World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

Mitigating third-party cyber risk using AI-powered threat intelligence and compliance analytics

Samson A. Adegbenro [1,*], Whenume O. Hundeyin [2], Florence Olinmah [3] and Chinedu A. Adaba [4]

[1] *Third Party Risk Audit, Global Strategy and Enterprise Platforms, USA.*
[2] *Assurance, Technology Risk, USA.*
[3] *Risk Analytics Reporting, Risk Management, USA.*
[4] *Management Information Systems, Bowie State University, USA.*

World Journal of Advanced Research and Reviews, 2025, 26(02), 2909-2929

Publication history: Received on 30 March 2025; revised on 16 May 2025; accepted on 18 May 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.26.2.1968

Abstract

In an increasingly interconnected digital landscape, third-party relationships have emerged as a critical vector for cyber risk exposure across industries. Increasingly, organisations rely on outside vendors, hosted and cloud services and supply chain partners to deliver key services but this reliance brings with it vulnerabilities that traditional cybersecurity models are not generally effective at mitigating. Third-party ecosystems are complex in nature, and combined with varying compliance postures and disjointed threat visibility, it becomes clear that the approach has to move from static risk assessment to dynamic, intelligence-led protection. This paper details how AI can revolutionize third-party risk management through AI-powered threat intelligence and compliance analytics. We begin by contextualizing the third party risk landscape, highlighting key challenges such as insufficient vendor transparency, dynamic threat actors, and regulatory fragmentation. The study then delves into the architecture and functionality of AI-driven platforms that ingest multi-source threat feeds, behavioral indicators, and compliance metrics to produce predictive risk scores. Through machine learning algorithms, these systems continuously adapt to emerging attack patterns and detect anomalies indicative of compromise within vendor networks. A major focus is placed on integrating compliance analytics—enabling organizations to automatically assess vendors against frameworks such as NIST, ISO 27001, and GDPR. This fusion of threat intelligence with regulatory mapping allows for proactive risk prioritization and enhanced vendor segmentation. Case studies and real-world applications demonstrate how organizations using AI-based tools have reduced response times, improved audit readiness, and minimized breach propagation across digital supply chains. By combining predictive AI modeling with automated compliance enforcement, organizations can move beyond reactive controls to implement a continuous, risk-informed approach to third-party cyber governance.

Keywords: Third-Party Risk Management; AI-Powered Threat Intelligence; Compliance Analytics; Cybersecurity Automation; Vendor Risk Scoring; Supply Chain Security

1. Introduction

1.1. The Expanding Digital Ecosystem and Third-Party Dependencies

The digital transformation of industries has led to a rapidly expanding digital ecosystem where organizations increasingly rely on a complex web of third-party vendors, cloud providers, and IT service suppliers to operate efficiently. This interconnected environment offers significant operational advantages scalability, specialization, cost savings but simultaneously introduces new vulnerabilities and systemic risks [1]. As digital services become more

* Corresponding author: Samson A. Adegbenro

modular and outsourced, core business functions are now often hosted outside organizational perimeters, blurring the lines of accountability for cybersecurity.

The third-party utilities have also shifted from simple IT services to more essential tasks such as storing data, hosting applications, and process automation. While in second party control in such industries including healthcare, finance and manufacturing, third parties often maintain sensitive information, proprietary algorithms, and operation technology networks [2]. These dependencies expand the attack surface, and cascade risks, where a single breached vendor can compromise security for tens to 100s of customer organizations.

Recent trends, such as the rise of software-as-a-service (SaaS), platform-as-a-service (PaaS), and complex API integrations, have created digital supply chains with multiple tiers of subcontractors and service nodes. Organizations often have limited visibility into these nested relationships, making it difficult to assess inherited risks or enforce standardized security protocols [3].

Moreover, regulatory frameworks are often outpaced by technological change, leaving legal and compliance gaps in cross-border data flows and vendor accountability. The reliance on third-party digital services necessitates a rethinking of cybersecurity strategy moving from perimeter defense to ecosystem governance. As this article will show, third-party risk management must become a central tenet of organizational cyber resilience in today's digital age [4].

1.2. Rising Cyber Threats in Supply Chains

As digital interdependencies grow, supply chains have become prime targets for cybercriminals, hacktivists, and state-sponsored actors. Attacks on suppliers are often more effective than direct breaches of hardened corporate networks, as smaller vendors may lack robust cybersecurity defenses yet retain privileged access to critical systems [5]. This trend has elevated the urgency of supply chain cybersecurity to a national and global security concern.

High-profile incidents such as the SolarWinds breach and the Kaseya ransomware attack exposed the vulnerability of digital supply chains, where malicious code or unauthorized access infiltrated trusted software updates and maintenance protocols [6]. These attacks affected not just the targeted firms but cascaded across governments, financial institutions, and infrastructure providers, highlighting the systemic nature of the threat.

Cyber threats in supply chains manifest in various forms including malware injection, unauthorized access, phishing campaigns, and data manipulation. Threat actors exploit weak identity controls, poor patch management, and opaque vendor practices to gain footholds in enterprise environments [7]. Adding to that complexity is the fact that supply networks have become increasingly global, and companies also need to factor in jurisdictional risks, diverse regulatory requirements and geopolitical risks. Even with higher levels of awareness and sophisticated attacks, many businesses lack full third-party risk assessments, ongoing monitoring and contractual cybersecurity requirements in vendor agreements. However, as adversaries continue to take advantage of these gaps, supply chains have become one of the most exploited and under-secured threat vectors in today's cyber risk landscape.[8].

1.3. Purpose and Scope of the Article

The purpose of this article is to explore the evolving risks posed by third-party digital dependencies and to outline strategic measures for enhancing cybersecurity resilience in supply chains. As digital ecosystems become increasingly interconnected, the lines between internal and external risk ownership blur, making it essential for organizations to adopt holistic and adaptive cyber defense frameworks [9].

This article aims to provide a multi-dimensional analysis of third-party cyber risks, grounded in current threat landscapes, real-world case studies, and emerging best practices. It discusses the drivers behind the proliferation of digital dependencies and outlines the common vulnerabilities exploited by threat actors in supply chain environments. The article also presents key regulatory trends and governance models shaping vendor risk oversight, including zero-trust architecture, continuous assurance, and cyber insurance integration [10].

The scope of the analysis spans both technical and organizational domains. On the technical front, it explores innovations in real-time monitoring, automated threat detection, and vendor access control. It covers, inter alia, organisational accountability at board level, cross functional working and the contract-based enforcement of cyber security. In this way, this paper adds to our understanding of what is needed for public and industry actors to work together to secure digital value chains and it does so based on analysis of where there are gaps, challenges and opportunities in managing third-party cybersecurity risk. Policy recommendations and a forward-looking roadmap

for the development of supply chain cybersecurity frameworks that can operate in the constantly changing threat landscape is presented at the end of this section [11].

2. Understanding third-party cyber risk

2.1. Defining Third-Party Risk in Cybersecurity Context

Third-party risk in cybersecurity refers to the potential threats and vulnerabilities that arise from an organization's reliance on external entities to perform business functions, manage data, or provide technology services. These outsiders could be software suppliers, public cloud providers, the IT consultancy, logistics companies or even subcontractors who get on a company's network remotely. The difficulty is the third-party has escalated privileges and access to the infrastructure to a deeper extent with less focussed security scrutiny, in comparison to internal operations [5]. This risk is compounded in the hyperconnected world of today, where business models are increasingly reliant on outsourcing and digital supply chains. The risk from third-parties is not only limited to direct partners, it encompasses also sub-vendors and fourth-party suppliers, thus leading to a complex mesh of nested and stealthy dependency [6]. A weakness or violation at any layer within this broad infrastructure could weaken or breach the whole system, causing data loss, penalties, disruption to business and damage to reputation. Unlike typical insider threats, or external threats, third-party threats arise from trusted relationship, and as such, they are inherently difficult to detect and even more problematic to mitigate." These dangers are more than just technical; they include compliance lapses, breaches of contract, and a lack of due diligence. For instance, a failure by vendor to comply with data protection laws may result in the hiring organisation incurring regulatory fines [7]. As it continues to evolve, organizations will need a wider network security risk lens that includes third-party attack vectors in order to adopt a continuous risk assessment. Enabling and Managing Third-Party TrustThird-party risk management is now vital for all organizations, rather than optional, or Deciphering obligation from strategic imperative and compelling companies to maintain trust, compliance, and operational integrity in a digital ecosystem that is rapidly growing [8].

