International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 9 Issue 8 August - 2025, Pages: 44-62

A Rule-Based Expert System for Cybersecurity Threat Detection: Evolution, Applications, and the Hybrid AI Paradigm

Abdallah Quffa and Samy S. Abu-Naser

Information Technology Department
Faculty of Engineering and Information Technology
Al-Azhar University-Gaza

Abstract*: This paper examines the evolution and role of Rule-Based Expert Systems (RBES) in cybersecurity threat detection, highlighting their strengths, limitations, and the growing shift toward hybrid AI approaches. RBES have historically offered clear, rule-driven methods for identifying known threats, but their static nature struggles to keep pace with today's fast-changing cyber landscape—especially against zero-day exploits and advanced persistent threats (APTs). To address these challenges, researchers are increasingly turning to hybrid AI systems that combine symbolic reasoning with machine learning and deep learning. These neuro-symbolic models offer both adaptability and transparency, making them well-suited for high-stakes cybersecurity environments. This study explores the architecture of RBES, compares traditional and hybrid threat detection methods, and presents real-world applications and empirical findings. It also discusses ethical concerns such as bias, accountability, and explainability. Ultimately, the paper argues for the development of intelligent, adaptive, and trustworthy AI systems to strengthen cyber defense in an ever-evolving threat landscape.*

Keywords: Rule-Based Expert Systems, Cybersecurity, Threat Detection, Hybrid AI, Neuro-Symbolic AI, Machine Learning, Deep Learning, Knowledge-Based Systems, Explainable AI, Cyber Threats.

Introduction

In today's hyper-connected world, cybersecurity threats are evolving faster than ever. From ransomware and phishing to advanced persistent threats (APTs), organizations face a constant stream of malicious activity that puts critical infrastructure and sensitive data at risk. Detecting these threats effectively is no longer optional—it's essential.

Artificial Intelligence (AI) has long played a key role in cybersecurity, with Rule-Based Expert Systems (RBES) being among the earliest tools used to mimic human decision-making. These systems rely on clear "if-then" rules to identify known attack patterns, offering transparency and logical reasoning. However, as cyber threats become more dynamic and unpredictable, the limitations of static rule-based systems—such as poor adaptability and the difficulty of updating rules—have become increasingly apparent[1-4].

To keep pace, the field has shifted toward machine learning (ML) and deep learning (DL), which excel at recognizing patterns and adapting to new data. Yet, these approaches often operate as "black boxes," making it hard to understand how decisions are made—an issue that's especially problematic in high-stakes environments like cybersecurity [5-8].

This paper explores the emerging solution: hybrid AI, particularly neuro-symbolic systems that blend the strengths of symbolic reasoning with the adaptability of neural networks. These systems aim to deliver more accurate, flexible, and explainable threat detection.
By reviewing the evolution of RBES, analyzing current AI-driven methods, and examining ethical concerns such as bias and accountability, this research highlights the need for intelligent and trustworthy AI solutions in the ongoing battle against cyber threats [9-10].

Objectives

This research paper sets out to achieve the following key objectives:

- **Clarify the architecture and functionality of Rule-Based Expert Systems (RBES)** in the context of cybersecurity, detailing their core components and operational principles.
- **Analyze the historical role and limitations of RBES**, particularly their effectiveness in detecting known threats and their shortcomings in adapting to evolving cyberattack strategies.
- **Investigate the rationale behind hybrid AI approaches**, focusing on how symbolic AI can be integrated with machine learning and deep learning to enhance threat detection capabilities.
- **Review empirical studies and real-world case applications** that compare the performance of traditional RBES with modern hybrid AI systems in identifying and responding to cybersecurity threats.

- **Identify the major challenges** in deploying AI-driven cybersecurity solutions, including issues of scalability, adaptability, and the difficulty of acquiring and maintaining expert knowledge.
- **Examine the ethical dimensions** of AI in cybersecurity, such as the need for transparency, the risks of algorithmic bias, and the importance of accountability in automated decision-making.
- **Propose future research directions** aimed at developing intelligent, explainable, and adaptive AI systems that can effectively respond to the dynamic nature of cyber threats.

Problem Statement

Modern cyber threats are growing not only in volume but also in complexity and unpredictability. Traditional cybersecurity tools, particularly Rule-Based Expert Systems (RBES), have played a foundational role in detecting known threats using predefined rules. However, their static nature makes them increasingly ineffective against emerging threats like zero-day exploits and polymorphic attacks. Updating these systems requires significant manual effort and financial resources, creating a "knowledge acquisition bottleneck" that limits scalability and responsiveness.

Meanwhile, machine learning (ML) and deep learning (DL) have introduced powerful capabilities for identifying unknown threats by learning from data. Yet, these models often operate as opaque "black boxes," making it difficult to understand or explain their decisions—an issue that undermines trust, complicates incident response, and raises concerns about bias and accountability [11-15].

The core challenge is to develop cybersecurity systems that are both **adaptive** and **transparent**. While RBES offer clarity and logic, they lack flexibility. ML/DL models offer adaptability but sacrifice interpretability. This tension highlights the need for hybrid AI systems that combine the strengths of both approaches. Neuro-symbolic AI, in particular, offers a promising path forward—integrating symbolic reasoning with data-driven learning to create systems that are not only effective in detecting threats but also capable of explaining their decisions in a way that supports trust, compliance, and informed action.

Literature Review

Evolution of AI in Cybersecurity: From Rule-Based to Hybrid Systems

The application of Artificial Intelligence in cybersecurity has undergone a significant transformation, marked by distinct phases of technological advancement and adaptation to evolving threat landscapes [16-20]. This evolution highlights a continuous effort to overcome the limitations of preceding paradigms and build more robust and intelligent defense mechanisms.

