In parallel, recent surveys and systematic reviews have

synthesised the expanding literature on LLMs in cybersecurity, covering applications from CTI analysis and IR to vulnerability assessment and security automation. These studies consistently highlight challenges in reliability, evaluation, data quality, and adversarial robustness, underscoring the need for trustworthy, human-centred, and rigorously evaluated AI architectures as LLMs become embedded in operational CTI environments [88–90].

Table 8 summarises key strategic pillars shaping future research and development in AI-enhanced CTI.

Table 8. Strategic future pillars of AI in CTI.

Pillar	Description
Predictive intelligence	Forward-looking analysis supporting early warning, threat forecasting, and risk-based prioritisation
Autonomous agents	Semi-autonomous agents for continuous collection, enrichment, correlation, and triage under human oversight
Multidomain fusion	Integration of CTI with SIGINT, HUMINT, and cyber–physical intelligence to support cross-domain situational awareness
Sector-specific CTI	Context-aware models and adaptive dashboards tailored to industry-specific threat landscapes
Federated threat learning	Collaborative intelligence enabled by privacy-preserving AI techniques across organisational boundaries
Evaluation and benchmarking	Task-based and longitudinal evaluation of AI-driven CTI pipelines, focusing on robustness, timeliness, and operational utility
Human–AI collaboration	HITL and human-on-the-loop architectures supporting accountable, trustworthy, and decision-support-oriented CTI workflows

Overall, AI is increasingly shaping the evolution of CTI architectures and workflows. Rather than replacing human expertise, AI-driven approaches are most effective when deployed as part of hybrid intelligence systems that combine automation, analyst judgement, and continuous evaluation. Future progress in AI-enhanced CTI will therefore depend not only on advances in modelling, but also on robust evaluation practices, governance mechanisms, and human-centred system design.

7. Conclusions

This work synthesises AI techniques, CTI architectures, platform capabilities, and associated limitations into a unified conceptual model that clarifies how AI is reshaping the CTI lifecycle. By explicitly linking architectural layers, analytical techniques, and human decision-making, the proposed framework provides a system-level reference for the design, evaluation, and evolution of AI-enhanced CTI systems.

As with any state-of-the-art review, the selection of analysed works may introduce bias. Nevertheless, the focus on representative and widely-cited approaches aims to balance breadth of coverage with analytical depth, while supporting a structured and coherent synthesis of current research and practice.

The integration of AI into CTI should not be understood as a simple functional enhancement, but rather as a structural evolution in how threat intelligence is collected, processed, and operationalised. As cyber threats become more complex, distributed, and adaptive, the limitations of predominantly manual or rule-based CTI workflows become increasingly apparent. AI-driven techniques address scale, velocity, and analytical complexity, while simultaneously introducing new challenges related to trust, robustness, and

governance. In this context, the term "redefining" refers to a conceptual and architectural shift in how CTI systems are designed, operationalised, and governed, rather than to empirical performance gains demonstrated through new experimental benchmarks.

This study contributes an integrative architectural and conceptual framework that consolidates previously fragmented research on AI-enhanced CTI. Although no new detection algorithms are proposed, the contribution lies in clarifying system-level design choices, supporting informed platform evaluation, and identifying open research challenges across the CTI lifecycle. The comparative platform analysis is intentionally qualitative and criteria-driven; quantitative benchmarking is beyond the scope of this work.

Across use cases ranging from automated IoC classification and NLP-driven OSINT analysis to graph-based correlation and predictive threat assessment, AI techniques increasingly support a transition toward more anticipatory and intelligence-informed security operations. Both open-source and commercial platforms demonstrate diverse approaches to integrating AI capabilities, reflecting different design priorities, operational contexts, and maturity levels.

At the same time, these developments underscore the importance of addressing data quality, model explainability, adversarial robustness, and human–AI interaction. Effective CTI systems must balance automation with analyst oversight, ensuring transparency, accountability, and operational relevance. Hybrid intelligence models that combine AI-driven analytics with human expertise remain central to the responsible deployment of AI in CTI environments.