2.2. Historical Breaches and High-Profile Incidents

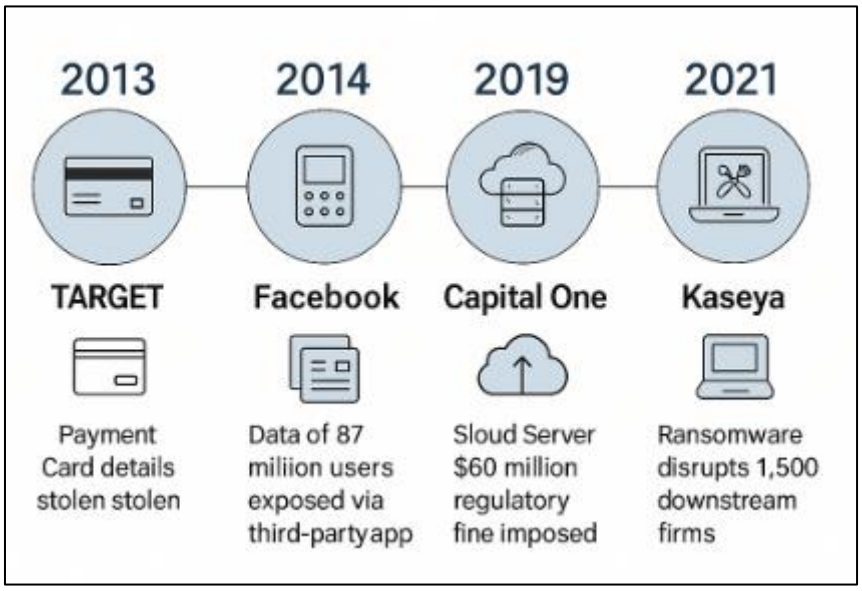

Figure 1 A timeline of such landmark incidents and their corresponding business impacts. These breaches reflect a recurring pattern: trusted third parties are often the weakest links in an organization's cybersecurity chain. Without robust oversight, these partnerships can inadvertently invite systemic risk, amplifying the urgency for comprehensive and continuous third-party risk governance [12]

A string of high-profile cyber attacks in the past decade has highlighted the catastrophic impact of insufficient third party risk management. These examples demonstrate how holes in vendor systems can be exploited to launch larger offensives that disrupt critical services and expose large caches of sensitive data. Some of the most notorious include the 2020 SolarWinds breach, in which attackers tampered with a software update for the Orion IT monitoring platform by placing malware in an otherwise routine software update. The software was used by more than 18,000 organizations around the world, including American federal agencies and Fortune 500 companies. It took months to

detect the breach giving the enemy deep penetration of internal systems [9]. Retail super power Target fell victim to a massive data breach in 2013 which compromised the credit card and personal information of more than 40 million customers. The intruders entered using a third-party HVAC vendor, demonstrating the potential cyber risk in suppliers that don't focus on IT. It caused hundreds of millions of dollars in losses and a huge reputation loss [10]. A second major hack took place in 2021 with Kaseya, a remote management software supplier. Ransomware should have increased by 'at least' 190% across the board Attackers exploited zero day bugs to unleash ransomware on more than 1,500 downstream organisations. This incident illustrated how service providers with extensive access rights can become a high-value target for supply chain attacks [11].

2.3. Categories of Third-Party Risk: Technical, Legal, and Reputational

Third-party risk in cybersecurity spans several interconnected domains, with the most critical being technical, legal, and reputational risks. Understanding these categories helps organizations prioritize controls and allocate resources effectively in managing their extended enterprise environment [13].

Technical risk refers to vulnerabilities that arise from the integration of third-party software, systems, or services into an organization's digital infrastructure. These may include unpatched software, misconfigured cloud environments, compromised APIs, or malicious updates. Technical risks are exacerbated when vendors lack basic security hygiene, such as multi-factor authentication or encryption protocols [14]. Attackers often exploit these technical entry points to gain unauthorized access or launch malware campaigns.

Legal risk involves exposure to compliance violations and regulatory penalties due to the actions or omissions of third-party providers. For example, under data protection frameworks such as the GDPR or HIPAA, an organization can be held liable for breaches caused by its vendors if it fails to conduct due diligence or lacks enforceable data processing agreements. Legal risk also encompasses breach of contract, intellectual property violations, and cross-border data transfer infractions [15].

Reputational risk results from public fallout following a third-party-related breach. Negative media coverage, customer distrust, and loss of investor confidence can have long-lasting effects that go beyond immediate financial losses. For instance, high-profile breaches can erode brand equity and lead to customer attrition, especially in industries like finance and healthcare, where data integrity is paramount [16].

These risk categories are not mutually exclusive. A single vendor incident can trigger technical disruptions, regulatory inquiries, and reputational damage simultaneously. Therefore, comprehensive risk assessments must incorporate multifactor evaluations that extend beyond traditional checklists and cover the full risk spectrum posed by third-party partnerships [17].

2.4. Challenges in Traditional Risk Management Approaches

Conventional risk management frameworks often fall short in addressing the complexities of third-party cybersecurity threats. Traditional approaches typically rely on static risk assessments conducted at the onboarding phase, with minimal follow-up or continuous oversight. This snapshot view fails to capture the evolving threat landscape, especially as vendors update systems, change subcontractors, or expand service offerings over time [18].

Another limitation is the over-reliance on self-assessment questionnaires and documentation, which may be incomplete, outdated, or overly optimistic. Without independent verification or real-time threat intelligence, organizations may unknowingly entrust sensitive data or access to vulnerable vendors. Moreover, siloed internal structures where procurement, legal, and IT departments operate independently further hinder integrated third-party risk management [19].

Resource constraints also play a role. Many organizations lack the tools, personnel, or expertise to monitor dozens or hundreds of vendor relationships. As a result, cybersecurity teams prioritize high-value vendors and overlook smaller suppliers who may still hold critical access privileges.

Finally, the absence of standardized benchmarks and interoperability frameworks complicates efforts to assess and compare vendor risk consistently across sectors. These gaps necessitate a shift toward dynamic, intelligence-led, and governance-driven third-party risk strategies that are better suited to today's hyperconnected digital environment [20].

3. The case for AI in cyber risk mitigation

3.1. Limitations of Static Risk Assessment Models

Static risk assessment models have historically formed the backbone of cybersecurity frameworks, particularly in third-party risk management. These models typically rely on pre-defined checklists, periodic questionnaires, and one-time security reviews conducted during the vendor onboarding process. While such assessments provide a snapshot of a vendor's cybersecurity posture at a given time, they fail to capture the dynamic nature of evolving threats and system changes [9].

One key limitation is their inability to account for real-time vulnerabilities or threat actor behaviors. As threat landscapes evolve rapidly with zero-day exploits, ransomware variants, and phishing techniques constantly emerging static models become outdated shortly after deployment. This time lag leaves organizations blind to risks introduced after the initial assessment, especially when vendors update software, onboard subcontractors, or change their security practices without notification [10].

Additionally, static models often rely heavily on self-reported data, which may be inaccurate, incomplete, or overly optimistic. Without continuous validation or live telemetry, organizations risk basing their decisions on outdated or unreliable information. These models also struggle to scale across increasingly complex digital supply chains, where hundreds of vendors operate with varying risk profiles [11].

Ultimately, the static approach fails to provide the situational awareness required for modern cybersecurity resilience. As threat actors grow more sophisticated and persistent, organizations must transition toward adaptive, real-time frameworks that leverage automation, behavioral analytics, and threat intelligence to keep pace with the constantly shifting cyber risk landscape [12].

3.2. Advantages of AI-Powered Threat Intelligence

AI-powered threat intelligence offers a transformative leap in how organizations detect, analyze, and respond to cybersecurity threats particularly those arising from third-party ecosystems. Unlike conventional tools, which operate on fixed rule sets or require human intervention, AI models use machine learning and data mining techniques to identify anomalies, forecast threats, and learn from evolving attacker behavior [13].

Table 1 Comparison of Conventional Threat Detection vs. AI-Powered Models

Feature	Conventional Detection	AI-Powered Detection
Detection Logic	Rule-based, static signatures	Behavioral, anomaly-based
Data Processing Volume	Limited	Massive, real-time
Response Time	Manual or delayed	Automated, near-instantaneous
Adaptability to New Threats	Requires manual updates	Continuously self-learning
False Positive Management	High and repetitive	Context-aware filtering
Supply Chain Risk Visibility	Low	High, cross-vendor pattern detection

One significant advantage of AI lies in its capacity to process massive volumes of structured and unstructured data in real time. Sources include firewall logs, endpoint telemetry, dark web monitoring, phishing email patterns, and threat feeds from vendor ecosystems. AI algorithms analyze these inputs to uncover hidden correlations and detect early indicators of compromise before damage occurs [14]. This real-time visibility is especially valuable in supply chains, where risks may emerge from obscure or lower-tier vendors.

Another strength is adaptability. AI models evolve by training on fresh data, enabling them to recognize previously unseen malware signatures, command-and-control domains, and attack vectors. This learning process empowers organizations to stay ahead of threat actors who constantly change their tactics to bypass traditional defenses [15]. For example, deep learning models can distinguish between normal and anomalous login behaviors across distributed vendor accounts, raising alerts on potential credential abuse.

AI also enhances prioritization. By assessing risk based on threat severity, asset criticality, and business context, AI helps security teams focus on high-impact incidents. This risk-based triaging minimizes alert fatigue and improves response time, especially when applied within Security Information and Event Management (SIEM) or Extended Detection and Response (XDR) platforms [16].

Moreover, AI-driven tools can automate remediation actions such as revoking vendor access, isolating compromised devices, or flagging suspicious API calls. These actions not only reduce manual workload but also accelerate containment, limiting lateral movement within networks [17].