Traditional Rule-Based Systems: Foundations and Early Applications

Early AI in cybersecurity was predominantly characterized by Rule-Based Expert Systems (RBES) [21-24]. These systems, which emerged from the broader field of Knowledge-Based Systems in the 1970s and proliferated in the 1980s, were designed to emulate human expert decision-making. RBES operate on a set of predefined "if-then" rules, explicitly encoding domain knowledge and logical reasoning. This approach provided transparency, allowing users to understand the rationale behind a system's conclusions.[7] Initial applications of RBES in cybersecurity included firewalls, intrusion detection systems (IDS), and antivirus software [25-30]. For instance, an early IDS like IDES would compare current user behaviors to historical profiles and use expert rules to define normal/suspicious behavior. Antivirus software relied on signature-based detection, matching known malware signatures against a rule database.[10]

A significant observation concerning RBES is the paradox of their early success. While expert systems were considered among the "first truly successful forms of AI software" and dominated AI research until the mid-1990s [31-36], their effectiveness was inherently constrained by their design. Their success in early cybersecurity applications stemmed directly from their ability to explicitly encode human expertise and provide transparent, logical decision-making, which was highly effective against *known* threats with *predefined* patterns [37]. However, this very strength became a critical limitation. These systems struggled with the inherent "brittleness" in handling out-of-domain problems, difficulties in knowledge acquisition, and the immense challenge of maintaining large knowledge bases. They proved inadequate against zero-day attacks and evolving threats that did not fit neatly into their predetermined rule sets [38]. This inflexibility ultimately contributed to the "second AI Winter" and a widespread recognition that traditional rule-based systems were no longer suitable for the dynamic and complex threats in modern cybersecurity environments. This historical trajectory illustrates that initial effectiveness in a controlled or predictable environment does not

guarantee long-term viability in dynamic domains like cybersecurity, underscoring the continuous need for adaptive intelligence.

Emergence of Machine Learning and Deep Learning in Cybersecurity

The limitations of static rule-based systems, particularly their inability to adapt to novel and evolving threats, spurred a fundamental shift towards machine learning (ML) and deep learning (DL) approaches in cybersecurity [39-40]. ML, defined as the field that gives computers the ability to learn without being explicitly programmed , analyzes vast datasets to identify patterns and anomalies, leading to adaptive security responses. Key types of ML include supervised learning, which is trained on labeled data (e.g., for spam detection), unsupervised learning, which discovers patterns in unlabeled data (e.g., for anomaly detection of novel threats), and reinforcement learning, which learns through trial and error (e.g., in robotics). Deep learning, a subset of ML, utilizes neural networks to process high-dimensional data, automatically extract features, and make complex decisions. It has revolutionized malware detection by scrutinizing code structure and behavior without relying on predefined signatures [41-44].

While ML and DL offer unprecedented adaptability and can detect previously unknown threats, a critical trade-off emerges: the "black box" problem. The complexity of these models, particularly deep neural networks, renders their decision-making processes opaque and difficult for humans to interpret. This opacity is a significant drawback in cybersecurity, where understanding *why* a particular threat was flagged is crucial for effective investigation, response, and forensic analysis. This lack of transparency can hinder accountability, complicate compliance with regulations that require explanations for AI decisions, and ultimately erode trust in the system, especially in critical security applications. Furthermore, the complexity can make these systems susceptible to adversarial attacks, where subtle manipulations of input data can deceive the model. The evolution from RBES to ML/DL, therefore, highlights a fundamental tension: gaining superior adaptability and detection capabilities often comes at the cost of interpretability. This inherent trade-off is a central factor driving the current development of hybrid AI [45-50].

Motivation for Hybrid AI: Bridging the Gap

The limitations inherent in both purely symbolic (rule-based) and purely sub-symbolic (machine learning/deep learning) AI approaches have driven the emergence of hybrid AI, also known as neuro-symbolic AI. Hybrid AI aims to combine the strengths of both paradigms: the logical reasoning, transparency, and low data hunger of symbolic AI with the pattern recognition, scalability, and adaptability of machine learning. This fusion is particularly valuable in enterprise settings where trust, interpretability, and compliance are critical, allowing for the creation of more robust, accurate, and context-aware AI systems [51-53].

A deeper understanding of the motivation for hybrid AI reveals an ambition beyond mere technical integration: it seeks to emulate human cognitive processes more closely. The theoretical foundation for neuro-symbolic AI is often linked to Daniel Kahneman's "System 1" (fast, intuitive, pattern recognition) and "System 2" (slow, deliberate, logical reasoning) models of thinking.[35] In this framework, deep learning is seen as excelling at System 1 cognition, adept at rapid pattern recognition, while symbolic reasoning is best suited for System 2, handling planning, deduction, and deliberative thinking. This conceptual alignment suggests that achieving true "intelligence" in AI, especially for complex tasks like cybersecurity, necessitates both intuitive pattern matching and explicit logical reasoning, mirroring how human experts operate. A system capable of both "intuitively" spotting anomalies and "logically" explaining *why* it flagged something would inherently be more effective and trustworthy. This pursuit of integrated, robust intelligence, drawing inspiration from human cognition, represents a significant driving force behind the neuro-symbolic AI paradigm [54-56].

Core Concepts of Rule-Based Expert Systems

Understanding the foundational principles of Rule-Based Expert Systems (RBES) is crucial for appreciating their role in cybersecurity and the subsequent evolution towards hybrid AI [57-60].

