University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)

ISSN: 2714-3627

A Journal of the Department of Computer Science, University of Ibadan, Ibadan, Nigeria

Volume 14 No. 1, June, 2025

journals.ui.edu.ng/uijslictr
http://uijslictr.org.ng/
uijslictr@gmail.com

University of Ibadan
Journal of Science and Logics in ICT Research

Artificial Intelligence in Cybersecurity: A Comparative Review of Its Role across the Cyber Kill Chain

[1]Azeez N. A., [2]✉Malomo O. S., [3]Aaron D.S., [4]Ademoye A. A., [5]Okerinde O. M., [6]Otolehi U. D., and [7]Lukman O. O.

[1,2,3,4,5,7] *Department of Computer Sciences, University of Lagos*
[6]*FSDH Merchant Bank*
[1]nurayhn1@gmail.com, [2]*oluwatobis.malomo@gmail.com*, [3]*249074144@live.unilag.edu.ng*,
[4]*249074057@live.unilag.edu.ng*, [5]*okerindeo@gmail.com*, [6]*uzomaotolehi@yahoo.com*, [7]ololaitan@gmail.com

Abstract
The adoption of Artificial Intelligence (AI) in diverse fields and the proliferation of interconnected devices have led to the emergence of highly sophisticated cyberattacks today. This new reality has compelled organisations to align their security policies by adopting cybersecurity frameworks. These frameworks provide organisations with models and methods for effectively managing digital security risks by promptly detecting and mitigating cyberattacks. The Cyber Kill Chain (CKC) decomposes cyberattacks into 7 phases, which cyber defenders can rely on when developing threat-informed strategies to mitigate cyberattacks. This paper presents a comprehensive overview of the CKC, highlighting the role Artificial Intelligence plays across each phase in terms of offensive and defensive cybersecurity operations. A comparative analysis of 3 cybersecurity frameworks, with justifications for each, was also examined. Drawing on real-world case studies and recent literature, this study further highlights current challenges with the fusion of AI into cybersecurity operations, ranging from data privacy, adversarial attacks, and AI explainability. The review concludes by advocating for the adaptation of dynamic, AI-driven modelling frameworks that better align with the rapidly evolving cyber threat landscape.

Keywords: *Cybersecurity; Artificial Intelligence; Cyber Kill Chain; Threat Modelling; AI-Based Threat Detection; Cyber Threat Intelligence*

1. Introduction

The digital age has brought about a proliferation of interconnected systems, leading to the surge of data-driven operations across diverse fields. This reality has further led to the increased frequency and complexity of cyberattacks. Today, adversaries or cyber attackers are equipped with more sophisticated tools and approaches, including social engineering, zero-day, and multi-vector attacks, which traditional security tools struggle to detect and contain in real-time. The evolving threat landscape, therefore, requires that smarter and more adaptive defense strategies be implemented [1]. The knowledge discovery process involves various selection steps which help in the efficient extraction of useful data from databases.

Organisations and cybersecurity teams are leveraging AI and Machine Learning (ML) tools to develop adaptive and scalable cybersecurity strategies, systems and architectures, given that these technologies are capable of scaling cybersecurity operations, especially in the areas of predictive analytics, anomaly detection, and automated threat response [2].

Cybersecurity frameworks provide cyber operators with methodologies and models to detect and mitigate cyber threats. Lockheed Martin developed the Cyber Kill Chain (CKC) framework in 2011 to equip organisations with a structured framework to understand and thwart cyberattacks [3]. The CKC improves situational awareness and helps to actively facilitate the development of defense strategies against future cyberattacks.

Azeez N. A., Malomo O. S., Otolehi U. D. and Lukman O. O. (2025). Artificial Intelligence in Cybersecurity: A Comparative Review of Its Role Across the Cyber Kill Chain. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)*, Vol. 14 No. 1, pp. 232 -245

This paper explores the influence of AI across both the offensive and defensive dimensions of the CKC framework, explores how the framework performs in comparison to alternative frameworks like the UKC, MITRE ATT&CK, and Diamond Model, evaluates the implications of integrating AI into cybersecurity operations, and highlights the emerging challenges and opportunities it presents for future cyber defense endeavours.

2. Literature Review

The continuous integration of AI into the cybersecurity landscape has significantly contributed to the growing sophistication of cyberattacks targeted at states, organisations, and individuals today. Hence, the foundational and traditional security mechanisms are increasingly being outpaced by attackers who are smarter and now better equipped to leverage emerging technologies and system vulnerabilities.

As digital infrastructures become increasingly interconnected and attackers adopt sophisticated Tactics, Techniques, and Procedures (TTPs) to carry out malicious activities, security experts appear to be accelerating the integration of AI into cybersecurity operations. The role of AI as a double-edged sword in cyber operations is now a widely discussed topic. While it strengthens cyber defenses through anomaly identification, intelligent threat detection, and automated response, it also serves as an important leverage for cyber attackers who utilise AI tools for automating cyberattacks, exploiting system vulnerabilities at scale, and evading detection [1].

The performance of cybersecurity frameworks has been enhanced by the introduction of AI during their adoption by organisations. This is equally true of the CKC framework, which has evolved, resulting in an enhanced level of dynamic intelligence, automation, adaptability, and responsiveness. However, this integration has also equipped cyber attackers with tools to carry out malicious activities across the kill chain in a more complex way.

Several studies have been conducted to examine the effectiveness of the CKC framework, given the evolution of threats and advancements in technology. According to Kazimierczak *et al.*, [4], the integration of AI within the CKC plays to the advantage of both cyber defenders and attackers as players on both sides of the divide can automate and optimise tasks at each phase. The findings of this study further underscore the continued relevance of the CKC framework while highlighting its limitations in modelling multi-vector attacks.

Caltagirone *et al.*, [5] identified the lack of relational depth needed for attacker attribution as a major challenge with the adoption of the CKC. The study suggested that organisations should consider integrating the Diamond Model within the CKC framework to develop a more dynamic and robust threat intelligence system.