By embracing AI-powered threat intelligence, organizations gain the agility, depth, and foresight needed to secure interconnected digital ecosystems. It redefines threat detection from a reactive exercise to a proactive defense strategy aligned with modern cyber risk realities [18].

3.3. AI and the Evolution of Continuous Monitoring Paradigms

Continuous monitoring represents a shift from episodic to ongoing cybersecurity vigilance an approach vital for managing third-party risks in an era of persistent threats and complex digital supply chains. AI plays a central role in this evolution by enabling real-time, intelligent surveillance of assets, users, and vendor ecosystems [19].

Traditional monitoring relied on periodic scans, scheduled audits, or manual log reviews, often resulting in delayed detection and reactive mitigation. In contrast, AI-integrated monitoring systems continuously ingest telemetry from endpoints, networks, APIs, and external threat feeds to build dynamic baselines of expected behavior. Deviations from these baselines such as abnormal data transfers, privilege escalations, or access attempts outside business hours are flagged instantly for investigation [20].

In third-party risk contexts, AI enhances visibility by mapping the digital footprint of vendors across systems, flagging policy violations, and assessing risk posture in real time. For instance, Natural Language Processing (NLP) algorithms can parse vendor contracts and flag ambiguous security clauses, while machine learning can evaluate historical breach patterns to forecast vendor risk scores [21].

AI-powered continuous monitoring also supports “trust but verify” principles through automated validation of vendor compliance. This includes tracking adherence to Service Level Agreements (SLAs), verifying encryption standards, and confirming endpoint protection. Alerts are integrated into dashboards, providing cybersecurity teams and executives with actionable intelligence for decision-making.

Importantly, continuous AI-driven monitoring facilitates rapid incident response. It enables the automatic suspension of third-party connections in the event of detected anomalies and triggers forensics pipelines to contain breaches early. This capability is essential given that modern breaches can propagate across networks within minutes [22].

As cyber threats grow more agile and supply chains more opaque, AI-powered continuous monitoring becomes indispensable. It offers a proactive, scalable, and context-aware defense architecture laying the groundwork for real-time governance and adaptive cyber resilience in third-party ecosystems.

4. Frameworks and architectures for AI-powered threat intelligence

4.1. Core Components of AI-Based Risk Management Platforms

AI-based risk management platforms are engineered to provide continuous, scalable, and context-aware oversight of cyber threats particularly in ecosystems with extensive third-party relationships. These platforms combine multiple technological layers that work in synergy to automate threat detection, prioritize risks, and enable timely interventions. At their core, such systems consist of data acquisition modules, machine learning (ML) engines, real-time dashboards, and integrated orchestration interfaces [14].

The first core component is the data ingestion layer, which collects structured and unstructured data from multiple sources internal logs, network telemetry, vendor APIs, system configurations, and cloud activity. This layer ensures that data volume, velocity, and variety are handled efficiently, enabling platforms to track changes in the attack surface over time [15].

Next is the feature engineering module, where raw data is transformed into meaningful attributes for machine learning. This involves extracting indicators such as login frequency, file access patterns, port activity, and geolocation anomalies, which serve as predictors of potential compromise. These engineered features enhance the quality and granularity of the models' learning capability [16].

The ML analytics engine is the intelligence core. It houses various algorithms for classification, clustering, and regression to detect malicious activity. This includes anomaly detection models that flag deviations from baseline behavior and supervised models trained on historical breaches. Some platforms also deploy ensemble techniques to improve model accuracy and reduce false positives [17].

Another essential component is the user interface and visualization dashboard, which presents risk scores, real-time alerts, and trend analyses to cybersecurity analysts and decision-makers. These dashboards often integrate with Security Orchestration, Automation, and Response (SOAR) platforms for automated incident handling [18].

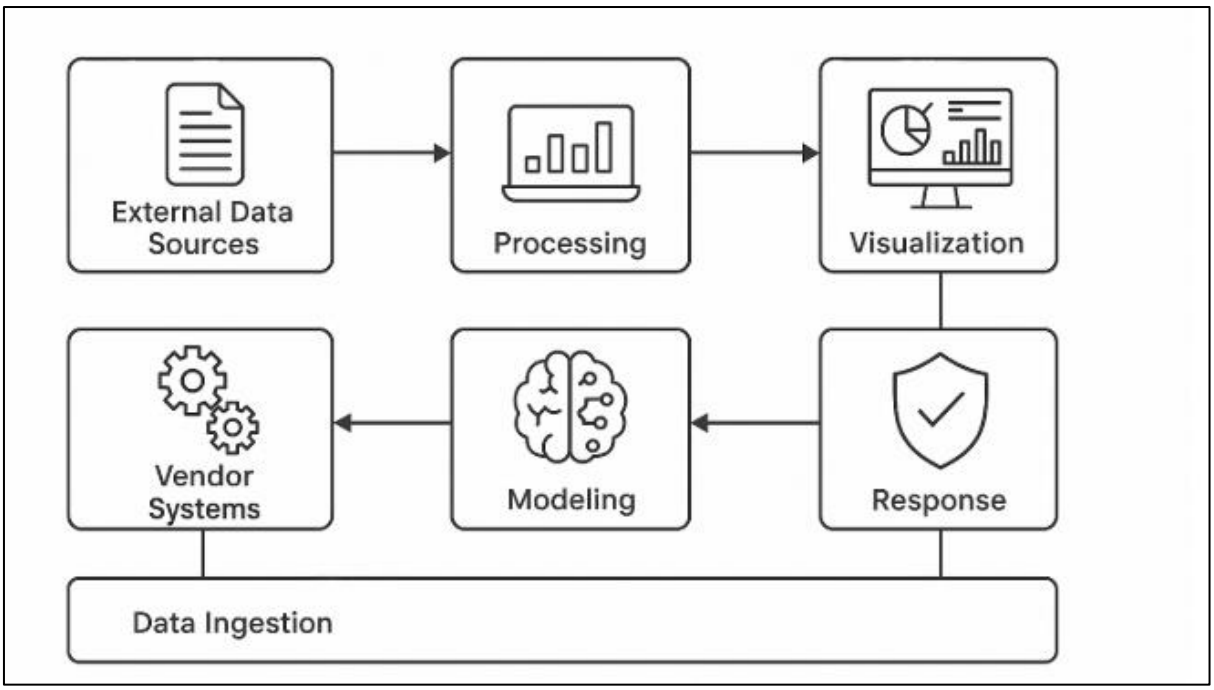

Figure 2 A typical system architecture of an AI-powered third-party cyber risk platform, showing the interconnection between ingestion, processing, modeling, visualization, and response mechanisms

Finally, **integration layers** with existing tools like SIEM, endpoint protection platforms, and governance risk and compliance (GRC) software ensure interoperability and scalability. These core components collectively transform AI risk platforms into real-time guardians of third-party cybersecurity hygiene, bridging gaps left by manual assessments and static defenses [19].

4.2. Data Ingestion, Feature Engineering, and Anomaly Detection

Effective AI-based third-party cyber risk platforms rely heavily on sophisticated data ingestion and feature engineering pipelines. These pipelines serve as the foundation upon which anomaly detection and threat attribution mechanisms are built. The quality, diversity, and timeliness of input data directly affect the accuracy and reliability of AI predictions [20].

Data ingestion begins with the continuous collection of data from multiple sources. These may include firewall logs, intrusion detection systems (IDS), vendor access records, endpoint sensors, and cloud service APIs. Additional sources may encompass identity and access management (IAM) tools, DevOps pipelines, and system health metrics. Stream processing frameworks like Apache Kafka or Flink are commonly employed to support real-time ingestion, enabling the detection of fast-moving threats before they propagate across systems [21].

Once collected, this data is passed to **feature engineering modules**, where it is structured and enriched to provide meaningful context. Feature engineering may involve aggregating access frequency by user, encoding login time patterns, or calculating entropy in command-line behavior. Derived metrics such as average dwell time, access deviation index, or protocol frequency become valuable predictors of risk behavior in vendor environments [22].

Anomaly detection algorithms then analyze these engineered features to flag deviations from expected patterns. These may include statistical models such as Gaussian Mixture Models (GMM) or K-means clustering, and ML-based approaches such as Isolation Forests or Autoencoders. For example, if a third-party service account initiates privileged actions outside of its baseline activity window, the system can trigger alerts and even automate containment actions [23].

Some platforms also incorporate hybrid detection frameworks combining rule-based thresholds with adaptive learning algorithms to balance precision and recall. This is crucial in third-party ecosystems, where behavioral baselines may vary widely between vendors due to differing service roles, geographies, or operational norms [24].

The combination of rigorous data ingestion, advanced feature extraction, and context-aware anomaly detection equips organizations to detect and respond to potential threats in real-time. It enables cyber teams to move beyond signature-based models and embrace behavioral insights tailored to the specific risks posed by external entities.

4.3. Integrating Open-Source Intelligence (OSINT) and Proprietary Feeds

Integrating Open-Source Intelligence (OSINT) and proprietary threat feeds significantly enhances the threat detection capabilities of AI-based third-party risk platforms. OSINT refers to publicly available information sources such as threat forums, CVE repositories, security blogs, leaked credential databases, and social media chatter. These sources offer early indicators of threat actor intent, emerging malware strains, and compromised vendor accounts [25].