Architecture and Components

A Rule-Based Expert System (RBES) is a type of Knowledge-Based System (KBS) that emulates human expert decision-making.[56] Its core architecture typically comprises several key components:

- **Knowledge Base:** This serves as the central repository of domain-specific facts and rules, most commonly represented as "if-then" statements, also known as production rules.[7] It encapsulates the accumulated "know-how" of human experts in a particular domain.[8]

- **Inference Engine:** Often referred to as the "brain" of the system, the inference engine is responsible for applying the rules from the knowledge base to the current facts to deduce new information or arrive at conclusions and decisions. Its primary function is to link the defined rules with the available facts and perform logical reasoning [61-62].

- **Working Memory (or Database/Facts):** This is a dynamic component that holds the current facts or data being processed by the system at any given time. As the inference engine applies rules and new information is derived, the working memory is updated accordingly. [63]
- **User Interface:** This layer facilitates communication and interaction between the user and the expert system, providing mechanisms for inputting data and displaying outputs or recommendations [64].
- **Explanation Module:** A distinctive and crucial feature of expert systems, this module provides users with justifications for the system's conclusions and explains the step-by-step reasoning process that led to a particular decision. This transparency is vital for building user trust and understanding [65].

The explicit inclusion of a "User Interface" and, more notably, an "Explanation Module" within the core architecture of RBES [7] reveals a foundational design principle: these systems were built with human interaction and understanding as central tenets. The explanation module's role in providing justification for conclusions and detailing the reasoning process stands in stark contrast to the later "black box" problem prevalent in pure machine learning models, where decisions are often uninterpretable [66]. This human-centric design, despite the eventual limitations of RBES in dynamic environments, established a crucial precedent for the concept of Explainable AI (XAI), which is now a major research area in hybrid AI. The historical emphasis on transparency in RBES therefore provides a conceptual lineage for current efforts to make advanced AI systems more understandable and trustworthy [67].

Table 1: Core Components of a Rule-Based Expert System and Their Role in Cybersecurity

Component	Description	Role in Cybersecurity Threat Detection
Knowledge Base	Repository of domain-specific facts and "if-then" rules, encapsulating human expert knowledge.	Stores rules defining known attack patterns (e.g., specific malware signatures, unusual login attempts) and normal system behaviors.
Inference Engine	Applies rules to facts in the knowledge base to derive conclusions or make decisions.	Processes real-time network traffic and system logs against predefined rules to identify potential threats. Determines if observed activity matches a malicious pattern.
Working Memory	Dynamic storage for current facts or data being processed by the system.	Holds current network events, user activities, system states, and temporary data relevant to the ongoing threat analysis. Updated as new information is gathered or inferred.
User Interface	Facilitates interaction between the user and the system for input and output.	Provides a dashboard for security analysts to input parameters, view alerts, and receive recommendations. Enables human oversight and intervention.
Explanation Module	Provides justifications for the system's conclusions and explains the reasoning process.	Offers transparency by detailing why a specific activity was flagged as a threat, tracing the sequence of rules and facts that led to the detection. Aids in forensic analysis and trust-building.

Knowledge Acquisition and Representation Techniques

Knowledge acquisition is the process of gathering and formalizing domain-specific knowledge, typically from human experts through interviews, document analysis, or data mining techniques. Once acquired, this knowledge must be represented in a formal

notation that the system can process and utilize.[9] Common representation techniques include:

- **Rule-based notation (IF-THEN rules):** This is the most prevalent form, representing knowledge as conditional statements (e.g., "IF a system receives more than X connection requests within Y seconds from a single IP address THEN flag as DDoS attack") [68].
- **Ontologies:** These are formal representations that define concepts, categories, and the relationships between them within a specific domain [69]. In cybersecurity, ontologies can define relationships between threats, vulnerabilities, defenses, and assets, aiding in the classification of detected threats, cross-referencing with past attacks, and suggesting appropriate countermeasures.
- **Frame-based notation:** This technique uses structured representations of concepts, similar to object-oriented programming classes, with "slots" for attributes and their associated values.
- **Semantic Networks:** These are graphical representations of knowledge, using nodes to represent concepts and links to represent relationships between them.

A significant challenge in the development of traditional RBES was the "knowledge acquisition bottleneck" This refers to the inherent difficulty and labor-intensiveness of manually acquiring and formalizing domain-specific knowledge from human experts. This process is complex, time-consuming, and expensive, particularly as the problem domain grows in complexity. This limitation was a primary factor driving the shift away from purely rule-based systems towards machine learning, which offered the promise of learning patterns directly from data.[17] However, the current evolution towards hybrid AI reintroduces the value of structured knowledge. In this new paradigm, machine learning can be leveraged to *automate* parts of the knowledge acquisition process, for instance, by automatically learning rules and ontologies from data. Alternatively, pre-existing human knowledge can be seamlessly *integrated* to guide and enhance the learning processes of machine learning models. This cyclical pattern in AI development, where past challenges are revisited with new computational tools, suggests a mature understanding that neither purely manual knowledge engineering nor purely data-driven learning is sufficient in isolation for complex, dynamic domains [70].

Types of Rules and Reasoning Mechanisms

Rules within RBES can be categorized by their functional purpose:

- **Deductive Rules:** These are used for logical deductions, where if a condition is true, a specific conclusion must logically follow (e.g., "If a file matches a known malware signature, then it is malicious") [71].
- **Reactive Rules:** These rules are designed for event-driven actions, triggering a specific response when a certain event occurs (e.g., "When network traffic from a single source exceeds 1000 packets per second, trigger an alert").
- **Production Rules:** A general category of rules used for decision-making and problem-solving, typically in the "IF-THEN" format, where specific conditions lead to corresponding actions.