Looking forward, the continued evolution of CTI will depend not only on advances in AI techniques but also on progress in evaluation methodologies, governance frameworks, and human-centred system design. Framing CTI as a strategic intelligence capability—rather than solely a technical security function—may further support its role in anticipating and mitigating cyber threats within an increasingly interconnected and adversarial digital landscape.

Author Contributions: Conceptualization, M.B.-G., J.M.A.-P., M.Á.P.-J. and E.C.-d.-B.; methodology, M.B.-G., J.M.A.-P. and M.Á.P.-J.; validation, M.B.-G., J.M.A.-P., M.Á.P.-J. and E.C.-d.-B.; investigation, M.B.-G. and J.M.A.-P.; supervision, J.M.A.-P.; writing—original draft preparation, M.B.-G., J.M.A.-P. and M.Á.P.-J.; writing—review and editing, M.B.-G., J.M.A.-P., M.Á.P.-J. and E.C.-d.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AI	Artificial Intelligence
APIs	Application Programming Interface
APTs	Advanced Persistent Threats
ARIMA	AutoRegressive Integrated Moving Average
ASN	Autonomous System Number
ATT&CK	Adversarial Tactics, Techniques, and Common Knowledge
BERT	Bidirectional Encoder Representations from Transformers
BiLSTM	Bidirectional Long Short-Term Memory
C2	Command and Control
CEMA	Cyber Electromagnetic Activities
CNNs	Convolutional Neural Networks

ComplEx	Complex Embeddings
CTI	Cyber Threat Intelligence
CVEs	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
DL	Deep Learning
DNS	Domain Name System
EDR/XDR	Endpoint Detection and Response/Extended Detection and Response
ENISA	European Union Agency for Cybersecurity
FL	Federated Learning
GDPR	General Data Protection Regulation
GNNs	Graph Neural Networks
GPT	Generative Pre-trained Transformer
HITL	Human-In-The-Loop
HUMINT	Human Intelligence
ICS	Industrial Control Systems
IDS/IPS	Intrusion Detection System/Intrusion Prevention System
IoCs	Indicators of Compromise
IoT	Internet of Things
IP	Internet Address
IR	Incident Response
ISACs	Information Sharing and Analysis Centres
KPIs	Key Performance Indicators
LIME	Local Interpretable Model-agnostic Explanations
LLMs	Large Language Models
LSTMs	Long Short-Term Memory
MISP	Malware Information Sharing Platform & Threat Sharing
ML	Machine Learning
MLP	Multilayer Perceptron
NER	Named Entity Recognition
NIST	National Institute of Standards and Technology
NLP	Natural Language Processing
OpenCTI	Open Cyber Threat Intelligence
OSINT	Open-Source INTelligence
OT	Operational Technology
OTX	Open Threat Exchange
PII	Personally Identifiable Information
RBAC	Role-based Access Control
RHML	Reciprocal Human–Machine Learning
RL	Reinforcement Learning
RoBERTa	Robustly Optimised BERT Pretraining Approach
SHAP	SHapley Additive exPlanations
SIEMs	Security Information and Event Management
SIGINT	Signals Intelligence)
SLAs	Service Level Agreement
SOAR	Security Orchestration, Automation, and Response
SOCs	Security Operations Centres
STIX/TAXII	Structured Threat Information Expression/ Trusted Automated Exchange of Indicator Information
SVM	Support Vector Machines
TIP	Threat Intelligence Platforms
TLP	Traffic Light Protocol
TTPs	Tactics, Techniques, and Procedures
TRAM	TTP and Role Annotated Model
TransE	Translating Embeddings

UI	User Interface
URL	Uniform Resource Locator
XAI	Explainable Artificial Intelligence
YETI	Your Everyday Threat Intelligence