By mining OSINT, AI models can dynamically update threat landscapes and detect correlations between observed vendor activity and known malicious behaviors. For instance, if a vendor's IP address is flagged in a threat intelligence feed for hosting phishing infrastructure, the platform can cross-reference that with internal access logs for rapid threat attribution.

Proprietary feeds from cybersecurity vendors add another layer of precision. These include curated threat signatures, threat actor profiles, vulnerability intelligence, and exploit detection frameworks. Combining OSINT with these commercial datasets provides richer context for predictive models, allowing the system to differentiate between benign anomalies and credible threats [26].

Effective integration involves aligning OSINT and proprietary feed ingestion with AI feature sets, creating fused indicators of compromise (IOCs) and enhancing detection logic. This fusion supports pre-emptive defenses and strengthens situational awareness across third-party ecosystems.

4.4. Role of Natural Language Processing (NLP) in Threat Attribution

Natural Language Processing (NLP) plays a pivotal role in enhancing the analytical depth of AI-based risk platforms, particularly in the domain of **threat attribution**. As cyber threat intelligence often resides in unstructured text advisories, incident reports, dark web conversations NLP techniques are used to extract actionable insights and map them to specific threat actors or campaigns [27].

For example, NLP algorithms can parse vendor audit logs, news articles, or breach disclosures to detect indicators like threat actor aliases, attack vectors, or exploited CVEs. Named Entity Recognition (NER) and topic modeling help classify entities such as malware types, attack techniques (e.g., phishing or privilege escalation), and temporal cues that establish when the threat was observed [28].

Sentiment analysis and semantic similarity models can further identify urgency or credibility within threat narratives. This is particularly useful when assessing chatter about zero-day vulnerabilities linked to known vendors. Once processed, this intelligence feeds into attribution engines, allowing security teams to correlate third-party activity with specific threat campaigns.

Ultimately, NLP bridges the gap between human-readable threat intelligence and machine-driven risk scoring, transforming textual data into structured knowledge that sharpens attribution accuracy and response prioritization across complex vendor landscapes [29].

5. Compliance analytics and automated governance

5.1. Overview of Regulatory and Industry Standards

The integration of AI into third-party cyber risk management must align with existing regulatory and industry frameworks to ensure legal compliance, data integrity, and ethical accountability. Key global standards such as NIST (National Institute of Standards and Technology), ISO/IEC 27001, GDPR (General Data Protection Regulation), **ISO/IEC 42001** and HIPAA (Health Insurance Portability and Accountability Act) set foundational requirements for information security, risk assessment, privacy, and vendor governance [19].

These frameworks, though diverse in scope, share common goals: to enhance data protection, promote transparency, and ensure organizational accountability. NIST's Cybersecurity Framework emphasizes risk-based approaches to identify, detect, respond, and recover from cyber incidents. ISO/IEC 27001 provides a structured information security management system (ISMS) applicable across sectors. Meanwhile, GDPR mandates explicit data processing controls and accountability for third-party data handling within the EU. HIPAA, primarily focused on healthcare, outlines stringent rules for protected health information (PHI), including third-party access and auditing [20].

When applied to AI-driven platforms, these frameworks guide how systems must manage data privacy, algorithmic accountability, and control enforcement. For instance, GDPR requires AI models to ensure explainability and provide data subjects with the right to access, correct, or delete personal data constraints that influence both model architecture and data retention policies [21].

Table 2 Maps each major standard to relevant AI analytics functions

Framework	Mapped AI Functions
NIST CSF	Anomaly detection, response automation, risk scoring
ISO 27001	Asset inventory classification, continuous compliance tracking
GDPR	Data minimization, explainable AI, consent validation
HIPAA	Access logging, PHI encryption modeling, audit trails
ISO/IEC 42001	AI lifecycle governance, algorithmic accountability, bias mitigation, human oversight

Adhering to these frameworks ensures that AI-enhanced third-party risk platforms support not only advanced analytics but also regulatory trustworthiness and ethical resilience [22].

5.2. Dynamic Vendor Risk Profiling and Scoring

Dynamic vendor risk profiling is a cornerstone of AI-enabled third-party risk management, offering real-time evaluation of a vendor's cybersecurity posture across technical, legal, and operational dimensions. Traditional risk scoring often based on annual surveys or subjective assessments fails to reflect the rapidly changing threat landscape or the evolving risk behavior of vendors post-onboarding [23].

AI-driven profiling models utilize live telemetry, threat feeds, behavioral baselines, and contextual metadata to create dynamic and continuously updated vendor risk scores. These scores are calculated based on weighted indicators such as patch cadence, endpoint security configuration, third-party software exposure, data residency compliance, and historical incident frequency. Scores can be tailored by criticality level, sector, or data access scope, providing granularity that enables precision targeting of mitigation efforts [24].

For example, a vendor providing back-end IT services with elevated admin privileges would be assessed differently from a vendor handling marketing analytics. Risk models assign risk multipliers based on access levels, sectoral risk (e.g., healthcare or finance), and regional threat vectors, such as geopolitical instability or data sovereignty concerns. Advanced systems use Bayesian inference and reinforcement learning to refine risk weights as new data becomes available [25].

Integrating this scoring into procurement and compliance processes enables smarter vendor selection and performance-based contract renewals. Organizations can also set threshold alerts, triggering automated policy reviews

or access limitation when a vendor's risk score crosses a predefined threshold. Dashboards visualize scoring trends over time, helping executives track vendor risk exposure and prioritize response actions [26].

Beyond the technical layer, dynamic scoring contributes to broader governance objectives by enabling transparency, audit readiness, and regulatory alignment. It transforms vendor management from a reactive process to a forward-looking, risk-aware discipline that evolves with both external threats and internal priorities [27].

5.3. Continuous Control Monitoring (CCM) and Audit Automation

Continuous Control Monitoring (CCM) is an AI-driven practice that enables real-time validation of cybersecurity controls across internal and third-party systems. Traditional audits and compliance reviews performed quarterly or annually often fail to capture violations or lapses that occur in between review periods. CCM addresses this by offering persistent oversight through automated checks, anomaly detection, and policy enforcement [28].

In AI-enhanced CCM platforms, control checks are embedded into digital workflows. These include monitoring for multi-factor authentication enforcement, encryption key rotations, user access reviews, and vulnerability patching intervals. Rules are defined based on compliance frameworks (e.g., NIST, ISO 27001) and internal policy requirements, allowing systems to automatically flag control breaches and suggest remediation steps [29].

Machine learning further enhances CCM by learning from prior violations and adapting rule sensitivity based on historical compliance patterns. For example, if a vendor consistently fails to meet endpoint compliance standards, the system may increase the frequency or depth of checks specific to that vendor. This targeted approach reduces noise and improves audit efficacy [30].

Audit automation builds on CCM by generating evidence-ready reports, control validation logs, and digital compliance dashboards. These outputs streamline regulatory reporting, reduce manual workload, and ensure continuous readiness for internal and external audits. Integration with GRC tools further ensures traceability and centralized documentation [31].

By embedding CCM and automated audits into third-party ecosystems, organizations reduce compliance risk, improve control effectiveness, and foster a culture of continuous assurance crucial in today's fast-evolving threat and regulatory environment.

5.4. Case Study: Enhancing Supply Chain Compliance with AI Tools

A multinational pharmaceutical company implemented an AI-based third-party risk platform to manage its global vendor network, comprising over 800 suppliers across 40 countries. Prior to deployment, compliance reviews were manual, fragmented, and reactive, leading to audit delays and undetected vulnerabilities [32].

The AI platform integrated vendor telemetry, OSINT feeds, and real-time control monitoring to produce dynamic risk scores and continuous compliance indicators. Anomaly detection models flagged irregular login attempts from third-party contractors in high-risk regions, triggering automated access suspensions and incident investigations. The platform's NLP engine also parsed vendor contracts to identify missing data protection clauses, prompting legal review workflows [33].

Over 12 months, the organization achieved a 40% reduction in vendor-related audit findings and a 60% improvement in time-to-resolution for access control violations. Dashboards provided executives with up-to-date compliance heat maps, allowing targeted risk remediation across supply chain tiers.

This case underscores how AI tools transform supply chain governance from a static function into a proactive compliance ecosystem. The integration of predictive analytics, automated audits, and contextual threat intelligence enabled the company to meet regulatory demands while improving vendor accountability and operational continuity [34].

6. AI models in practice: technical deep dive

6.1. Machine Learning Models Used (Random Forest, XGBoost, Neural Nets)

AI-powered third-party cyber risk management platforms rely on a diverse set of machine learning (ML) models to detect anomalies, forecast breaches, and prioritize vendor vulnerabilities. Among the most commonly used algorithms are Random Forest, XGBoost (Extreme Gradient Boosting), and various types of Neural Networks each offering unique strengths in pattern recognition, classification, and predictive analysis [23].

Random Forest is a widely used ensemble learning technique that builds multiple decision trees and aggregates their results for classification or regression tasks. In cybersecurity, Random Forest is particularly effective in modeling access behavior, login frequency anomalies, and user-entity activity patterns. Its resistance to overfitting and ability to handle high-dimensional data make it a strong choice for environments with complex, non-linear relationships between features [24].