Inference engines, the core reasoning component, employ different mechanisms to apply these rules and derive conclusions:

- **Forward Chaining:** This data-driven approach starts with a set of available facts or input data and applies rules to deduce new information or reach a conclusion. For example, if a system detects an unusual spending pattern, forward chaining would apply rules to determine if this pattern indicates fraudulent activity [71].
- **Backward Chaining:** This goal-driven approach starts with a goal or hypothesis and works backward to determine the facts or evidence required to prove or disprove it. For instance, to confirm a specific type of cyberattack, backward chaining would seek evidence that matches the attack's known characteristics.
- **Hybrid Inference Engines:** These advanced engines combine the strengths of both forward and backward chaining, often employed in complex applications that necessitate multiple reasoning strategies. This approach allows for both data-driven discovery and goal-driven verification.

The existence of "hybrid inference engines" even within traditional RBES, which combine forward and backward chaining, signifies an early, implicit recognition that a single, rigid reasoning approach is often insufficient for complex problem-solving. This internal hybridization within the symbolic AI paradigm conceptually paved the way for the broader integration of symbolic and sub-symbolic AI. It suggests that the desire for flexible, multi-faceted reasoning and integrated, robust intelligence is not a new development but a persistent theme in AI research, continuously adapting its form as new computational paradigms emerge. This historical progression illustrates a foundational understanding that complex problems benefit from diverse reasoning strategies [73].

Cybersecurity Threat Landscape

Common Cyber Threats and Their Characteristics

The contemporary cybersecurity landscape is characterized by a diverse and rapidly evolving array of threats. Understanding these threats is fundamental to designing effective detection systems. Common categories of cyber threats include [65]:

- **Malware:** Abbreviation for "malicious software," this broad category includes viruses, worms, trojans, spyware, and ransomware. Malware is designed to infiltrate systems, compromise confidentiality, integrity, or availability of data, and can cause widespread damage and disruption. Spyware, specifically, aims to violate privacy by tracking personal activities or facilitating financial fraud.
- **Ransomware:** A specific type of malware that encrypts a user's or organization's systems or data, denying access until a ransom (often in cryptocurrency) is paid for a decryption key. Ransomware is difficult to detect before it's too late, and its techniques constantly evolve. Human-operated ransomware, where attackers gain access to an entire network, is a growing concern.[2]
- **Distributed Denial of Service (DDoS) Attacks:** These attacks aim to make an online service unavailable by overwhelming it with excessive traffic from multiple compromised systems (botnets), causing website response times to slow down or preventing access entirely. DDoS attacks are often used as a distraction for other forms of fraud or cyber intrusion.
- **Phishing and Social Engineering:** Phishing involves sending fraudulent emails or messages pretending to come from a trusted source to trick individuals into revealing sensitive information or downloading malicious code. Social engineering encompasses broader psychological manipulation tactics, including baiting, pretexting, vishing (voice phishing), and smishing (SMS phishing), to gain unauthorized access or information. These attacks are becoming more sophisticated and personalized, often leveraging generative AI.
- **Insider Threat:** Not all threats originate externally. Insider threats involve trusted individuals with authorized access who inadvertently or maliciously harm an organization by compromising data or systems.
- **Identity-Based Attacks:** These attacks involve compromising user identities, where cyberattackers steal or guess credentials to gain unauthorized access to systems and data. Credential stuffing and brute-force attacks fall into this category.
- **Supply Chain Attacks:** Attackers target an organization by tampering with software or hardware supplied by a third-party vendor, introducing malicious code or vulnerabilities into the supply chain.[2]
- **Code Injection:** Exploiting vulnerabilities in how source code handles external data, cybercriminals inject malicious code into an application, such as SQL injection or cross-site scripting.

The increasing complexity of these threats, coupled with the rapid adoption of new technologies and the widening cyber skills gap, contributes to a more opaque and unpredictable risk landscape. Generative AI, in particular, is noted as a catalyst for cybercrime, enabling more sophisticated and scalable attacks, lowering the cost of phishing, and streamlining the process from vulnerability exploitation to malware deployment.

Traditional Threat Detection Methods

Cybersecurity threat detection systems employ various methods to identify and mitigate malicious activities. These methods often form the foundation upon which more advanced AI-driven solutions are built [66]:

- **Signature-based Detection:** This method relies on a database of known indicators of compromise (IOCs), such as unique file hashes, IP addresses, or specific patterns in network traffic or software. It is fast and reliable for detecting *known* threats that match predefined signatures (e.g., antivirus software detecting known malware). However, its primary limitation is its ineffectiveness against novel, zero-day, or polymorphic attacks that do not have a pre-existing signature.[1]
- **Anomaly-based Detection:** This approach flags deviations from expected or "normal" patterns in network traffic, system performance, or user activity. By establishing a baseline of normal behavior, any significant departure from this norm can indicate a potential threat. This method is particularly effective for spotting stealthy, novel, or zero-day threats that signature-based systems would miss.
- **Behavior-based Detection:** Similar to anomaly detection, this method monitors typical user or system behavior over time to detect suspicious shifts. It focuses on actions and sequences of events rather than static patterns. Examples include unusual access to sensitive data, lateral movement across systems, or anomalous login times. This approach is valuable for identifying insider threats or sophisticated attacks that mimic legitimate activity.
- **Intelligence-driven Detection:** This method integrates external threat intelligence feeds—data streams highlighting current and potential cyberattacks, tactics, techniques, and procedures (TTPs)—to identify emerging threats earlier. It allows security teams to take a more proactive approach by leveraging up-to-date information on cybercriminal behaviors and trends

Most modern cybersecurity platforms layer these approaches to improve overall visibility and reduce false positives. This multi-

layered strategy acknowledges that no single detection method is sufficient against the full spectrum of modern cyber threats.