References

1. Balasubramanian, P.; Liyana, S.; Sankaran, H.; Sivaramakrishnan, S.; Pusuluri, S.; Pirttikangas, S.; Peltonen, E. Generative AI for Cyber Threat Intelligence: Applications, Challenges, and Analysis of Real-World Case Studies. *Artif. Intell. Rev.* **2025**, *58*, 336. [CrossRef]
2. Shah, S.; Khoda Parast, F. AI-Driven Cyber Threat Intelligence Automation. *arXiv* **2024**, arXiv:2410.20287. [CrossRef]
3. Sánchez del Monte, A.; Hernández-Álvarez, L. Analysis of Cyber-Intelligence Frameworks for AI Data Processing. *Appl. Sci.* **2023**, *13*, 9328. [CrossRef]
4. Binbeshr, F.; Imam, M.; Ghaleb, M.; Hamdan, M.; Rahim, M.A.; Hammoudeh, M. The Rise of Cognitive SOCs: A Systematic Literature Review on AI Approaches. *IEEE Open J. Comput. Soc.* **2024**, *6*, 360–379. Available online: https://ieeexplore.ieee.org/document/10858372 (accessed on 15 November 2025). [CrossRef]
5. Lima, M.; Viana, C.; Santos, W.R.M.; Neves, F.; Campos, J.R.; Aires, F. Toward Using Cyber Threat Intelligence with Machine and Deep Learning for IoT Security: A Comprehensive Study. *J. Supercomput.* **2025**, *81*, 1404. [CrossRef]
6. Aljahdali, A.O. A Cyber Threat Intelligence Model Using MISP and Machine Learning in a SOC Environment. *Int. J. Adv. Appl. Sci.* **2025**, *12*, 1–11. [CrossRef]
7. Naseer, I. Machine Learning Applications in Cyber Threat Intelligence: A Comprehensive Review. *Asian Bull. Big Data Manag.* **2024**, *3*, 190–200. [CrossRef]
8. Preuveneers, D.; Joosen, W. Sharing Machine Learning Models as Indicators of Compromise for Cyber Threat Intelligence. *J. Cybersecur. Priv.* **2021**, *1*, 140–163. [CrossRef]
9. Al-Taleb, N.; Saqib, N.A. Towards a Hybrid Machine Learning Model for Intelligent Cyber Threat Identification in Smart City Environments. *Appl. Sci.* **2022**, *12*, 1863. [CrossRef]
10. Sarker, I.H. AI-driven Cybersecurity and Threat Intelligence. In *AI-Driven Cybersecurity and Threat Intelligence*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 1–20. [CrossRef]
11. Sufi, F. A New Social Media-Driven Cyber Threat Intelligence. *Electronics* **2023**, *12*, 1242. [CrossRef]
12. Arazzi, M.; Arikkat, D.R.; Nicolazzo, S.; Nocera, A.; Rehiman, K.A.R.; Vinod, P.; Conti, M. NLP-Based Techniques for Cyber Threat Intelligence. *arXiv* **2023**, arXiv:2311.08807. [CrossRef]
13. Rani, N.; Saha, B.; Maurya, V.; Shukla, S.K. TTPXHunter: Actionable Threat Intelligence Extraction as TTPs from Finished Cyber Threat Reports. *arXiv* **2024**, arXiv:2403.03267. [CrossRef]
14. Marchiori, F.; Conti, M.; Verde, N.V. STIXnet: A Novel and Modular Solution for Extracting All STIX Objects in CTI Reports. In *Proceedings of the 18th International Conference on Availability, Reliability and Security (ARES '23), Benevento, Italy, 29 August–1 September 2023*; Association for Computing Machinery (ACM): New York, NY, USA, 2023. [CrossRef]
15. Guo, Y.; Liu, Z.; Huang, C.; Wang, N.; Min, H.; Guo, W.; Liu, J. A Framework for Threat Intelligence Extraction and Fusion. *Comput. Secur.* **2023**, *132*, 103371. [CrossRef]
16. Wagner, T.D.; Mahbub, K.; Palomar, E.; Abdallah, A.E. Cyber Threat Intelligence sharing: Survey and Research Directions. *Comput. Secur.* **2019**, *87*, 101589. [CrossRef]
17. Alrawashdeh, K.; Purdy, C.N. Fast activation function approach for Deep Learning based online anomaly intrusion detection. In *Proceedings of the IEEE BigDataSecurity/HPSC/IDS Conference, Omaha, NE, USA, 3–5 May 2018*; IEEE: Piscataway, NJ, USA, 2018; pp. 5–13. [CrossRef]
18. Ferrag, M.A.; Maglaras, L.; Moschoyiannis, S.; Janicke, H. Deep Learning for Cyber Security Intrusion Detection: Approaches, Datasets, and Comparative Study. *J. Inf. Secur. Appl.* **2020**, *50*, 102419. [CrossRef]
19. Aldhaheri, A.; Alwahedi, F.; Ferrag, M.A.; Battah, A. Deep Learning for Cyber Threat Detection in IoT Networks: A review. *Internet Things Cyber-Phys. Syst.* **2024**, *4*, 110–128. [CrossRef]
20. Xiao, P. Malware Cyber Threat Intelligence System for Internet of Things (IoT) Using Machine Learning. *J. Cyber Secur. Mobil.* **2023**, *13*, 53–90. [CrossRef]
21. El Jaouhari, S.; Etiabi, Y. FedCTI: Federated Learning and Cyber Threat Intelligence on the Edge for Secure IoT Networks. In *Proceedings of the International Conference on the Internet of Things, Nagoya, Japan, 7–10 November 2023*; ACM (Association for Computing Machinery): New York, NY, USA, 2023; pp. 98–104. [CrossRef]
22. Muheidat, F.; Mallouh, M.A.; Al-Saleh, O.; Al-Khasawneh, O.; Tawalbeh, L.A. Applying AI and Machine Learning to Enhance Automated Cybersecurity. *Procedia Comput. Sci.* **2024**, *251*, 287–294. [CrossRef]
23. Alevizos, L.; Dekker, M. Towards an AI-enhanced Cyber Threat Intelligence Processing Pipeline. *Electronics* **2024**, *13*, 2021. [CrossRef]