XGBoost, an advanced implementation of gradient boosting algorithms, offers enhanced performance for structured data and is known for its speed, accuracy, and scalability. In third-party risk platforms, XGBoost is often employed to predict the likelihood of a vendor breach based on multiple weighted inputs such as patch management delays, region-specific threat indicators, and control audit failures. Its built-in regularization capabilities help prevent overfitting, especially in imbalanced datasets where breaches are rare events [25].

Neural Networks, particularly deep learning models like feedforward neural nets and Long Short-Term Memory (LSTM) architectures, are well-suited for time-series data and unstructured inputs such as logs, textual threat intelligence, and behavioral sequences. These models are applied in tasks like log anomaly detection, NLP-based contract analysis, and real-time correlation of vendor activity with external threat signals [26].

Model selection depends on the specific analytic task e.g., Random Forest for risk classification, XGBoost for breach likelihood scoring, and Neural Nets for anomaly detection in sequential datasets. Often, hybrid or ensemble systems that combine these models are deployed to improve detection precision and resilience across diverse threat vectors [27].

The integration of these models enables AI platforms to perform both granular analytics and broad pattern recognition providing a multi-layered defense mechanism for third-party cybersecurity management.

6.2. Risk Prediction vs. Threat Detection Models

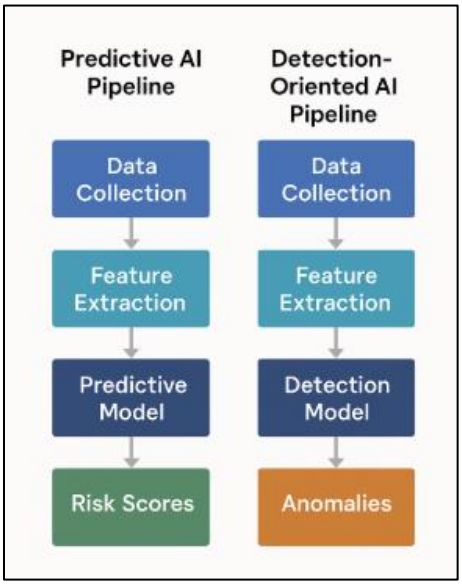

Figure 3 Contrast, showing side-by-side data flows, feature sets, and outcomes for both model types.

Machine learning models in third-party cybersecurity ecosystems can be broadly categorized into two functional groups: risk prediction models and threat detection models. Although both serve complementary purposes, their operational logic, input features, and deployment strategies differ significantly [28].

Risk prediction models focus on forecasting the probability of future security incidents or vendor failures. These models typically use historical data, vendor profiles, compliance records, and behavioral telemetry to estimate breach likelihood or control performance degradation. For example, XGBoost or logistic regression can be trained to output a vendor breach risk score based on variables like MFA enforcement, location-specific threat trends, and number of unresolved vulnerabilities [29]. Risk prediction enables proactive mitigation measures, such as conditional access policies or prioritization of high-risk vendors for audits.

Threat detection models, in contrast, operate in real time or near-real time to identify deviations from expected behavior. These include supervised classifiers like Random Forest and unsupervised methods such as Isolation Forests or Autoencoders, which flag anomalies like unusual file transfers, geographic login anomalies, or unauthorized API calls from vendor accounts [30].

The fundamental difference lies in temporal orientation prediction models look forward based on aggregate history, while detection models focus on identifying immediate risks as they occur.

While prediction enhances strategic planning and long-term vendor oversight, detection supports operational resilience and incident response. The integration of both types is essential for a comprehensive third-party cyber risk framework that spans prevention, monitoring, and real-time defense.

6.3. Model Training, Bias Handling, and Accuracy Evaluation

Training machine learning models for third-party risk management involves rigorous data preprocessing, hyperparameter tuning, and ongoing validation to ensure reliability and fairness. These models are typically trained on historical datasets that include vendor performance logs, security incident records, access behaviors, and threat intelligence signals [31].

Model training begins with data normalization and cleaning, followed by feature selection to isolate the most predictive variables. Cross-validation techniques, such as k-fold validation, are used to mitigate overfitting and improve generalization. For time-series models like LSTMs, windowed sampling is applied to preserve temporal dependencies [32].

A critical challenge in training is bias mitigation. Imbalanced datasets where breach incidents are rare compared to normal activity can lead to models that favor the majority class, undermining their effectiveness. Techniques such as Synthetic Minority Over-sampling Technique (SMOTE), cost-sensitive learning, and balanced class weights are applied to address this issue. Fairness constraints may also be introduced to prevent systemic bias against vendors from specific regions or sectors [33].

Accuracy evaluation is conducted using multiple metrics: precision, recall, F1-score, Area Under the Receiver Operating Characteristic Curve (AUC-ROC), and Matthews Correlation Coefficient (MCC). These metrics provide a balanced view of model performance, especially in skewed datasets where high accuracy may mask poor detection of true positives [34].

Additionally, explainability is crucial especially under regulations like GDPR. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are used to interpret model decisions, increasing transparency and trust.

By incorporating robust training, bias handling, and performance evaluation protocols, AI models can maintain integrity and reliability while enhancing decision-making in third-party cyber risk ecosystems [35].

7. Implementation strategies and integration

7.1. Vendor Risk Assessment Workflows in Large Enterprises

In large enterprises, vendor risk assessment workflows are complex, multi-tiered processes that involve a combination of procurement, legal, cybersecurity, and compliance teams. These workflows are designed to evaluate the potential

risk a third-party vendor poses to the organization's data security, regulatory compliance, and operational resilience. With hundreds or thousands of active vendors, large enterprises must rely on structured, repeatable workflows enhanced by automation and analytics [27].

The typical vendor risk assessment process begins with pre-screening, during which vendors are evaluated based on sector, geography, service criticality, and regulatory exposure. This step may include querying external threat intelligence sources or leveraging pre-existing vendor risk databases [28]. Once shortlisted, vendors undergo due diligence, involving detailed cybersecurity questionnaires, audits, and analysis of their technical capabilities, including encryption protocols, access controls, and incident response procedures.

Modern workflows increasingly incorporate AI-driven risk scoring, which supplements qualitative inputs with real-time telemetry and threat indicators. These scores are dynamically updated based on vendor behavior, breach disclosures, and control audit results. Tools like NLP can parse vendor contracts to flag incomplete or missing data protection clauses, while anomaly detection algorithms monitor access patterns during the onboarding phase [29].

Following assessment, vendors are either approved, conditionally accepted (pending remediation), or rejected. Conditional vendors are assigned mitigation tasks tracked through workflow tools with embedded SLA monitoring. Approved vendors enter continuous monitoring loops, with access reviews and compliance re-verification at set intervals.

By leveraging AI and automation in these workflows, enterprises not only enhance the accuracy of risk profiling but also streamline the decision-making process, enabling faster, evidence-based vendor onboarding and ongoing oversight [30].

7.2. Interoperability with Existing GRC and SIEM Platforms

For AI-based third-party cyber risk tools to be effective within large enterprises, interoperability with existing Governance, Risk, and Compliance (GRC) platforms and Security Information and Event Management (SIEM) systems is essential. These integrations ensure seamless data flow, centralized risk visibility, and coordinated incident response, reducing operational friction and improving decision accuracy [31].

GRC platforms such as RSA Archer, MetricStream, or ServiceNow serve as the backbone for policy enforcement, audit management, and compliance tracking. Integrating AI tools into these ecosystems enables automated ingestion of vendor risk scores, control assessment outputs, and anomaly detection alerts. For instance, risk thresholds flagged by the AI system can automatically trigger GRC workflows such as contract review, escalation notices, or updated risk acceptance documentation [32].

SIEM systems, such as Splunk or IBM QRadar, collect and correlate security events across the organization. AI-based tools enhance this capability by supplying enriched, behavior-based threat intelligence on third-party access patterns, suspicious file movements, or geolocation anomalies. This real-time input enables more precise rule-building and automated detection rules for third-party-related incidents [33].

Table 3 Key Integration Points for AI Tools within GRC Ecosystems

Functionality	GRC System Integration Point	SIEM System Integration Point
Risk Score Sync	Vendor profiles & onboarding modules	Threat prioritization engine
Alert Correlation	Compliance incident workflows	Event correlation logic
Control Violation Reporting	Automated audit log generation	Real-time alert console
Policy Trigger Automation	SLA & contract enforcement workflows	Intrusion prevention triggers

Interoperability also facilitates bidirectional learning, allowing AI models to refine predictions using historical audit data and incident response logs from GRC and SIEM platforms. This closed-loop system supports dynamic risk governance and reinforces regulatory compliance [34].

7.3. Cost, Scalability, and ROI Considerations

Adopting AI-based third-party cyber risk management tools requires a thorough evaluation of cost, scalability, and return on investment (ROI) factors that significantly influence enterprise adoption decisions. While initial investments can appear high due to licensing, integration, and training costs, the long-term savings and risk reduction benefits often outweigh upfront expenses [35].