Table 2: Comparison of Traditional Cybersecurity Threat Detection Methods

Method	Principle	Strengths	Weaknesses	Role in Cybersecurity
Signature-based	Matches known indicators of compromise (IOCs) against a database of predefined patterns.	Fast and reliable for known threats; low false positives for exact matches.	Ineffective against novel, zero-day, or polymorphic attacks; requires constant updates.	Foundational for antivirus, firewalls, and basic IDS; identifies common, well-documented threats.
Anomaly-based	Flags deviations from established baselines of normal network, system, or user behavior.	Effective against novel and zero-day threats; can detect previously unseen attacks.	High potential for false positives if baselines are not accurately established or environment changes rapidly.	Detects unusual activities that may indicate a threat; complements signature-based methods.
Behavior-based	Monitors typical user/system actions over time to identify suspicious shifts or sequences of activity.	Good for detecting insider threats, sophisticated attacks, and lateral movement; focuses on malicious intent.	Requires extensive data collection and analysis to establish accurate behavioral profiles; can be resource-intensive.	Identifies malicious intent or compromised accounts by analyzing activity patterns; used in UEBA.
Intelligence-driven	Integrates external threat intelligence (TTPs, IoCs) to identify emerging threats.	Proactive defense against advanced attacks; leverages collective knowledge of cybercriminal activities.	Relies on timely and accurate external feeds; may not cover highly targeted or unique attacks.	Informs threat hunting, risk assessment, and strategic defense planning; enhances early detection.

Rule-Based Expert Systems in Cybersecurity: Applications and Limitations

Historical Applications and Performance

Historically, Rule-Based Expert Systems (RBES) played a foundational role in cybersecurity, particularly in the development of early intrusion detection systems (IDS) and antivirus software. Early IDS, such as the Intrusion Detection Expert System (IDES), operated by comparing current user behaviors against historical profiles and applying expert-defined rules to identify deviations indicative of suspicious activity. These systems aimed to automate intrusion detection and even detect zero-day attacks, which previously required extensive human analysis. Similarly, antivirus software of the era relied heavily on signature-based detection,

where a database of known malware signatures was matched against executable files or network traffic to identify threats [74].

While these systems were effective in their time, particularly against known and well-defined threats, their performance was inherently limited by their static nature. For instance, a rule-based antivirus could only detect a virus if its signature was already present in its database. Rule-based systems for malware detection, such as those using YARA rules, have shown effectiveness in identifying a wide range of malware families and variants through pattern matching. Recent efforts to automate YARA rule generation, leveraging Large Language Models (LLMs), have demonstrated promising results, with one system generating 763 rules (452 YARA and 311 Semgrep) achieving a precision of 85.2% and a recall of 91.8% in identifying malicious packages, outperforming state-of-the-art tools. This indicates that even in modern contexts, rule-based approaches, especially when augmented by advanced AI, can still contribute to cybersecurity [75].

Challenges: Adaptability, Scalability, and Maintenance in Evolving Threat Landscapes

Despite their historical significance and foundational contributions, purely rule-based expert systems face significant challenges that limit their effectiveness in the dynamic modern cybersecurity landscape [76]:

- **Lack of Adaptability to Evolving Threats:** The primary limitation of RBES is their reliance on predefined, static rules.[10] Cybercriminals constantly modify their attack patterns, creating novel or polymorphic variants that do not match existing signatures or rules. This inflexibility makes RBES ineffective against zero-day exploits and rapidly evolving threats. The static nature of these systems means they cannot learn from new data or adapt to unforeseen scenarios without manual intervention.
- **Knowledge Acquisition Bottleneck:** As discussed previously, the process of acquiring and formalizing domain-specific knowledge into explicit rules is labor-intensive, time-consuming, and expensive. This "bottleneck" becomes increasingly severe as the complexity and volume of cyber threats grow, making it impractical to manually encode rules for every new threat or variant.
- **Scalability Issues:** As the number of rules and the volume of data to be processed increase, RBES can become cumbersome and difficult to manage. The computational overhead of matching and executing a vast number of rules can degrade system performance, making real-time threat detection challenging in large-scale networks.
- **Rule Conflict Resolution:** In large and complex rule bases, inconsistencies or conflicts between rules can arise, leading to ambiguous or incorrect decisions. Resolving these conflicts requires significant manual effort and can be challenging to manage effectively.[9]
- **Maintenance Burden:** Maintaining and updating rule bases to ensure their accuracy and relevance in a constantly changing threat environment requires substantial manual involvement and significant financial resources. This ongoing effort makes RBES less cost-effective and efficient compared to systems that can learn and adapt autonomously.

These limitations underscore why traditional rule-based systems, while foundational, are no longer sufficient as standalone solutions for comprehensive cybersecurity threat detection. They highlight the imperative for more adaptive and intelligent approaches that can overcome these inherent rigidities.