24. Dingankar, S.; Shankar, B.M.; Kalnawat, A.; Sani, A.; Dongre, Y.V.; Nagargoje, V.J. Enhancing Cyber Threat Intelligence with AI and ML: An Ensembled Approach. In *Smart Innovation, Systems and Technologies*; Springer: Berlin/Heidelberg, Germany, 2025; pp. 517–526. [CrossRef]
25. UNSW-NB15 Dataset. Available online: https://research.unsw.edu.au/projects/unsw-nb15-dataset (accessed on 15 November 2025).
26. Kshetri, N. Transforming Cybersecurity with Agentic AI. *Telecommun. Policy* **2025**, *49*, 102976. [CrossRef]
27. ENISA. ENISA Threat Landscape Report 2024. Available online: https://www.enisa.europa.eu (accessed on 15 November 2025).
28. Lekssays, A.; Sencar, H.T.; Yu, T. From Text to Actionable Intelligence: Automating STIX Entity and Relationship Extraction. *arXiv* **2025**, arXiv:2507.16576. [CrossRef]
29. Ramsdale, A.; Shiaeles, S.; Kolokotronis, N. A Comparative Analysis of Cyber-Threat Intelligence Sources, Formats and Languages. *Electronics* **2020**, *9*, 824. [CrossRef]
30. Connolly, J.; Davidson, M.; Schmidt, C. *The Trusted Automated Exchange of Indicator Information (TAXII)*; The MITRE Corporation: McLean, VA, USA, 2014; pp. 1–20.
31. Anomali ThreatStream. Available online: https://www.anomali.com/products/threatstream (accessed on 15 November 2025).
32. Recorded Future. Available online: https://www.recordedfuture.com/ (accessed on 15 November 2025).
33. VirusTotal. Available online: https://www.virustotal.com/gui/home/upload (accessed on 15 November 2025).
34. Chang, Y.; Wang, G.; Zhu, P.; He, J.; Kong, L. Research on Unified Cyber Threat Intelligence Entity Recognition Method Based on Multiple Features. In *Proceedings of the 2023 4th International Conference on Computers and Artificial Intelligence Technology (CAIT), Macau, China, 13–15 December 2023*; IEEE: Piscataway, NJ, USA, 2023; pp. 233–240. [CrossRef]
35. Lourenço, B.; Adão, P.; Ferreira, J.F.; Marques, M.M.; Vaz, C. Structuring Security: A Survey of Cybersecurity Ontologies, Semantic Log Processing, and LLMs Application. *arXiv* **2025**, arXiv:2510.16610. [CrossRef]
36. Alaeifar, P.; Pal, S.; Jadidi, Z.; Hussain, M.; Foo, E. Current Approaches and Future Directions for Cyber Threat Intelligence Sharing: A Survey. *J. Inf. Secur. Appl.* **2024**, *83*, 103786. [CrossRef]
37. Hagen, R.A.; Helkala, K. Complexity of Contemporary Indicators of Compromise. In *Proceedings of the European Conference on Cyber Warfare and Security (ECCWS), Jyväskylä, Finland, 27–28 June 2024*; European Conference on Cyber Warfare and Security (ECCWS): Jyväskylä, Finland, 2024; Volume 23, pp. 697–707. [CrossRef]
38. Phishtank. Available online: https://phishtank.org/ (accessed on 15 November 2025).
39. Sufi, F. A global Cyber-threat Intelligence System with Artificial Intelligence and CNN. *Decis. Anal. J.* **2023**, *9*, 100364. [CrossRef]