Furthermore, the framework performs poorly when faced with sophisticated cloud-based attacks and insider threats, largely because of its linearity and tendency to be biased towards external threats [6]. To tackle this issue, it was also recommended that the CKC framework be augmented with iterative loops and behavioural indicators that more closely reflect real-world attack patterns.

2.1 Overview of Artificial Intelligence

Artificial Intelligence, or simply AI, is the theory and application of computer systems that perform tasks that would normally require human intelligence and interference [7]. These tasks include learning, reasoning, decision-making, and problem-solving, and are carried out by analysing data to identify patterns. According to Sharma *et al.*, [8], ML involves making algorithms learn by acquiring knowledge from previous experiences and is one branch of AI that has advanced greatly over the last 30 years. Deep Learning, Computer Vision, Natural Language Processing (NLP), and Generative AI (GenAI), among other fields, have also evolved significantly in recent years.

2.2. Types of Learning

Depending on the kind of project at hand, the following are types of learning one can deploy for a Machine Learning project:

1. Supervised Learning: Involves training an algorithm with a labelled dataset. The goal here is to predict the labels of unseen data – to generalise accurately.

2. Unsupervised Learning: It is used when the project involves drawing inferences from unlabelled datasets. It is best used for pattern recognition and predictive modelling.
3. Reinforcement Learning: This type of learning is used in robotics, whereby the reward and punishment approach is adopted in training AI agents.

2.2 AI in Cybersecurity: A Foundational Perspective

Predictive analytics helps cybersecurity teams across multiple organisations to anticipate threats using known behavioural patterns of an attacker; hence, they get to approach cyber threats from a position of strength (Sarker *et al.*, [9]). AI has become a critical enabler within the cybersecurity landscape, leading to the development of improved cybersecurity strategies in behavioural analytics and anomaly detection irrespective of the size of the datasets involved (Buczak and Guven [10]). The ability to detect known threats and zero-day threats will help organisations drastically reduce threat response time and mitigate potential attacks [17].

Artificial Intelligence supports the automation of cybersecurity operations through Security Orchestration, Automation, and Response (SOAR) systems. According to Mohamed [11], these systems are known to streamline cybersecurity tasks such as alert triaging, log analysis, and incident response. These advancements greatly ease the burdens on human analysts and greatly enhance the scalability of defense strategies. In terms of identity and access management, AI systems are used to power adaptive authentication systems, which make use of behavioural biometrics to provide a real-time response to suspicious login activities across security infrastructures [12].

3. Overview of Cybersecurity Frameworks

A cybersecurity framework is a model, a structured set of documents and methodologies that assist an organisation in managing digital security risks, identifying breach attempts, and mitigating cyber threats promptly. These frameworks allow cybersecurity teams to integrate policies, technological structures and controls into organisational cybersecurity strategies that align with operational guidance on emerging threat landscapes [13, 14]. In addition, cybersecurity frameworks assist in the critical task of creating a uniform standard for cybersecurity procedures across various sectors and industries. As compliance regulations become more stringent, cybersecurity frameworks support compliance governance, threat modelling, incident response and even prevention approaches. In a growing number of jurisdictions where organisations and states increase efforts to secure digital infrastructure, these frameworks aid in the organisation of defense strategies to support situational understanding by providing a systematic, consistent, reproducible, and replicable approach.

Cybersecurity frameworks can be categorised into two:

1. Threat Modelling and Operational Defense Frameworks: These frameworks attempt to describe, categorise, and defend an adversary activity. Examples are the Unified Kill Chain (UKC), MITRE ATT&CK, Diamond Model, and CKC.
2. Governance and Risk Management Frameworks: These are focused on system principles guiding the creation of an Information Security Management System (ISMS). They help integrate cybersecurity policies with the organisation's goals. Unlike the threat modelling frameworks, they focus on policy development, enterprise-wide risk management, and asset protection. Examples include NIST Cybersecurity Framework (CSF), ISO/IEC 27001, and Control Objectives for Information Technologies (COBIT).

Cybersecurity frameworks have continued to evolve, becoming more context-sensitive, adversary-aware, and dynamic, in response to the increasing complexity of cyber threats, particularly with the integration of AI.

3.1 The Cyber Kill Chain Framework

The Cyber Kill Chain is a conceptual framework designed to model the sequence of steps adversaries typically follow to execute a cyberattack successfully [3]. Inspired by traditional military kill chains, the CKC decomposes the cyberattack lifecycle into 7 linear stages. One of the most important strong points of this framework is its capacity to guide cyber defense strategies by promptly

identifying opportunities to detect, deny, disrupt, degrade, or deceive adversaries or attackers at each phase of an attack's lifecycle. The evolution of AI brought about a rapid change in cybersecurity operations at each stage of the CKC framework. In recent years, cybersecurity operations have improved significantly, especially in the aspects of threat intelligence, threat hunting, and incident response. CKC offers a structured lens through which defenders can understand attack progression and swiftly implement mitigation measures before adversaries can achieve the objectives of their cyberattack efforts. The linear nature of the framework has been heavily criticised in recent years as it fails to capture non-linear and recursive behaviours of modern threat actors, especially within the scope of Advanced Persistent Threats (APTs). This has prompted calls for more dynamic models such as the UKC, MITRE ATT&CK framework, and hybridisation of multiple models for effectiveness [15, 16].

Therefore, it is imperative to examine the CKC framework from a dynamic cybersecurity perspective, where AI serves both as a defensive tool and a potential threat vector on the cybersecurity spectrum.

3.1.1 Phases of the Cyber Kill Chain

The following are the stages of the CKC framework:

1. Reconnaissance: Offenders gather information to identify vulnerabilities in the target system
2. Weaponization: A tailored payload is created to exploit the vulnerabilities
3. Delivery: Malware is transmitted using phishing emails or infected websites
4. Exploitation: Triggers malicious code to exploit the system vulnerability
5. Installation: Malware is installed to ensure persistent access to the target system
6. Command and Control (C2): Communication is established with an external server
7. Actions on Objectives: The attacker achieves their end goal, exfiltrating data or disrupting operations.