Table 3: Advantages and Disadvantages of Rule-Based Expert Systems in Cybersecurity

Aspect	Advantages	Disadvantages
Interpretability/ Transparency	Highly transparent; decision-making process is explicit and human-understandable ("if-then" logic).[7] Facilitates debugging and trust.	Limited flexibility; incapable of autonomous learning or adapting to new situations without manual rule modification.[13]
Knowledge Handling	Can combine and preserve the knowledge of human experts.[7] Can operate with uncertain or incomplete	**Knowledge Acquisition Bottleneck:** Difficult and time-consuming to

	knowledge.[8]	acquire and formalize domain knowledge into rules.[55]	
Adaptability to Threats	Effective for detecting *known* threats with predefined patterns.[10]	**Poor Adaptability to Novel/Evolving Threats:** Struggles with zero-day attacks, polymorphic malware, and rapidly changing attack patterns.[10]	
Scalability & Maintenance	Modularity allows for easier extension of rules without affecting existing ones.[12]	**Scalability Issues:** Becomes cumbersome and difficult to manage as the number of rules and data volume increase.[11]	**High Maintenance Burden:** Requires substantial manual effort and financial resources to update rules.[10]
Precision & Accuracy	Can provide high precision for well-defined problems with clear rules.[15]	Prone to making incorrect assumptions if rules are not immaculately defined.[13] May lead to inaccuracies with uncertain or ambiguous information.[15]	

Hybrid AI and Neuro-Symbolic Approaches for Enhanced Threat Detection

The limitations of both purely symbolic and purely sub-symbolic AI paradigms have spurred the development of hybrid AI, particularly neuro-symbolic AI, as a promising solution for advanced cybersecurity threat detection. These approaches aim to combine the strengths of logical reasoning with data-driven learning to create more robust, explainable, and adaptable systems.[2]

Integration Strategies and Architectural Patterns

Hybrid AI models employ various strategies to combine machine learning with domain knowledge, leveraging expert information at different stages of the AI pipeline. Henry Kautz's taxonomy provides a useful framework for categorizing these neuro-symbolic architectures [77]:

- **Symbolic Neural Symbolic:** This approach is common in large language models (LLMs) used in natural language processing, where words or subword tokens serve as direct input and output for neural networks.
- **Symbolic[Neural]:** Symbolic techniques invoke neural techniques. A prominent example is AlphaGo, where Monte Carlo tree search (symbolic) calls upon neural networks to evaluate game positions. In cybersecurity, this could involve symbolic rules triggering ML models for deeper analysis of suspicious events.
- **Neural | Symbolic:** A neural architecture interprets perceptual data, converting it into symbols and relationships that are then processed symbolically. For instance, a neural network might recognize objects in an image (e.g., a malicious file icon) and then apply logical rules to understand relationships between those objects and system behaviors.
- **Neural: Symbolic → Neural:** Symbolic reasoning generates or labels training data, which is subsequently learned by a deep learning model. This approach can address the data scarcity problem often faced by ML models, especially for rare cyberattack types.
- **Neural_Symbolic_:** A neural network is directly generated from symbolic rules. Examples include Neural Theorem Provers and Logic Tensor Networks, which encode logical formulas as differentiable functions, allowing symbolic knowledge to be integrated with gradient-based learning.
- **Neural:** A neural model directly calls a symbolic reasoning engine to perform an action or evaluate a state. This could be seen

in systems where a neural network detects an anomaly, then calls a rule-based engine to determine the appropriate response based on compliance or safety protocols.

Beyond Kautz's taxonomy, other integration strategies include:

- **Pre-processing or Feature Engineering with Domain Knowledge:** Domain expertise is used to design or select relevant features, transform raw data, or filter noise, improving model input quality and learning efficiency.
- **Incorporating Domain Knowledge as Constraints During Model Training:** Rules, physical laws, or logical constraints are embedded directly into the learning process (e.g., via constrained optimization or knowledge-based loss functions) to guide model parameters and ensure predictions respect essential domain principles.
- **Post-processing ML Outputs with Rule-Based Corrections:** After an ML model generates predictions, domain knowledge can be applied through rule-based systems to validate, adjust, or correct results, enhancing reliability and preventing outputs that violate domain-specific rules. This is exemplified by hybrid fraud detection systems where ML identifies unusual spending, and symbolic rules flag potentially fraudulent transactions based on predefined compliance standards.

These architectural patterns demonstrate the diverse ways in which hybrid AI systems can be designed to leverage the complementary strengths of symbolic and sub-symbolic approaches, moving towards more comprehensive and intelligent solutions.

Benefits of Hybridization:

Explainability, Robustness, Data Efficiency, and Generalization

The integration of symbolic and sub-symbolic AI in hybrid systems offers several significant benefits, directly addressing the limitations of their standalone counterparts [78]:

- **Enhanced Explainability and Transparency:** Hybrid models provide clearer explanations for their decisions by linking outcomes to known rules, concepts, or expert reasoning. This addresses the "black box" problem of deep learning, increasing user trust and facilitating regulatory compliance and auditing. For example, a hybrid system can explain a loan denial by citing a specific rule violation (e.g., "debt-to-income ratio violates Rule 12.4").
- **Improved Accuracy and Robustness:** Incorporating domain knowledge guides model learning and inference, reducing errors caused by noisy or insufficient data. This results in more accurate and stable predictions across varying conditions, making systems more resilient to errors and unexpected inputs.
- **Faster Learning with Less Data (Data Efficiency):** Domain knowledge acts as prior information, allowing models to learn meaningful patterns more quickly and effectively from limited datasets, thus reducing data dependency. This is particularly valuable in cybersecurity where labeled data for novel threats might be scarce. Neuro-symbolic models can achieve high accuracy with a fraction of the data required by pure neural networks.
- **Better Generalization in Real-World Scenarios:** By combining data-driven learning with explicit rules, hybrid models can generalize better to unseen or out-of-domain scenarios. Symbolic components can enforce logical constraints, ensuring that predictions align with known principles even when data is sparse or unusual.
- **Handling Uncertainty and Ambiguity:** Hybrid systems can effectively handle both unstructured data (via neural networks) and complex reasoning tasks (via symbolic logic), offering improved decision-making in ambiguous or uncertain situations.[22] Probabilistic methods can be incorporated alongside traditional logic to make informed decisions with incomplete or noisy data.
- **Adaptability with Integrity:** Neural layers allow the system to learn and evolve from new data, while symbolic layers ensure that core ethical, legal, and safety standards remain intact. This creates systems that can adapt without compromising foundational principles.