40. Djenna, A.; Barka, E.; Benchikh, A.; Khadir, K. Unmasking Cybercrime with AI-driven Cybersecurity Analytics. *Sensors* **2023**, *23*, 6302. [CrossRef]
41. Balasubramanian, P.; Nazari, S.; Kholgh, D.K.; Mahmoodi, A.; Seby, J.; Kostakos, P. A Cognitive Platform for Collecting CTI. *Decis. Anal. J.* **2025**, *14*, 100545. [CrossRef]
42. Kaur, R.; Gabrijelčič, D.; Klobučar, T. Artificial Intelligence for Cybersecurity: Literature review. *Inf. Fusion* **2023**, *97*, 101804. [CrossRef]
43. Srinivas, S.; Kirk, B.; Zendejas, J.; Espino, M.; Boskovich, M.; Bari, A.; Dajani, K.; Alzahrani, N. AI-Augmented SOC: A Survey of LLMs and Agents for Security Automation. *J. Cybersecur. Priv.* **2025**, *5*, 95. [CrossRef]
44. Ferrag, M.A.; Alwahedi, F.; Battah, A.; Cherif, B.; Mechri, A.; Tihanyi, N.; Bisztray, T.; Debbah, M. Generative AI in Cybersecurity: A Comprehensive Review. *Internet Things Cyber-Phys. Syst.* **2025**, *5*, 1–46. [CrossRef]
45. Zacharis, A.; Gavrila, R.; Patsakis, C.; Douligeris, C. Optimising AI Models for Intelligence Extraction. *J. Inf. Secur. Appl.* **2025**, *90*, 104037. [CrossRef]
46. Khan, T.; Alam, M.; Akhunzada, A.; Hur, A.; Asif, M.; Khan, M.K. Towards Augmented Proactive Cyberthreat Intelligence. *J. Parallel Distrib. Comput.* **2019**, *124*, 47–59. [CrossRef]
47. TheHive. Available online: https://strangebee.com/thehive/ (accessed on 15 November 2025).
48. MISP. Available online: https://www.misp-project.org/ (accessed on 15 November 2025).
49. OpenCTI. Available online: https://www.opencti.io/ (accessed on 15 November 2025).
50. YETI. Available online: https://yeti-platform.github.io/ (accessed on 15 November 2025).
51. ThreatQuotient (ThreatQ). Available online: https://www.threatq.com/ (accessed on 15 November 2025).
52. IBM X-Force Threat Intelligence. Available online: https://exchange.xforce.ibmcloud.com (accessed on 15 November 2025).
53. IBM QRadar. Available online: https://www.ibm.com/qradar (accessed on 15 November 2025).
54. IBM Watson. Available online: https://www.ibm.com/watson (accessed on 15 November 2025).
55. Spyros, A.; Koritsas, I.; Papoutsis, A.; Panagiotou, P.; Chatzakou, D.; Kavallieros, D.; Tsikrika, T.; Vrochidis, S.; Kompatsiaris, I. AI-based Holistic Framework for Cyber Threat Intelligence Management. *IEEE Access* **2025**, *13*, 20820–20846. [CrossRef]
56. Alzahrani, I.Y.; Lee, S.; Kim, K. Enhancing Cyber-threat Intelligence in the Arab World. *Electronics* **2024**, *13*, 2526. [CrossRef]
57. Iacovazzi, A.; Wang, H.; Butun, I.; Raza, S. Towards Cyber Threat Intelligence for the IoT. In *Proceedings of the 2023 IEEE 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT 2023), Paphos, Cyprus, 19–21 June 2023*; IEEE: Piscataway, NJ, USA, 2023; pp. 483–490. [CrossRef]