CKC models a cyberattack as a series of sequential phases an adversary progresses through to achieve malicious objectives. These 7 phases present a structured approach for identifying, understanding, and mitigating threats at each step of an attack lifecycle.

Reconnaissance involves attackers gathering intelligence about the target system, network, or personnel using techniques such as open-source intelligence and social engineering. The stealthy nature of this stage of the CKC, especially when passive methods are adopted, represents a huge challenge for cyber defenders, as such activities often leave no traceable footprint [4]. However, the deployment of sophisticated network monitoring tools can help organisations detect anomalies such as repetitive scanning and/or unusual user behaviour. This, therefore, presents the organisation with an opportunity to promptly disrupt potential attacks at their inception.

In the *Weaponization* phase, adversaries create a malicious payload, mostly by combining a vulnerability exploit with a random, deliverable medium such as a script or document. This stage of the CKC typically occurs outside the defender's network, thus limiting visibility and making it a herculean task for traditional intrusion detection systems to respond [4].

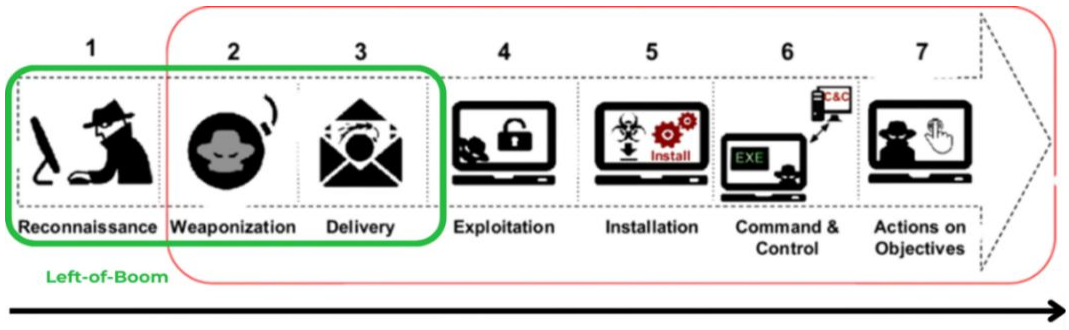

Figure 1: Phases of the CKC (Adapted from Dargahi *et al.*, [20])

However, this limitation notwithstanding, an understanding of prevalent malware construction techniques can help defenders anticipate the types of payloads being prepared and adjust their defenses promptly with the use of proactive threat intelligence.

The transmission of the weaponized payload to the target is carried out in the *Delivery Phase*, mostly via methods such as phishing emails, removable media or compromised websites. The CKC model proves valuable here, as this is one of the stages where organizations can exert a level of control. Using antivirus software, email filters, and network security gateways, organizations can intercept and quarantine malicious content before it reaches a host or an end user. However, the growing use of fileless malware and encrypted traffic further complicates threat detection efforts by organisations and individuals alike, as these methods are often known to allow threats to bypass traditional signature-based security solutions [18]. As illustrated in Figure 1, the left-of-boom (Reconnaissance, Weaponisation and Delivery phases) presents organisations with an opportunity to prevent cyberattacks by taking proactive measures instead of being reactive once the attacker has gained access to their infrastructure.

The *Exploitation* stage is where the attack is executed on the target machine by leveraging vulnerabilities in software, user behaviour, or system configurations. Attackers often exploit gaps when system patches are outdated. While it is possible to stop an attack at this stage, especially through endpoint protection tools and vulnerability management, the increasing prevalence of zero-days continues to render many conventional defenses ineffective [19]. Hence, the extremely narrow window between the discovery and exploitation phases requires a swift and automated response to thwart attacks effectively.

The next phase after a successful exploitation is the *Installation* phase, where the attacker places malware or backdoors that will enable persistent access to the victim's environment. This can be in the form of rootkits, keyloggers, or remote access tools. As a control measure, well-configured Endpoint Detection and Response (EDR) solutions can identify unauthorised changes to registry keys and system files. However, sophisticated malware is often obfuscated or embedded in legitimate processes, thereby allowing it to evade detection while maintaining persistence over time 19].

The *Command and Control*, known simply as the *C2 stage*, allows an attacker to manipulate compromised systems remotely, typically through covert channels such as DNS or HTTP tunnelling. With the CKC framework, cybersecurity officers can identify anomalous network behaviours, which often serve as Indicators of Compromise. More recently, adversaries have adapted by encrypting their communications or using trusted cloud services for C2. This action helps them blend their actions into legitimate network traffic thereby evading detection completely [4]. This challenge further highlights the limitations of static monitoring systems and underscores the need for behavioural analysis.

Finally, *Actions on Objectives* refers to the stage where an attacker achieves their objective, including data exfiltration, data destruction, or further lateral movement within the network. Once here, the attacker has already bypassed most defenses, hence response and containment efforts become more difficult and costly. The CKC model emphasises earlier stages of attack detection to prevent escalation. However, researchers have argued that this reactive orientation overlooks proactive defense strategies and downplays threats such as insider abuse and supply chain vulnerabilities that do not conform to the linear kill chain model [21].

Table 1 presents a comparative overview of each phase of the CKC framework, identifying the impact of AI across both offensive and defensive cybersecurity operations.

Table 1. Impact of AI across the phases of the CKC

CKC Phase	AI Impact	Offensive AI Techniques	Defensive AI Techniques
Reconnaissance	Automated, large-scale info-gathering	Web crawlers, NLP profiling	AI-based threat intelligence, anomaly detection
Weaponization	Tailored malware generation	GAN-based malware, evasion testing	Predictive patching, AI sandboxes
Delivery	Personalised and covert transmission	Deepfakes, spear-phishing automation	NLP email filters, AI-based SEG
Exploitation	Adaptive and intelligent exploitation	AI fuzzing, reinforcement learning	EDR, automated patching
Installation	Covert, polymorphic malware deployment	Dynamic malware, AI-coded droppers	AI HIDS, behavioural detection
Command & Control	Sophisticated and stealthy communication	Encrypted AI C2, botnet orchestration	AI network monitoring, DPI
Actions on Objectives	Strategic and prioritised attacks	Smart data exfiltration, timed extraction	AI DLP, UEBA systems

3.1.2 Strengths and Weaknesses of the CKC

As the cybersecurity landscape evolves with the introduction of more sophisticated adversarial techniques, it is important to understand the strengths and weaknesses of the CKC framework.