Cost considerations typically include platform licensing fees (often tiered by the number of vendors or data volume), professional services for implementation, and indirect costs such as staff training and change management. Organizations must also budget for infrastructure, especially if deploying AI models in on-premise or hybrid cloud environments [36].

However, AI tools provide substantial ROI through labor savings and breach avoidance. Automating vendor onboarding, risk scoring, and compliance tracking reduces manual workloads and accelerates time-to-decision. This leads to more efficient procurement cycles and stronger vendor accountability. Moreover, by identifying and mitigating third-party vulnerabilities before exploitation, organizations can avoid costly incidents, reputational damage, and regulatory fines [37].

Scalability is a major strength of AI platforms. These systems are designed to ingest vast amounts of telemetry and adapt to complex vendor ecosystems. AI models can scale across global operations, supporting multilingual interfaces, region-specific compliance modules, and customizable workflows suited to local regulations.

Enterprises also benefit from predictive budgeting as AI tools provide visibility into long-term vendor risk trajectories, enabling strategic planning and risk-adjusted investment allocation. Ultimately, the convergence of cost-efficiency, scalability, and predictive insight ensures that AI-enabled risk platforms deliver measurable value over time, transforming third-party governance into a strategic business enabler [38].

8. Real-world applications and impact

8.1. Sector-Specific Use Cases: Healthcare, Finance, and Energy

AI-powered third-party risk platforms are increasingly tailored to address the unique threat landscapes and compliance demands of critical sectors such as healthcare, finance, and energy. Each of these industries operates under stringent regulatory frameworks and manages high-value data, making targeted risk reduction essential [32].

In healthcare, organizations manage electronic health records (EHR), medical IoT devices, and external service providers such as telehealth platforms and billing processors. AI systems in this sector prioritize real-time monitoring of third-party data flows and anomaly detection in protected health information (PHI) access patterns. Natural Language Processing (NLP) tools assist in scanning Business Associate Agreements (BAAs) to ensure compliance with HIPAA and to flag missing privacy clauses before onboarding [33].

In the financial sector, the focus shifts to vendor-related fraud prevention, regulatory compliance (e.g., SOX, PCI-DSS), and insider threat detection. AI models are used to monitor transactional behaviors of third-party fintech integrations, assess geo-fraud risk, and evaluate the enforcement of multi-factor authentication across vendor access points. Predictive analytics also help identify liquidity risks stemming from third-party software vulnerabilities that could impact operations [34].

The energy sector especially utilities and oil and gas relies heavily on third-party contractors for field operations, pipeline monitoring, and smart grid management. AI is deployed to monitor remote access behaviors, detect anomalies in SCADA interactions, and validate that external contractors adhere to NERC CIP (Critical Infrastructure Protection) controls. Additionally, AI-powered supply chain analysis flags equipment vendors located in regions associated with geopolitical risk [35].

Each sector benefits from the adaptive capabilities of AI, but successful deployment hinges on fine-tuning algorithms to sector-specific data types, access models, and regulatory frameworks. This contextualization ensures that risk scoring, anomaly detection, and control monitoring are both relevant and actionable.

8.2. Metrics for Evaluating AI Effectiveness in Risk Reduction

Evaluating the effectiveness of AI in third-party risk reduction requires a multi-dimensional metrics framework that captures not only technical performance but also operational and strategic impact. Quantitative indicators are essential for demonstrating ROI, compliance readiness, and cybersecurity posture improvements [36].

One critical metric is the Mean Time to Detect (MTTD) third-party anomalies. AI models significantly reduce MTTD by automating behavioral analysis and continuously ingesting telemetry. A shorter MTTD indicates improved visibility and faster incident escalation. Complementary to this is the Mean Time to Respond (MTTR), which gauges the efficiency of AI-triggered remediation workflows and automated policy enforcement [37].

Reduction in Vendor-Induced Incidents is another key metric. By comparing pre- and post-deployment incident frequencies, organizations can quantify how AI interventions decrease successful phishing attempts, credential abuse, or control violations originating from third parties.

False Positive Rate (FPR) and True Positive Rate (TPR) are also used to evaluate model precision and accuracy. High FPRs can overwhelm analysts, while low TPRs may allow threats to go undetected. Balancing these metrics is vital for operational efficiency.

Compliance KPIs such as SLA adherence, third-party audit completion rates, and automated policy enforcement rates help measure how well AI tools support regulatory alignment.

Figure 4 A KPI dashboard mock-up visualizing these metrics, including risk trends by vendor, anomaly heatmaps, and SLA violation alerts enabling C-suite and security teams to track AI efficacy across the vendor ecosystem [38]

8.3. Lessons Learned and Operational Insights from Deployment

Deploying AI for third-party risk management across sectors has yielded critical operational insights and lessons for organizations pursuing scalable, adaptive cybersecurity strategies. One consistent lesson is the importance of cross-functional collaboration between IT, procurement, legal, and compliance teams. Without shared data ownership and clearly defined responsibilities, AI systems may underperform due to incomplete context or siloed implementation [39].

Another insight is that contextual training data dramatically improves model accuracy. Sector-specific models that incorporate localized behaviors, asset criticality, and compliance requirements outperform generic solutions. For example, in the energy sector, including SCADA telemetry and contractor access patterns produced more relevant alerts than models trained only on enterprise IT logs [40].

Integration planning also emerged as a key success factor. Organizations that treated AI implementation as an ecosystem enhancement interfacing with GRC, SIEM, and identity management tools saw greater efficiency gains and faster incident containment. Poor integration, by contrast, resulted in alert overload and fragmented risk visibility.

Additionally, AI tools are most effective when coupled with human-in-the-loop oversight. Analysts are essential for refining model outputs, contextualizing anomalies, and adjusting decision thresholds. AI augments human expertise but does not replace strategic judgment.

Lastly, early adopters noted that regulatory engagement improved when AI platforms were equipped with audit-friendly features like explainable models, traceable logs, and compliance dashboards—helping organizations demonstrate due diligence in third-party risk management and build trust with regulators [41].

9. Limitations, challenges, and ethical considerations

9.1. Data Privacy and AI Transparency Challenges

As organizations deploy AI-driven platforms for third-party cyber risk management, data privacy and transparency emerge as critical challenges. These systems rely heavily on continuous data ingestion, including logs, network activity, vendor credentials, and even contract metadata. When handling personally identifiable information (PII) or sensitive business data, privacy risks increase particularly in jurisdictions governed by strict frameworks like GDPR and CCPA [35].

A major concern lies in the opacity of AI models. Complex algorithms—especially deep neural networks may make high-impact risk predictions without clear, interpretable logic, which conflicts with legal requirements for algorithmic explainability and individual rights to contest automated decisions. The lack of transparency also makes it difficult to audit or validate decisions made by AI about third-party vendors, which can erode stakeholder trust [36].

Additionally, organizations face difficulty in managing data minimization and purpose limitation principles when AI systems continuously collect telemetry beyond originally intended scopes. These concerns underscore the need for integrating privacy-by-design principles and explainable AI (XAI) techniques, such as SHAP or LIME, into vendor risk analytics workflows [37].

Balancing security insights with ethical data governance is essential for compliant, responsible use of AI in third-party ecosystems ensuring platforms enhance resilience without compromising privacy.

9.2. Adversarial Attacks and Model Robustness Risks

AI models used in third-party cyber risk platforms are increasingly susceptible to adversarial attacks, which exploit model weaknesses through manipulated inputs to produce misleading outcomes. In the context of vendor surveillance, attackers may inject specially crafted data such as disguised access logs or benign-looking API activity to bypass anomaly detection systems [38].

Adversarial machine learning techniques, such as evasion attacks and data poisoning, pose a direct threat to model **robustness**. In evasion attacks, threat actors subtly alter input features to fall within acceptable thresholds, deceiving the AI into classifying malicious behavior as legitimate. Data poisoning involves corrupting training data to misguide the model during learning, undermining predictive accuracy and increasing false negatives [39].

These vulnerabilities not only reduce detection effectiveness but can also erode trust in automated systems if unaddressed. Organizations deploying AI tools must therefore implement **robustness testing**, such as adversarial training, input validation, and simulation of edge-case scenarios, to strengthen model resilience [40].

Additionally, integrating human review checkpoints and anomaly escalation layers can mitigate the operational risk of overreliance on fully autonomous systems. As cyber adversaries evolve, defending the AI itself becomes as critical as using it to defend infrastructure.

9.3. Ethical Use of AI in Vendor Surveillance

While AI tools significantly enhance third-party risk visibility, they also raise important concerns about the ethical boundaries of surveillance, especially when applied to vendors, contractors, and small service providers. Unlike internal employee monitoring, vendor surveillance often lacks clearly defined consent frameworks and may encroach upon contractual or jurisdictional boundaries [41].

The ethical dilemma intensifies when AI is used to analyze behavioral patterns, geolocation data, or communication logs from third-party systems. Without mutual agreement or transparent disclosure, such practices can be seen as intrusive or disproportionate particularly when vendors are smaller entities with limited capacity to challenge terms or practices [42].