58. Wagner, C.; Dulaunoy, A.; Wagener, G.; Iklody, A. MISP: The Design and Implementation of a Collaborative Threat Intelligence Sharing Platform. In *Proceedings of the 3rd ACM Workshop on Information Sharing and Collaborative Security (WISCS '16), Vienna, Austria, 24 October 2016*; Association for Computing Machinery (ACM): New York, NY, USA, 2016; pp. 49–56. [CrossRef]
59. Anastopoulos, V. Using Social Network Analysis for Cyber Threat Intelligence. In *CCDCOE Research Paper 2022*; NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE): Tallinn, Estonia, 2022; pp. 1–12. Available online: https://ccdcoe.org/uploads/2022/07/Research_paper.pdf (accessed on 15 November 2025).
60. Stojkovski, B.; Lenzini, G.; Koenig, V.; Rivas, S. What's in a Cyber Threat Intelligence Sharing Platform?: A Mixed-Methods User Experience Investigation of MISP. *Asia-Pac. Comput. Syst.* **2021**, *8*, 385–398. [CrossRef]
61. Chatziamanetoglou, D.; Rantos, K. Weighted Quality Criteria for Cyber Threat Intelligence: Assessment and Prioritisation in the MISP Data Model. *Int. J. Inf. Secur.* **2025**, *24*, 160. Available online: https://link.springer.com/article/10.1007/s10207-025-01080-6 (accessed on 15 November 2025). [CrossRef]
62. van Haastrecht, M.; Golpur, G.; Tzismadia, G.; Kab, R.; Priboi, C.; David, D.; Răcătăian, A.; Baumgartner, L.; Fricker, S.; Ruiz, J.F.; et al. A Shared Cyber Threat Intelligence Solution for SMEs. *Electronics* **2021**, *10*, 2913. [CrossRef]
63. Amanov, R.; Isaev, R.; Doszhanov, E.; Abdykerimov, A. Using the MISP Platform to Collect Incident Data. *Preprints* **2025**, 2025050785. Available online: https://www.preprints.org/manuscript/202505.0785/v1 (accessed on 15 November 2025).
64. Dulaunoy, A.; Wagener, G.; Iklody, A.; Mokaddem, S.; Wagner, C. An Indicator Scoring Method for MISP Platforms. In *Proceedings of TNC18 Conference, Trondheim, Norway, 10–14 June 2018*; GÉANT: Trondheim, Norway, 2018; pp. 1–12. Available online: https://tnc18.geant.org/getfile/tnc18_paper_MISPCW.pdf (accessed on 15 November 2025).
65. Ruohonen, S.; Kirichenko, A.; Komashinskiy, D.; Pogosova, M. Instrumenting OpenCTI with a Capability for Attack Attribution Support. *Forensic Sci.* **2024**, *4*, 12–23. [CrossRef]
66. Argon, E.; Guiglionia, L.; Filigran. Case Study: Leveraging OpenCTI to Investigate a Phishing Attack. 2024. Available online: https://filigran.io/leveraging-opencti-to-investigate-a-phishing-attack-targeting-filigran-repo/ (accessed on 15 November 2025).
67. Onyagu, C.L.; Chibuike, I.L. From Reactive to Resilient: An OpenCTI-Driven Cyber Threat Intelligence Framework for Academic Institutions. *Adv. Int. J. Res.* **2025**, *6*, 1112. [CrossRef]
68. Manoj, J.; Keshav, V.A.; Peter, S.V.; Rajan, D.M. OPENCTI Plus: AI-Driven Cyber Threat Intelligence. *Int. J. Sci. Res. Sci. Technol.* **2025**, *12*, 394–403. Available online: https://ijsrst.com/paper/13808.pdf (accessed on 15 November 2025).
69. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why Should I Trust You?" Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016*; ACM: San Francisco, CA, USA, 2016; pp. 1135–1144. [CrossRef]
70. Baniecki, H.; Biecek, P. Adversarial Attacks and Defenses in Explainable Artificial Intelligence: A Survey. *Inf. Fusion* **2024**, *107*, 102303. [CrossRef]
71. Yuan, X.; He, P.; Zhu, Q.; Li, X. Adversarial Examples: Attacks and Defenses for Deep Learning. *arXiv* **2018**, arXiv:1712.07107. [CrossRef] [PubMed]
72. Gilbert, C.; Gilbert, M.A. The Impact of AI on Cybersecurity Defense Mechanisms: Future Trends and Challenges. *Glob. Sci. J.* **2024**, *12*, 9. Available online: https://www.globalscientificjournal.com/researchpaper/The_Impact_of_AI_on_Cybersecurity_Defense_Mechanisms_Future_Trends_and_Challenges_.pdf (accessed on 15 November 2025). [CrossRef]
73. Karunamurthy, A.; Kiruthivasan, R.; Gauthamkrishna, S. Human-in-the-Loop Intelligence: Advancing AI-Centric Cybersecurity for the Future. *Quing Int. J. Multidiscip. Sci. Res. Dev.* **2023**, *2*, 20–43. [CrossRef]
74. Cohen, D.; Te'eni, D.; Yahav, I.; Zagalsky, A.; Schwartz, D.; Silverman, G.; Mann, Y.; Elalouf, A. Human–AI enhancement of Cyber Threat Intelligence. *Int. J. Inf. Secur.* **2025**, *24*, 99. [CrossRef]
75. AI Act. Shaping Europe's Digital Future. Available online: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (accessed on 15 November 2025).
76. Gilbert, C.; Gilbert, M.A. Artificial Intelligence (AI) and Machine Learning (ML) for Predictive Cyber Threat Intelligence (CTI). *Int. J. Res. Publ. Rev.* **2025**, *6*, 584–617.
77. Almahmoud, Z.; Yoo, P.D.; Alhussein, O.; Farhat, I.; Damiani, E. A Holistic and Proactive Approach to Forecasting Cyber Threats. *Sci. Rep.* **2023**, *13*, 35198. Available online: https://www.nature.com/articles/s41598-023-35198-1 (accessed on 15 November 2025). [CrossRef]
78. Balantrapu, S.S. AI for Predictive Cyber Threat Intelligence. *Int. J. Multidiscip. Sci. Dev.* **2024**, *7*, 590–602. Available online: https://www.ijsdcs.com/index.php/IJMESD/article/download/590/228 (accessed on 15 November 2025).
79. Al-Turjman, F.; Zahmatkesh, H.; Mostarda, L. Enhancing Offloading with Cybersecurity in Edge Computing for Digital Twin-Driven Patient Monitoring. *IET Wirel. Sens. Syst.* **2024**, *14*, 363–380. [CrossRef]
80. Al-Turjman, F.; Zahmatkesh, H.; Mostarda, L. AI-Enabled Healthcare and Enhanced Computational Resource Management with Digital Twins into Task Offloading Strategies. *IEEE Access* **2024**, *12*, 90353–90370. [CrossRef]
81. Ekundayo, F.; Atoyebi, I.; Soyele, A.; Ogunwobi, E. Predictive Analytics for Cyber Threat Intelligence in Fintech Using Big Data and Machine Learning. *Int. J. Res. Publ. Rev.* **2024**, *5*, 5934–5948. [CrossRef]