Strengths of the CKC

1. Early Detection and Prevention: The proactive approach of the CKC framework aligns greatly with intelligence-driven defense strategies [21]. With a strong emphasis on left-of-boom phases (reconnaissance, weaponisation, and delivery), the framework enables security teams to stop cyberattacks as they arise, long before attackers can carry out their malicious objectives.
2. Sequential and Structured Understanding of Attacks: By projecting cyberattacks in a linear, sequential manner, the CKC framework allows organisations to identify and mitigate even the most complex threats before they can affect cybersecurity infrastructures [3].
3. Incident Response and Forensics: The post-attack analysis offered by the framework helps security analysts to reconstruct cyberattacks, thereby gaining insights into compromised units [6].
4. Threat Intelligence: The framework enables cybersecurity teams across organisations to map cyberattacks to the threat intelligence obtained from well-known threat incidents [22]. This enhances situational awareness and the contextual relevance of Indicators of Compromise (IOCs).
5. Foundation for Advanced Defensive Models: The CKC framework has inspired the development of complementary and improved models, including the AIVC (Diamond Model), MITRE ATT&CK, and UKC frameworks [23]. The much-improved frameworks are built upon the foundational principles of the CKC by addressing its limitations.

Limitations of the CKC

1. Linear Bias and Lack of Flexibility: A major criticism of the CKC framework is its assumption of a strictly linear progression of attack stages. In real-world scenarios, attackers can skip stages, loopback, or simultaneously perform multiple phases [24]. A typical example of this is when malware downloaded via phishing includes both delivery and exploitation, thereby undermining the model's granularity.
2. Ineffectiveness against Insider Threats: The CKC is primarily designed to detect external threats that require entry into the system; hence, it is known to struggle to model insider threats where malicious activity begins post-authentication and may bypass several early stages entirely [25]. This has led to calls for supplemental frameworks like the Insider Threat Matrix of the MITRE ATT&CK.

3. Limited Applicability in Cloud and Decentralised Environments: The rise of cloud-native applications, SaaS platforms and microservices has continued to render CKC's perimeter-based assumptions outdated. In recent years, attacks have often targeted identity and access tokens or exploited third-party APIs. These entry points are not succinctly represented in the original chain [6].
4. Overemphasis on Malware-Centric Attacks: The CKC framework is suited mainly to malware-driven, targeted APT-style campaigns, making it less applicable to misconfiguration exploitation, social engineering, or supply-chain compromises that are known not to follow the traditional mechanisms of payload delivery [3].
5. Limited Support for Real-Time Defense Adaptation: In contrast to the increasingly dynamic adversary behaviour and AI-driven attacks, it lacks the flexibility to accommodate evolving tactics. The CKC framework does not support real-time threat scoring and adaptive feedback, which is increasingly vital for modern cybersecurity defense strategies [4].

3.2 The Course of Action Matrix

The COA Matrix is tailored to evaluate different defensive strategies in terms of risk, operational impact, deployment time, feasibility, cost, and overall impact on an organisation or state. Actions are mapped to specific threat behavioural responses or tactics, primarily drawn from cybersecurity frameworks, to aid decision-making in selecting a countermeasure needed to address cyber threats [3, 22].

The matrix complements the CKC framework by assigning targeted defensive strategies against adversarial actions at each phase of the sequence. It also provides structured response profiles for cyberattacks by giving room for justification of control allocation. This becomes increasingly relevant when working under resource-constrained and time-sensitive conditions [22, 29].

3.2.1 The COA Matrix and the CKC Framework

Each stage of the CKC provides security response teams with the opportunity to detect, disrupt, or contain cyber threats; hence, integrating the COA Matrix across the framework will help in defining and selecting the best suitable countermeasures at each phase [3].

The COA Matrix in Table 2 describes 6 operational responses security teams can take at each phase of the CKC framework, alongside corresponding defensive measures. For instance, since the Host Intrusion Detection System (HIDS) can detect exploitation attempts passively, they are implemented at the exploitation phase. To block adversaries from gaining access to a cybersecurity infrastructure, the security team must ensure timely system patching. The matrix captures the wide range of defensive tools available to security teams, including Network Intrusion Detection Systems (NIDS), firewalls, ACLs, and even system hardening techniques like audit logging, which are considered traditional. More importantly, it also emphasises the critical role of human vigilance, acknowledging that alert users can be instrumental in identifying and responding to suspicious activity.

3.3 Comparative Review of Threat Modelling Frameworks

The evolution of cyber threats, fuelled by the interflow of AI and cybersecurity, and the limitations of the CKC framework have led to the emergence of alternative frameworks. In the following subsection, these frameworks are reviewed to understand their differences, complementary strengths, and applicability in AI-enhanced cybersecurity environments.

3.3.1 The MITRE Adversarial Tactics, Techniques, and Common Knowledge

The MITRE ATT&CK organises attack behaviours into a matrix of tactics (goals) and techniques (methods) in a detailed manner. It is continuously updated with observations from the real world to ensure its applicability across various platforms, including cloud, mobile, Linux, and Windows systems [21]. The behavioural taxonomy of the framework enhances detection systems by mapping AI techniques to telemetry data [6]). The constant training of AI models with the ATT&CK datasets significantly enhances threat correlation and detection.