Additionally, AI-driven surveillance can perpetuate bias if models flag specific vendors more frequently based on geographic or sectoral factors embedded in the training data. This could lead to discriminatory scrutiny, uneven risk scoring, or exclusion from contracts raising questions about fairness and due process in automated vendor management systems [43].

To mitigate these risks, organizations must adopt AI ethics charters, involve legal and compliance teams in surveillance design, and offer vendors avenues for recourse or clarification. Embedding fairness, proportionality, and transparency into AI governance ensures that the pursuit of cyber resilience does not come at the expense of ethical integrity.

10. Future trends in AI-driven third-party risk management

10.1. Predictive Governance and Digital Supply Chain Twins

As third-party cyber risks grow more complex, the concept of predictive governance leveraging AI to foresee disruptions, compliance lapses, or breaches before they occur is gaining traction. A cornerstone of this approach is the implementation of digital supply chain twins, virtual replicas of real-world vendor ecosystems that simulate operational, technical, and geopolitical risk scenarios in real time [38].

These twins aggregate data from internal telemetry, vendor portals, and external threat intelligence feeds to create a dynamic map of third-party interdependencies. By running simulations through AI algorithms, organizations can model what-if scenarios, such as how a regulatory failure or infrastructure outage at a key vendor would cascade through the enterprise. This forward-looking analysis supports proactive mitigation, SLA renegotiation, and budget allocation for resilience upgrades [39].

Digital twins also enhance incident planning and resource optimization. For example, if a supplier in a high-risk zone shows increased latency and API error rates, the twin can forecast potential data availability issues or service-level breaches. This level of continuous, AI-enabled foresight strengthens vendor governance and makes supply chain oversight more adaptive to real-world changes [40].

By aligning predictive governance with digital simulation, enterprises create an integrated framework that supports early intervention, resource efficiency, and resilient procurement strategies—moving risk management from reactive defense to proactive orchestration.

10.2. Federated Learning and Privacy-Preserving Risk Modeling

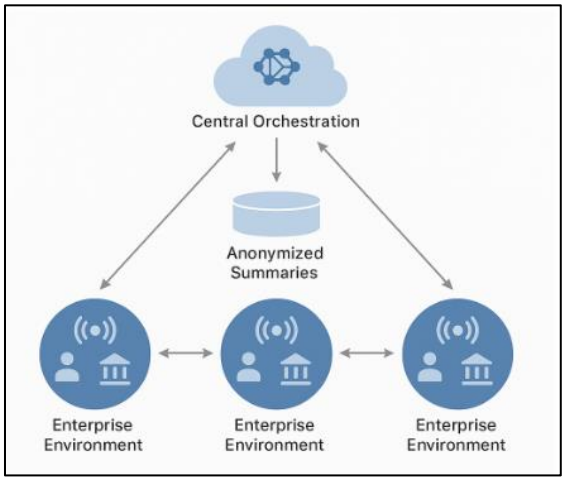

Figure 5 A conceptual model where AI agents embedded across enterprise networks collaboratively contribute to decentralized model updates. Risk signals such as access violations or abnormal vendor behavior are processed locally and anonymized before contributing to broader risk scoring engines

Federated learning (FL) represents a groundbreaking shift in AI-based third-party risk modeling by enabling collaborative intelligence without data centralization. In contrast to traditional models that aggregate sensitive vendor data into a central repository, FL allows individual organizations to train models locally on their own datasets while sharing only model parameters with a central orchestrator [41].

This decentralized approach significantly enhances privacy and data sovereignty, especially in regulated sectors like healthcare and finance, where cross-border data transfers may violate jurisdictional laws. With federated learning, risk models can learn from a broad network of enterprise environments each contributing to model robustness without exposing raw telemetry or contract details [42].

In a federated ecosystem, updates from each organization are aggregated using secure multiparty computation or homomorphic encryption techniques, preserving confidentiality throughout the training cycle. The resulting global model reflects diverse third-party risk profiles, enabling more generalizable and equitable threat detection across varied supply chain contexts [43].

By integrating FL into cyber risk platforms, organizations achieve a rare trifecta: enhanced model accuracy, compliance with privacy regulations, and broader insight into evolving threats without compromising proprietary or sensitive information.

10.3. Integration with Blockchain for Immutable Compliance Records

The integration of blockchain technology with AI-based third-party risk management introduces a new paradigm in compliance accountability. Blockchain's inherent immutability and transparency make it an ideal mechanism for recording vendor compliance actions, audit logs, and risk events in a tamper-proof ledger [44].

Each step in the vendor lifecycle from onboarding and contract approval to anomaly detection and remediation can be encoded as a timestamped transaction. Smart contracts can automate compliance workflows by triggering specific actions (e.g., access suspension or audit initiation) when predefined risk thresholds are breached. This real-time enforcement fosters trust among stakeholders while ensuring verifiable audit trails for regulators [45].

For enterprises operating across jurisdictions, blockchain provides a single source of truth that satisfies cross-border compliance and internal governance requirements. Moreover, vendors themselves can access these records to dispute risk assessments or demonstrate remediation progress, promoting transparency and fairness.

When paired with AI, blockchain ensures that risk scoring, model decisions, and corrective actions are not only accurate but also accountable and traceable. This synergy strengthens resilience and governance in increasingly decentralized, high-risk digital ecosystems bridging the gap between automation, privacy, and institutional integrity [46].

11. Conclusion

11.1. Summary of Key Findings

This article has explored the evolution of AI-powered systems in managing third-party cyber risks within complex, interconnected digital ecosystems. It examined the limitations of static assessments and highlighted how advanced machine learning models such as Random Forest, XGBoost, and neural networks enhance real-time threat detection and dynamic risk profiling. Sector-specific applications in healthcare, finance, and energy demonstrated AI's contextual versatility, while integration with GRC and SIEM platforms facilitated operational interoperability. Key challenges, including data privacy, adversarial attacks, and ethical concerns, were addressed alongside emerging innovations like federated learning and blockchain for decentralized compliance. Metrics such as MTTD, MTTR, and risk score variance were identified as essential for evaluating AI effectiveness. Through predictive governance, continuous control monitoring, and digital supply chain twins, AI platforms are reshaping vendor risk oversight from reactive to proactive. Collectively, these findings underscore the transformative potential of AI in fortifying enterprise resilience and aligning cybersecurity strategies with evolving regulatory and operational realities.

11.2. Strategic Recommendations for Implementation

To effectively implement AI-driven third-party cyber risk platforms, organizations should begin by aligning internal stakeholders IT, legal, procurement, and compliance under a unified governance framework. They must prioritize data quality and interoperability by integrating AI tools with existing GRC, SIEM, and IAM systems. Investment in explainable AI is crucial for ensuring regulatory compliance and stakeholder trust, especially in highly regulated sectors. Model

training should incorporate diverse, context-rich data sources, with special attention to bias mitigation and sector-specific calibration. Organizations should adopt a layered defense strategy that blends real-time anomaly detection with long-term risk prediction and incorporates human oversight at key decision points. Regular audits of AI decision logs and integration of adversarial testing will help validate performance and robustness. Lastly, piloting federated learning and blockchain components can elevate privacy and accountability while scaling risk intelligence across diverse vendor networks. A phased, modular deployment approach will ensure measurable impact and operational scalability.

11.3. Final Remarks on the Role of AI in Cyber Risk Governance

AI is no longer a theoretical add-on in cybersecurity it is a foundational component of modern risk governance. Its ability to ingest, process, and interpret complex data at scale makes it uniquely capable of addressing the evolving threat landscape posed by third-party relationships. From dynamic risk scoring to predictive governance and immutable compliance, AI offers a multidimensional toolkit for securing digital ecosystems. As enterprises grow increasingly interconnected and exposed, leveraging AI responsibly and strategically will be key to building adaptive, transparent, and resilient cyber risk management frameworks for the future.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Johnson B. Artificial Intelligence and Cybersecurity in Banking Sector: Opportunities and Risks.

[2] Ekolama SM, Ebregbe D. Application of Artificial Intelligence (AI) Model to Mitigate Security threats of Internet of Things (IoT): A Review.

[3] Adepoju Adekola George, Adepoju Daniel Adeyemi. Biomarker discovery in clinical biology enhances early disease detection, prognosis, and personalized treatment strategies. Department of Health Informatics, Indiana University Indianapolis, Indiana, USA; 2024. doi: https://doi.org/10.5281/zenodo.15244690

[4] Hussain D, Hajjar L. Cybersecurity and Big Data Analytics: Strategies for Securing Business Intelligence in the Digital Era.

[5] George S. Artificial Intelligence-Powered Cybersecurity: The Future of How Threats Are Detected and Responded. InLeveraging Large Language Models for Quantum-Aware Cybersecurity 2025 (pp. 247-276). IGI Global Scientific Publishing.

[6] Chukwunweike Joseph, Salaudeen Habeeb Dolapo. Advanced Computational Methods for Optimizing Mechanical Systems in Modern Engineering Management Practices. International Journal of Research Publication and Reviews. 2025 Mar;6(3):8533-8548. Available from: https://ijrpr.com/uploads/V6ISSUE3/IJRPR40901.pdf

[7] Goffer MA, Uddin MS, Hasan SN, Barikdar CR, Hassan J, Das N, Chakraborty P, Hasan R. AI-Enhanced Cyber Threat Detection and Response Advancing National Security in Critical Infrastructure. Journal of Posthumanism. 2025 Apr 17;5(3):1667-89.