These benefits highlight the potential of hybrid AI to overcome the inherent weaknesses of monolithic AI approaches, paving the way for more intelligent, reliable, and context-aware cybersecurity systems.

Empirical Studies and Real-World Applications

Hybrid AI systems are increasingly being deployed across various industries, demonstrating their practical potential in complex decision-making tasks, including those relevant to cybersecurity.

- **Fraud Detection in Finance:** Financial institutions utilize hybrid AI systems to combine machine learning's ability to identify unusual spending patterns with predefined symbolic AI rules for flagging potentially fraudulent transactions. This dual-model

approach helps prevent fraud and reduces false positives, allowing human experts to follow up on questionable transactions with investigations. Large banks use AI-powered behavior analytics to monitor login patterns and transaction anomalies, flagging suspicious activity for immediate review.

- **Phishing Campaigns and Email Security:** Healthcare providers have implemented AI-driven email filtering tools that scan message context, tone, and metadata, not just keywords or known blacklists. These systems have successfully blocked spear phishing emails impersonating executives, preventing credential theft and potential ransomware deployment. AI models analyze email metadata, content, and attachment behavior, leveraging natural language processing (NLP) to evaluate the tone, structure, and intent of messages, identifying subtle social engineering tactics.[75]
- **Malware Detection:** While traditional rule-based systems (like YARA) are effective for known malware, hybrid approaches are emerging. Research on RuleLLM, which leverages Large Language Models (LLMs) to automate YARA and Semgrep rule generation, has shown promising empirical results. RuleLLM generated 763 rules with a precision of 85.2% and a recall of 91.8% in identifying malicious packages, significantly outperforming baseline tools. This demonstrates how ML can enhance the traditional rule-based approach. Dynamic deep learning methods combined with heuristic approaches have also shown improved performance in classifying and detecting modern malware families.
- **Intrusion Detection Systems (IDS):** AI-driven IDS improve network security by monitoring traffic patterns and detecting anomalies that signal intrusions. Machine learning models distinguish between normal and malicious network behaviors, identifying threats like DDoS attacks and lateral movement. Hybrid IDS architectures that adaptively apply interpretable regulations, combining misuse and anomaly-based detection, have been suggested to provide a more understandable illustration of classification models. Empirical studies on ML-based IDS using datasets like UNSW-NB15 have evaluated performance metrics such as accuracy, precision, recall, and F1-score, with models like XGBoost and CatBoost achieving high accuracy (87%) and interpretability.
- **IoT Threat Detection:** Global manufacturers have deployed AI-based anomaly detection on their production networks to learn normal device communication patterns and flag abnormal activity, such as unexpected firmware updates or lateral movement attempts. This helps avoid costly downtime and equipment damage in smart factories and industrial control systems.
- **Incident Response and Security Operations Center (SOC) Automation:** AI helps SOCs filter out false positives and prioritize real threats. AI-driven Security Orchestration, Automation, and Response (SOAR) platforms automatically correlate logs, enrich alerts with threat intelligence, and trigger incident response workflows without human intervention. One enterprise SOC reported a 70% reduction in average response time, freeing analysts to focus on higher-level threats.
- **User Behavior Analytics (UBA):** AI is used to establish behavioral baselines for individual users and detect deviations that could indicate insider threats or compromised accounts, such as unusual login times or data access patterns.
- **Autonomous Systems (e.g., Self-Driving Cars):** While not directly cybersecurity, this domain illustrates hybrid AI's capability. Self-driving cars combine symbolic systems for safe driving rules with machine learning for object detection (pedestrians, vehicles). This highlights the integration of learned perception with explicit safety constraints.
- **Medical Diagnosis:** Although a different domain, neuro-symbolic AI has shown promise in medical diagnosis, where Logical Neural Networks (LNNs) integrate domain-specific knowledge through logical rules with learnable weights. Studies have demonstrated enhanced diagnostic accuracy and explainable diagnostic pathways, bridging the gap between accuracy and interpretability. This serves as a strong parallel for the potential of hybrid AI in cybersecurity, where both accuracy and explainability are paramount.

These empirical results and real-world applications underscore the growing maturity and effectiveness of hybrid AI approaches in addressing complex, high-stakes problems, providing a strong foundation for their continued development and deployment in cybersecurity threat detection.

Table 4: Key Hybrid AI Architectures and Integration Strategies for Cybersecurity

Architecture/Strategy	Description	Example/Mechanism	Benefits for Cybersecurity
Symbolic[Neural]	Symbolic techniques invoke neural networks for specific sub-tasks.	Symbolic rules trigger ML models for deeper analysis of suspicious network flows or unusual user behavior.	Combines logical control with powerful pattern recognition; allows targeted application of ML.