Table 2: COA Matrix of Operational Responses Across the CKC Phases

Phase	Detect	Deny	Disrupt	Degrade	Deceive	Destroy
Reconnaissance	Web Analytics	Firewall ACL				
Weaponization	NIDS	NIPS				
Delivery	Vigilant User	Proxy Filter	In-line AV	Queuing		
Exploitation	HIDS	Patch	DEP			
Installation	HIDS	"chroot" jail	AV			
C2	NIDS	Firewall ACL	NIPS	Tarpit	DNS redirect	
Actions on Objectives	Audit log			Quality of Service	Honeypot	

Strengths of the MITRE ATT&CK

1. Granular catalogue of adversarial behaviours
2. Threat intelligence can be mapped to red teaming and other SOC operations
3. Improved alignment with detection rules, SIEMs, and EDR tools

Limitations of the MITRE ATT&CK

1. The complexity of the framework can overwhelm security teams
2. Not designed for high-level strategic communication or training

3.3.2 The Unified Kill Chain (UKC)

The UKC, proposed by Hutchins and collaborators, expands the CKC framework by adding more phases from both real-world campaigns and the ATT&CK framework to make up 18 unique phases. The framework categorises the cyberattack lifecycle into three primary stages: initial foothold, network propagation, and action on objectives [26, 27].

The detailed phase structure and expanded scope of the framework provide a rich basis for training AI models. For instance, Reinforcement Learning systems can simulate the behaviours of an attacker across the stages of the framework to test and optimise defense postures [4].

Strengths of the UKC

1. Integrates insights from both ATT&CK and the CKC frameworks
2. Captures privilege escalation and lateral movement more accurately
3. Better suited for modelling persistent and multi-phase cyber attacks

Limitations of the UKC

1. Increased complexity makes implementation and interpretation more difficult
2. Not widely adopted like CKC or ATT&CK

3.3.3 The Diamond Model of Intrusion Analysis

Introduced in 2013, the framework helps to understand cyber intrusions through four interrelated features: *adversary*, *capability*, *infrastructure*, and *victim* [5]. Rather than the sequence of attack, it focuses on relationships and causality. It is well-suited for AI-powered threat correlation and clustering by analysing datasets to infer relationships between adversaries, infrastructure, and techniques [28].

Strengths of the Diamond Model

1. The framework supports advanced intelligence analysis and threat attribution
2. Uses observable characteristics to perform hypothesis-driven investigations
3. Emphasises adversary intent and infrastructure, and aids campaign tracking

Limitations of the Diamond Model

1. The framework is less focused on specific stages of an attack or technical mitigations
2. Not ideal for response workflows and SOC operations

Table 3 compares the four cybersecurity frameworks in terms of their unique strengths, weaknesses and relevance with the prevalence of AI in cybersecurity today. This further underlines the effectiveness of hybridised frameworks in mitigating evolving cyber threats. While CKC remains valuable for structured incident response, MITRE ATT&CK and the UKC offer much deeper operational insight, especially in dynamic, AI-enhanced threat environments

Table 3: Comparative Overview of the Cybersecurity Frameworks

Framework	Focus	Strengths	Weaknesses	AI Relevance
CKC	Linear attack progression	Simplicity; good for defense planning	Outdated for clouds and insider threats	Useful for mapping AI to known attack stages
MITRE ATT&CK	Techniques & tactics matrix	Granularity; updated threat behaviour	Complexity Less strategic	Ideal for training AI models in detection & analysis
Unified Kill Chain	Extended attack lifecycle	Full campaign coverage; includes lateral movement	Complex Less adopted	Strong AI mapping across persistent and adaptive threats
Diamond Model	Adversary-infrastructure mapping	Adversary profiling Intelligence-driven	Less suited for defense tool alignment	Excellent for AI-driven attribution and clustering

The Diamond Model complements these by supporting strategic attribution and intelligence development. It is expedient for organisations seeking to build AI-driven security architectures to adopt a hybrid approach by leveraging CKC for detection stages, ATT&CK for technical defenses, and the Diamond Model for threat intelligence enrichment.

4. Applications Areas of AI in Cybersecurity

The important role AI plays in cybersecurity cannot be overemphasised. Similarly, it has contributed greatly to the enhancement of both the effectiveness and agility of actions across the stages of the CKC. This subsection explores the integral role AI plays across the different application areas in cybersecurity.

4.1AI in Offensive Cybersecurity Operations

The offensive application of AI in cyber operations has ushered in new dimensions of the threat landscape. Attackers now employ AI systems to develop polymorphic malware that can change its code to avoid detection, automate reconnaissance and vulnerability assessment, and conduct highly targeted phishing attacks [4].

Cyberattacks are much more effective and persistent today as adversaries adapt their tactics in real-time while making efforts to breach security systems. Social engineering has become much more sophisticated with Generative AI deepfakes making it difficult to distinguish between fake and real content [4, 24].

1. *Reconnaissance:* The initial phase of the CKC can be time-consuming as it involves information gathering. AI techniques are now used to automate Open-Source Intelligence (OSINT) collection across social media platforms, databases and websites [4]. By mining and synthesising information such as employee names, email addresses, and software stacks, among others, attackers can develop precise attack vectors to curate believable phishing messages [1].
2. *Exploitation:* Through a process known as fuzzing, adversaries input malformed data into software to uncover unknown vulnerabilities. This empowers adversaries to optimise test case generation and prioritise likely failure points, which drastically reduces the time needed to exploit weaknesses even in the most complex security systems [33]). In reverse engineering endeavours, attackers use AI models to understand and manipulate binary code, firmware, or proprietary protocols to find entry points into systems.
3. *Evasion:* AI has also become a critical enabler of evasion techniques, particularly in the use of Adversarial Machine Learning. Attackers craft inputs, such as modified malware samples or poisoned datasets that deliberately deceive AI-based security models like Intrusion Detection Systems (IDS) or malware classifiers. By altering features subtly, adversarial attackers can bypass defenses without triggering alerts before proceeding to exploit the vulnerabilities in the learning algorithms directly [34].
4. *Phishing and Malware Customisation:* With GPT-based models, adversaries can now generate personalised phishing

campaigns that mimic the writing styles of victims. Adversaries use deepfakes to impersonate people and executives of organisations to carry out scams such as the Business Email Compromise (BEC). Furthermore, AI-enhanced malware delivered to a target system can alter the system's configurations such that it delays execution to avoid sandbox detection [1].