[8] Radanliev P, De Roure D, Maple C, Nurse JR, Nicolescu R, Ani U. AI security and cyber risk in IoT systems. Frontiers in Big Data. 2024 Oct 10;7:1402745.

[9] Tanikonda A, Pandey BK, Peddinti SR, Katragadda SR. Advanced AI-Driven Cybersecurity Solutions for Proactive Threat Detection and Response in Complex Ecosystems. Journal of Science & Technology. 2022 Jan;3(1).

[10] Ugwueze VU, Chukwunweike JN. Continuous integration and deployment strategies for streamlined DevOps in software engineering and application delivery. Int J Comput Appl Technol Res. 2024;14(1):1–24. doi:10.7753/IJCATR1401.1001.

[11] Battu GG. AI-Driven Data Analytics in Custody Services: Enhanced Reporting, Compliance, and Risk Management.

[12] Shaffi SM, Vengathattil S, Sidhick JN, Vijayan R. AI-Driven Security in Cloud Computing: Enhancing Threat Detection, Automated Response, and Cyber Resilience. arXiv preprint arXiv:2505.03945. 2025 May 6.

[13] Ejedegba Emmanuel. Innovative solutions for food security and energy transition through sustainable fertilizer production techniques. World Journal of Advanced Research and Reviews. 2024 Dec;24(3):1679–1695. Available from: https://doi.org/10.30574/wjarr.2024.24.3.3877

[14] Rangaraju S. Ai sentry: Reinventing cybersecurity through intelligent threat detection. EPH-International Journal of Science And Engineering. 2023 Dec 1;9(3):30-5.

[15] Olutimehin AT. Advancing cloud security in digital finance: AI-driven threat detection, cryptographic solutions, and privacy challenges. Cryptographic Solutions, and Privacy Challenges (February 13, 2025). 2025 Feb 13.

[16] Adegboye O. Integrating renewable energy in battery gigafactory operations: Techno-economic analysis of net-zero manufacturing in emerging markets. World J Adv Res Rev. 2023;20(02):1544–1562. doi: https://doi.org/10.30574/wjarr.2023.20.2.2170.

[17] Kayode-Ajala O. Applications of Cyber Threat Intelligence (CTI) in financial institutions and challenges in its adoption. Applied Research in Artificial Intelligence and Cloud Computing. 2023 Aug 4;6(8):1-21.

[18] Kolhar A. Future Trends and Innovation in Machine Intelligence for Cyber Risk Management. InMachine Intelligence Applications in Cyber-Risk Management 2025 (pp. 415-438). IGI Global Scientific Publishing.

[19] Paul F. AI-Powered Threat Detection in Hybrid and Multi-Cloud Environments: Overcoming Security Challenges.

[20] Emmanuel Ochuko Ejedegba 'INTEGRATED STRATEGIES FOR ENHANCING GLOBAL FOOD SECURITY AMID SHIFTING ENERGY TRANSITION CHALLENGES', International Journal of Engineering Technology Research & Management (ijetrm), (2024) 08(12). doi: 10.5281/zenodo.14502251,

[21] Chinta PC, Jha KM, Velaga V, Moore C, Routhu K, SADARAM G. Harnessing Big Data and AI-Driven ERP Systems to Enhance Cybersecurity Resilience in Real-Time Threat Environments. Available at SSRN 5151788. 2024.

[22] Ake A. Enhancing US energy sector performance through advanced data-driven analytical frameworks. Int J Res Publ Rev. 2024 Dec;5(12):3336-56. Available from: https://doi.org/10.55248/gengpi.5.1224.250111

[23] Jeyachandran P. Implementing AI-Driven Strategies for First-and Third-Party Fraud Mitigation. Available at SSRN 5076791. 2024 Jul 3.

[24] Kshetri N. Transforming cybersecurity with agentic AI to combat emerging cyber threats. Telecommunications Policy. 2025 Apr 22:102976.

[25] Enemosah A, Chukwunweike J. Next-Generation SCADA Architectures for Enhanced Field Automation and Real-Time Remote Control in Oil and Gas Fields. Int J Comput Appl Technol Res. 2022;11(12):514–29. doi:10.7753/IJCATR1112.1018.

[26] Chukwunweike J, Lawal OA, Arogundade JB, Alade B. Navigating ethical challenges of explainable AI in autonomous systems. International Journal of Science and Research Archive. 2024;13(1):1807–19. doi:10.30574/ijsra.2024.13.1.1872. Available from: https://doi.org/10.30574/ijsra.2024.13.1.1872.

[27] Paul F, Al-Farsi Y. Automating Compliance Management with AI-Powered Risk Assessment Models.

[28] Olanrewaju AG, Ajayi AO, Pacheco OI, Dada AO, Adeyinka AA. AI-driven adaptive asset allocation: A machine learning approach to dynamic portfolio optimization in volatile financial markets. Int J Res Finance Manag. 2025;8(1):320-32. Available from: https://www.doi.org/10.33545/26175754.2025.v8.i1d.451

[29] Ayodele OF, Adelaja AO. Advancing Cybersecurity Governance: Adaptive Resilience and Strategic Third-Party Risk Management in Financial Services.

[30] Ejedegba Emmanuel Ochuko. Synergizing fertilizer innovation and renewable energy for improved food security and climate resilience. Global Environmental Nexus and Green Policy Initiatives. 2024 Dec;5(12):1–12. Available from: https://doi.org/10.55248/gengpi.5.1224.3554

[31] Salako AO, Fabuyi JA, Aideyan NT, Selesi-Aina O, Dapo-Oyewole DL, Olaniyi OO. Advancing information governance in AI-driven cloud ecosystem: Strategies for enhancing data security and meeting regulatory compliance. Asian Journal of Research in Computer Science. 2024 Dec 7;17(12):66-88.

[32] Olanrewaju AG. Artificial Intelligence in Financial Markets: Optimizing Risk Management, Portfolio Allocation, and Algorithmic Trading. Int J Res Publ Rev. 2025 Mar;6(3):8855-70. Available from: https://doi.org/10.55248/gengpi.6.0325.12185

[33] Ok E. Addressing Security Challenges in AI-Driven Cloud Platforms: Risks and Mitigation Strategies.

[34] Adegboye Omotayo Abayomi. Development of a pollution index for ports. Int J Sci Res Arch. 2021;2(1):233–258. Available from: https://doi.org/10.30574/ijsra.2021.2.1.0017

[35] Dhruvitkumar VT. Enhancing data security and regulatory compliance in AI-driven cloud ecosystems: Strategies for advanced information governance.

[36] Adegboye Omotayo, Arowosegbe Oluwakemi Betty, Prosper Olisedeme. AI Optimized Supply Chain Mapping for Green Energy Storage Systems: Predictive Risk Modeling Under Geopolitical and Climate Shocks 2024. International Journal of Advance Research Publication and Reviews. 2024 Dec;1(4):63-86. Available from: https://ijarpr.com/uploads/V1ISSUE4/IJARPR0206.pdf

[37] Jaggi K. Advancing Cybersecurity Strategies: Balancing Threat Detection, Compliance, and Resilient Architectures. Compliance, and Resilient Architectures (February 04, 2025). 2025 Feb 4.

[38] Ejedegba Emmanuel Ochuko. Advancing green energy transitions with eco-friendly fertilizer solutions supporting agricultural sustainability. International Research Journal of Modernization in Engineering, Technology and Science. 2024 Dec;6(12):1970. Available from: https://www.doi.org/10.56726/IRJMETS65313

[39] Ofili BT, Obasuyi OT, Osaruwenese E. Threat intelligence and predictive analytics in USA cloud security: mitigating AI-driven cyber threats. Int J Eng Technol Res Manag. 2024 Nov;8(11):631.

[40] Mbah GO, Evelyn AN. AI-powered cybersecurity: Strategic approaches to mitigate risk and safeguard data privacy.

[41] Ofili BT, Erhabor EO, Obasuyi OT. Enhancing Federal Cloud Security with AI: Zero Trust, Threat Intelligence, and CISA Compliance. World Journal of Advanced Research and Review. 2025.

[42] Paul F. The Future of Cloud Security: AI-Powered Predictive Analytics for Proactive Threat Management.

[43] Adeyeye OJ, Akanbi I, Emeteveke I, Emehin O. Leveraging secured AI-driven data analytics for cybersecurity: Safeguarding information and enhancing threat detection. International Journal of Research and Publication and Reviews. 2024;5(10):3208-23.

[44] Clement M. Ethical and Legal Risks in AI-Powered Compliance Systems.

[45] Kaul D. AI-Powered Autonomous Compliance Management for Multi-Region Data Governance in Cloud Deployments. Journal of Current Science and Research Review. 2024 Dec 19;2(03):82-98.

[46] Mohammed A. Elevating Cybersecurity Audits: How AI is Shaping Compliance and Threat Detection. Aitoz Multidisciplinary Review. 2023 Apr 6;2(1):35-43.