82. Oriaro, S.; Mishra, S. Improving Cybersecurity Through Explainable Artificial Intelligence: A Systematic Literature Review. *Issues Inf. Syst.* **2025**, *3*, 387–400. Available online: https://iacis.org/iis/2025/3_iis_2025_387-400.pdf (accessed on 15 November 2025).
83. Tom, A.K.; Khraisat, A.; Jan, T.; Whaiduzzaman, M.; Nguyen, T.D.; Alazab, A. Survey of Federated Learning for Cyber Threat Intelligence in Industrial IoT: Techniques, Applications and Deployment Models. *Future Internet* **2025**, *17*, 409. [CrossRef]
84. Alam, M.T.; Bhusal, D.; Nguyen, L.; Rastogi, N. CTIBench: A Benchmark for Evaluating Large Language Models in Cyber Threat Intelligence. *arXiv* **2024**, arXiv:2406.07599. [CrossRef]
85. Cheng, Y.; Liu, Y.; Li, C.; Song, D.; Gao, P. CTIArena: Benchmarking LLM Knowledge and Reasoning Across Heterogeneous Cyber Threat Intelligence. *arXiv* **2025**, arXiv:2510.11974. [CrossRef]
86. Mohsin, A.; Janicke, H.; Ibrahim, A.; Sarker, I.H.; Camtepe, S. A Unified Framework for Human–AI Collaboration in Security Operations Centers with Trusted Autonomy. *arXiv* **2025**, arXiv:2505.23397. [CrossRef]
87. Singh, R.; Tariq, S.; Jalalvand, F.; Baruwal Chhetri, M.; Nepal, S.; Paris, C.; Lochner, M. LLMs in the SOC: An Empirical Study of Human–AI Collaboration in Security Operations Centres. *arXiv* **2025**, arXiv:2508.18947. [CrossRef]
88. Habibzadeh, A.; Feyzi, F.; Atani, R.E. Large Language Models for Security Operations Centers: A Comprehensive Survey. *arXiv* **2025**, arXiv:2509.10858. [CrossRef]
89. Jaffal, N.O.; Alkhanafseh, M.; Mohaisen, D. Large Language Models in Cybersecurity: A Survey of Applications, Vulnerabilities, and Defense Techniques. *AI* **2025**, *6*, 216. [CrossRef]
90. Xu, H.; Wang, S.; Li, N.; Wang, K.; Zhao, Y.; Chen, K.; Yu, T.; Liu, Y.; Wang, H. Large Language Models for Cyber Security: A Systematic Literature Review. *arXiv* **2024**, arXiv:2405.04760. [CrossRef]