4.2 AI in Defensive Cybersecurity Operations

The offensive application of AI in cyber operations has ushered in new dimensions of the threat landscape by enabling attackers to develop polymorphic malware that can change its code to avoid detection, automate reconnaissance and vulnerability assessment, and conduct highly targeted phishing attacks [4].

To effectively mitigate evolving cyber threats, the following systems have been adopted:

1. *Anomaly Detection (IDS/IPS):* Cybersecurity teams in organisations and states are beginning to discard signature-based Intrusion Detection/Prevention Systems for more dynamic and current alternatives that do not rely on predefined patterns. Supervised Learning and Unsupervised Learning ML algorithms are now employed to accurately detect anomalies in network traffic. Considering that there is a continuous flow of traffic within the network, Intrusion Detection Systems are developed using the Online Learning ML paradigm, thereby always ensuring adaptability to novel threats [11].
2. *Malware Detection and Classification:* In the past, antivirus programs used known malware signatures to scan systems, which has proven to be ineffective in detecting masked cyberattacks. AI-based malware detection systems are trained on large datasets to generalise accurately by constantly training the models with new datasets obtained from continuous network monitoring, as described in Figure 2. To build adaptive systems, security teams can develop hybrid models by combining static, dynamic, and contextual analysis or utilise Deep Learning architectures, such as CNN and RNNs that are known to be highly effective in detecting masked variations and zero-days in cyber threats [1].
3. *User and Entity Behaviour Analytics (UEBA):* This involves the use of statistical analysis and ML techniques to detect and analyse anomalies in both user and entity behaviour across systems and networks. By monitoring the behavioural patterns, the ML model learns about the network and promptly flags noticeable deviations such as compromised accounts, lateral movements, or unusual network traffic. UEBA drastically reduces false positives by constantly improving its network traffic classification. To enhance the threat visibility of a network, security teams are encouraged to deploy UEBA with the SIEM tool [35].
4. Automating Threat Intelligence and Incident Response: This helps security teams to obtain more context about cyber threats. NLP tools are now used for mining, analysing and extracting actionable intelligence from unstructured data. The intelligence obtained is used to correlate disparaging indicators of compromise (IOCs), monitor trends to predict threats, and prioritise alerts based on the context and severity of attacks. Furthermore, the automation of incident response minimises human interference by executing actions such as the isolation of compromised systems and IP addresses blocking [4].

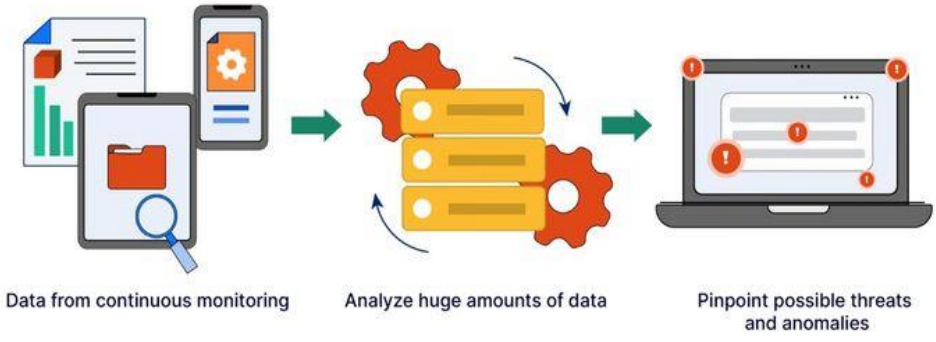

Figure 2: Illustration of the role of AI in cybersecurity defense operation (adapted from Astra [32])

4.3 Case-Based Incidents

Real-world cyber incidents highlight the growing role of AI across cybersecurity operations. This section presents key examples of high-profile attacks and the corresponding influence of AI techniques. These case studies exemplify the dual role of AI in amplifying attack effectiveness and in empowering timely detection and mitigation.

4.3.1 Stuxnet (2010)

Stuxnet, although predating the contemporary rise of AI, laid foundational strategies that are now being enhanced through AI. Its multi-stage attack mirrored the CKC, from reconnaissance to physical destruction. If executed today, AI would augment its evasion techniques using behavioural cloaking and optimise payload delivery through AI-powered reconnaissance of SCADA systems. Defensively, AI-driven anomaly detection tools could flag the abnormal operation of programmable logic controllers (PLCs) far earlier than traditional signature-based methods.

4.3.2 SolarWinds Attack (2020)

The SolarWinds supply chain attack demonstrated the potential scale of stealthy, long-term infiltration. AI could have played a role in both improving the attacker's ability to identify low-noise lateral movement and evasion tactics and in helping defenders detect anomalies in network behaviour and irregular data exfiltration. AI-enhanced threat intelligence platforms that cross-correlate system behaviours and code patterns across multiple clients may have reduced the detection time.

4.3.3 Microsoft Exchange Server Exploits (2021)

This attack campaign exploited zero-day vulnerabilities in Microsoft Exchange servers. AI-based fuzzing tools may have accelerated the exploitation process by identifying vulnerabilities faster than traditional methods. On the defensive end, AI-enabled Endpoint Detection and Response (EDR) platforms have been critical in post-exploit detection, helping security teams recognise behavioural anomalies and isolate affected systems before further lateral movement.

4.3.4 Colonial Pipeline Ransomware Attack (2021)

The 2021 Colonial Pipeline incident revealed weaknesses in ransomware preparedness. While AI may not have been used directly by attackers, the use of AI-enabled phishing kits to automate the targeting process is a growing trend. In contrast, AI-based email filters, UEBA tools, and ransomware behaviour classifiers could have been used defensively to detect suspicious activity in the early phases of the kill chain.

5. Conclusion

This paper presents an overview of the CKC framework, exploring its structure, areas of application, and integration with AI by critically analysing the double-edged role it plays across each stage of the CKC framework. In recent literature, the CKC has proven to remain a relevant framework in cybersecurity threat modelling as its linear, sequential approach enables security teams the ability to identify and mitigate even the most complex cyber threats. However, it is not without its limitations; hence, organisations should consider deploying frameworks such as the MITRE ATT&CK that offer a more granular and flexible taxonomy of the tactics and techniques used by cyber attackers.

5.1 Future Research Directions

The fusion of AI and cybersecurity has led to the development of much-improved strategies and systems to handle the sophisticated cyber threats in our world today. However, there exist critical gaps that should be explored further in academic investigation. Considering the linearity of the CKC, more studies should be carried out to explore how it performs when deployed in tandem with other frameworks. AI models perform best when trained when quality data and with access to data come issues such as privacy. More studies should be conducted on the best approach to enhance AI model transparency and explainability [30, 31]. The following are other areas future research studies should be directed at:

1. *Legal, Ethical, and Policy Frameworks Governing the Use of AI in Cybersecurity:* AI has equipped both the bad guys and the good actors with tools and techniques to carry out their objectives across the cybersecurity landscape, thereby raising ethical and legal questions. Intentional efforts should go into the development of frameworks for responsible AI in areas of discussion such as international norms, accountability mechanisms, and ethical red lines for the use of AI in cybersecurity applications.

2. *Explainable Artificial Intelligence (XAI):* The reliance on AI for decision-making does not look like slowing down in areas such as real-time threat detection or automated incident response. Therefore, future research must focus on bridging the gap between complex AI models and human understanding as this increases trust across systems such as the exploitation and C2 phases of the CKC framework.
3. *Design of Resilient AI Systems Against Adversarial Attacks:* AI models are inherently vulnerable to adversarial attacks, where small perturbations in input data can mislead systems into incorrect classifications or predictions. In the context of the CKC, such vulnerabilities could be exploited to bypass detection systems or manipulate automated responses. Future research must explore the development of robust AI architectures and adversarial training techniques that can withstand sophisticated evasion strategies.
4. *Development of Simulation Environments to Evaluate AI Models Across CKC Stages:* The development of a controlled simulation environment will give room for security teams to test how their security systems respond to threat incidents under diverse cyberattack scenarios across the phases of the CKC framework. By developing open-source, modular simulators, cyberattacks can be reproduced to fine-tune the architecture to withstand actual attacks.

References

[1] Gaber, M. G., Ahmed, M., & Janicke, H. (2024). Malware Detection with Artificial Intelligence: A Systematic Literature Review. *ACM Computing Surveys, 56*(Article 148), 1–33. https://doi.org/10.1145/3638552

[2] Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A Survey of Deep Learning Methods for Cyber Security. *Information, 10*, 122. https://doi.org/10.3390/info10040122

[3] Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. *Leading Issues in Information Warfare and Security Research, 1*, 80–106.

[4] Kazimierczak, M., Habib, N., Chan, J. H., & Thanapattheerakul, T. (2024). Impact of AI on the Cyber Kill Chain: A Systematic Review. *Heliyon, 10*, e40699. https://doi.org/10.1016/j.heliyon.2024.e40699

[5] Caltagirone, S., Pendergast, A., & Betz, C. (2013). *The Diamond Model of Intrusion Analysis* (Report No. ADA586960). U.S. Department of Defense. http://dx.doi.org/10.13140/RG.2.2.31143.265648 1

[6] Kim, H., Kwon, H., & Kim, K. K. (2019). Modified Cyber Kill Chain Model for Multimedia Service Environments. *Multimedia Tools and Applications, 78*, 3153–3170. https://doi.org/10.1007/s11042-018-5897-5

[7] Vieira, S., Pinaya, W. H., & Mechelli, A. (2019). Introduction to Machine Learning. In *Machine learning: Methods and Applications to Brain Disorders* (pp. 1–6). Elsevier. https://doi.org/10.1016/B978-0-12-815739-8.00001-8

[8] Sharma, N., Sharma, R., & Jindal, N. (2021). Machine Learning and Deep Learning Applications – A Vision. *Global Transitions Proceedings, 2*, 24–28. https://doi.org/10.1016/j.gltp.2021.01.004

[9] Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). Cybersecurity Data Science: An Overview from Machine Learning Perspective. *Journal of Big Data, 7*, 41. https://doi.org/10.1186/s40537-020-00318-5

[10] Buczak, A. L., & Guven, E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials, 18*, 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

[11] Mohamed, N. (2025). Artificial Intelligence and Machine Learning in Cybersecurity: A Deep Dive into State-of-the-Art Techniques and Future Paradigms. *Knowledge and Information Systems*. https://doi.org/10.1007/s10115-025-02429-y

[12] Awad, A. I., Babu, A., Barka, E., & Shuaib, K. (2024). AI-Powered Biometrics for Internet of Things Security: A Review and Future Vision. *Journal of Information Security and Applications, 82*, 103748. https://doi.org/10.1016/j.jisa.2024.103748

[13] Mehmood, K., Ashraf, Z., Iqbal, R., Rafique, A. A., Gul, H., & Khawaja, D. (2025). Cyber Security Governance as a Pillar of Enterprise Risk Management: Designing a Compliance-Driven Framework for Operational Resilience, Policy Enforcement, and Regulatory Alignment. *Annual Methodological Archive Research Review, 3*, 59–77. https://doi.org/10.63075/0jv35d33

[14] Toussaint, M., Krima, S., & Panetto, H. (2024). Industry 4.0 Data Security: A Cybersecurity Frameworks Review. *Journal of Industrial Information Integration, 39*, 100604. https://doi.org/10.1016/j.jii.2024.100604

[15] Odarchenko, R., Pinchuk, A., Polihenko, O., & Skurativskyi, A. (2025). A Comparative Analysis of Cyber Threat Intelligence Models. In *Proceedings of the Third International Conference on Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN 2024)* (pp. 3–12). CEUR-WS.org.

[16] Naik, N., Jenkins, P., Grace, P., & Song, J. (2022). Comparing Attack Models for IT Systems: Lockheed Martin's Cyber Kill Chain, MITRE ATT&CK Framework and Diamond Model. In *Proceedings of the 2022 IEEE International Symposium on Systems Engineering (ISSE)* (pp. 1–7). IEEE. https://doi.org/10.1109/ISSE54508.2022.10005490

[17] Sayed, M. A., Anwar, A. H., Kiekintveld, C., Bosansky, B., & Kamhoua, C. (2023). Cyber Deception Against Zero-Day Attacks: A Game Theoretic Approach. *arXiv*. https://arxiv.org/abs/2307.13107

[18] Bada, M., Sasse, M. A., & Nurse, J. R. (2019). Cybersecurity Awareness Campaigns: Why do they fail to Change Behaviour? *arXiv*. https://arxiv.org/abs/1901.02672

[19] Touré, A., Imine, Y., Semnont, A., Delot, T., & Gallais, A. (2024). A Framework for Detecting Zero-Day Exploits in Network Flows. *Computer Networks, 248*, 110476. https://doi.org/10.1016/j.comnet.2024.110476

[20] Dargahi, T., Dehghantanha, A., Bahrami, P. N., Conti, M., Bianchi, G., & Benedetto, L. (2019). A Cyber-Kill-Chain Based Taxonomy of Crypto-Ransomware Features. *Journal of Computer Virology and Hacking Techniques, 15*, 153–174. https://link.springer.com/article/10.1007/s11416-019-00338-7

[21] Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). *MITRE ATT&CK: Design and Philosophy*. MITRE Corporation.

[22] Mavroeidis, V., & Bromander, S. (2017). Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence. In *Proceedings of the 2017 European Intelligence and Security Informatics Conference (EISIC)*. IEEE. https://doi.org/10.1109/EISIC.2017.20

[23] Abo El Rob, M. F., Islam, M. A., Gondi, S., & Mansour, O. (2024). The Application of MITRE ATT&CK Framework in Mitigating Cybersecurity Threats in the Public Sector. *Issues in Information Systems, 25*, 62–80. https://doi.org/10.48009/3_iis_2024_106

[24] CrowdStrike. (2025). What is the Cyber Kill Chain? Retrieved June 2, 2025, from https://www.crowdstrike.com/en-us/cybersecurity-101/cyberattacks/cyber-kill-chain/

[25] Splunk. (2025). Cyber Kill Chains: Learn How They Work and How to Break Them. Retrieved June 2, 2025, from https://www.splunk.com/en_us/blog/learn/cyber-kill-chains.html

[26] Cipher Security. (2024, September 14). The Unified Kill Chain: A Comprehensive Approach to Cybersecurity Defense. Retrieved June 2, 2025, from https://ciphersecurity.com/unified-kill-chain-approach-to-cyber-defense/

[27] Pols, P. (2023). *The Unified Kill Chain: Raising Resilience against Advanced Cyberattacks* (Version 1.3). Retrieved June 2, 2025, from https://www.unifiedkillchain.com/

[28] Sánchez del Monte, A., & Hernández-Álvarez, L. (2023). Analysis of Cyber-Intelligence Frameworks for AI Data Processing. *Applied Sciences, 13*, 9328. https://doi.org/10.3390/app13169328

[29] Hubbard, D. W., & Seiersen, R. (2023). *How to Measure Anything in Cybersecurity Risk* (2nd ed.). Wiley. https://www.wiley.com/en-us/How+to+Measure+Anything+in+Cybersecurity+Risk%2C+2nd+Edition-p-9781119892304

[30] Ali, I. (2024). *AI Transparency and Explainability*. Frankfurt University of Applied Sciences. Retrieved June 2, 2025, from https://www.researchgate.net/publication/386416207

[31] Marey, A., Arjmand, P., Alerab, A. D. S., Elsami, M. J., Saad, A. M., Sanchez, N., & Umair, M. (2024). Explainability, Transparency and Black Box Challenges of AI in Radiology: Impact on Patient Care in Cardiovascular Radiology. *Egyptian Journal of Radiology and Nuclear Medicine, 55*, 183. https://doi.org/10.1186/s43055-024-01356-2

[32] Astra (2025). *AI in Cybersecurity: Benefits and Challenges*. Retrieved June 2, 2025, from https://www.getastra.com/blog/ai-security/ai-in-cybersecurity/

[33] Mohurle, S., & Patil, M. (2017). A Brief Study of WannaCry Threat: Ransomware Attack 2017. *International Journal of Advanced Research in Computer Science, 8*, 1938–1940.

[34] Sharma, D. P., Habibi Lashkari, A., Firoozjaei, M. D., Mahdavifar, S., & Xiong, P. (2025). Defense Methods for Adversarial Attacks and Privacy Issues in Secure AI. In *Understanding AI in Cybersecurity and Secure AI* (pp. 159–195). Springer. https://doi.org/10.1007/978-3-031-91524-6_9

[35] Khaliq, S., Tariq, Z. U. A., & Masood, A. (2020). Role of User and Entity Behaviour Analytics in Detecting Insider Attacks. In *Proceedings of the 2020 International Conference on Cyber Warfare and Security (ICCWS)* (pp. 1–6). IEEE.

[36] Azeez, N. A., & Ajetola, A. R. (2009). Exploration of the Gap Between Computer Science Curriculum and Industrial IT Skills Requirements. *International Journal of Computer Science and Information Security (IJCSIS), 4*(1 & 2), [n.p.], USA.

[37] Azeez, N. A., Venter, I. M., & Tiko, I. (2011). Grid Security Loopholes with Proposed

Countermeasures. In *Proceedings of the 26th International Symposium on Computer and Information Sciences* (ISCIS 2011), Imperial College, London, UK. Springer Verlag.

[38] Azeez, N. A., & Van Vyver, C. (2018). Security Challenges and Suggested Solutions for E-Health Information in Modern Society. In *Proceedings of the 5th EAI International Conference on IoT Technologies for HealthCare (HealthyIoT 2018)*, Guimarães, Portugal. http://healthyiot.org/accepted-papers/

[39] Azeez, N. A., & Anochirionye, E. C. (2017). Detecting Malicious and Compromised URLs in E-Mails using Association Rule. *Covenant Journal of Informatics and Communication Technology (CJICT)*, *5*(2), 36–48.