**Neural	Symbolic**	Neural networks interpret raw data into symbolic representations, which are then reasoned about symbolically.	A neural network identifies suspicious patterns in system logs (e.g., unusual API calls), which are then translated into symbols for a rule-based engine to classify and respond.
Neural: Symbolic → Neural	Symbolic reasoning generates or labels training data for neural networks.	Security experts define rules for generating synthetic attack data to train ML models, especially for rare or zero-day threats.	Addresses data scarcity issues for ML; ensures training data aligns with expert knowledge; improves learning efficiency.
Neural_Symbolic_	Neural networks are directly structured or constrained by symbolic rules.	Logic Tensor Networks encode logical formulas as differentiable functions within a neural network, allowing symbolic knowledge to guide learning.	Integrates explicit knowledge directly into the learning architecture; ensures predictions adhere to security policies and logical constraints.
Neural	A neural model directly calls a symbolic reasoning engine for specific actions or evaluations.	An ML-based anomaly detection system flags a high-risk event, then calls a rule-based engine to determine the appropriate automated response (e.g., isolate endpoint, block IP) based on predefined security playbooks.	Combines ML's adaptive detection with symbolic AI's precise, explainable, and policy-driven response.
Knowledge-Infused Feature Engineering	Domain expertise is used to create or select relevant features for ML models.	Security analysts define features (e.g., packet size ratios, frequency of specific port access) that are known indicators of certain attacks, improving ML model input quality.	Improves learning efficiency and accuracy by providing meaningful input features; reduces reliance on raw, noisy data.
Rule-Based Post-processing of ML Outputs	ML model predictions are validated or corrected by rule-based systems.	An ML model predicts a transaction as fraudulent, but a rule-based system verifies if it violates a specific compliance rule before an alert is triggered, reducing false positives.	Enhances reliability and compliance; provides a "guardrail" for ML predictions; adds a layer of explainability to final decisions.

Methodology

This research adopts a **systematic literature review** methodology to investigate the evolution, architecture, and effectiveness of Rule-Based Expert Systems (RBES) in cybersecurity threat detection, as well as the emergence of hybrid AI paradigms. The approach is structured to ensure comprehensive coverage, critical analysis, and evidence-based insights.

1. Scope Definition

The study focuses on RBES within the domain of cybersecurity, emphasizing their historical development, core components, limitations, and the transition toward hybrid AI systems—particularly neuro-symbolic AI. The scope includes both theoretical foundations and practical applications.

2. Information Retrieval

A broad and targeted search was conducted across academic databases, industry publications, and reputable online sources. Keywords included:

- "Rule-Based Expert Systems"
- "Cybersecurity Threat Detection"
- "Hybrid AI"
- "Neuro-Symbolic AI"
- "Machine Learning in Cybersecurity"
- "Explainable AI"
- "Knowledge-Based Systems"

Sources were selected based on relevance, credibility, and recency to ensure a robust foundation for analysis.

3. Data Extraction and Synthesis

Key information was extracted from the selected literature, focusing on:

- Definitions and conceptual frameworks
- System architectures and components
- Historical applications and performance metrics
- Integration strategies for hybrid AI
- Empirical results and case studies
- Ethical considerations in AI deployment

The extracted data were synthesized to identify recurring themes, trends, and causal relationships.

4. Critical Analysis

The synthesized findings were critically examined to uncover deeper insights. This included:

- Evaluating the trade-offs between interpretability and adaptability in AI systems
- Assessing how limitations in RBES influenced the rise of ML/DL
- Exploring how hybrid AI reconciles symbolic and sub-symbolic approaches
- Analyzing ethical implications such as transparency, bias, and accountability

This analysis aimed to contextualize technological developments within broader cybersecurity challenges.

5. Structure and Presentation

The paper is organized into standard academic sections: Introduction, Objectives, Problem Statement, Literature Review, Methodology, Results, Discussion, Conclusion, and References. Tables and diagrams are used to present comparative data and architectural models clearly and accessibly.

6. Citation and Referencing

All claims and data points are supported by rigorous citations, following a consistent referencing style. This ensures academic integrity and allows readers to trace sources for further exploration.

Results

The systematic literature review conducted in this study revealed a clear trajectory in the evolution of cybersecurity threat detection systems—from static, rule-based architectures to dynamic, learning-based models, and ultimately toward integrated hybrid AI solutions.

Performance of Rule-Based Expert Systems (RBES)

Historically, RBES played a pivotal role in early cybersecurity applications such as Intrusion Detection Systems (e.g., IDES) and antivirus software. These systems excelled at identifying **known threats** by matching predefined patterns and signatures. Their **transparency** and **logical reasoning** made them highly interpretable and trustworthy for human analysts.

However, their effectiveness diminished significantly in the face of **zero-day attacks**, **polymorphic malware**, and **rapidly evolving threat vectors**. The static nature of RBES, coupled with the **knowledge acquisition bottleneck**—the labor-intensive process of manually updating rule sets—limited their scalability and adaptability.

Advancements through Machine Learning and Deep Learning

The transition to **machine learning (ML)** and **deep learning (DL)** marked a major leap in threat detection capabilities. ML models, using supervised and unsupervised learning, demonstrated the ability to detect **anomalies** and **unknown threats** by analyzing large volumes of data. DL techniques, particularly those employing neural networks, enhanced malware detection by examining behavioral patterns rather than relying solely on signatures.

Empirical evidence supports these advancements:

- **Dynamic deep learning** combined with heuristic methods outperformed static models in malware classification.
- AI-driven cybersecurity tools are projected to save organizations **over $150 billion annually** by 2025.
- MIT reports that AI-based systems detect cyberattacks **85% faster** than traditional tools.

Despite these gains, ML/DL models introduced the **"black box" problem**, where decision-making processes became opaque, hindering interpretability, trust, and forensic analysis.

Emergence and Impact of Hybrid AI

To address the limitations of both RBES and ML/DL, **hybrid AI systems**—particularly **neuro-symbolic AI**—have emerged. These systems integrate symbolic reasoning with neural learning, offering both **adaptability** and **explainability**.

Key empirical findings include:

- **RuleLLM**, a hybrid system using large language models to generate YARA and Semgrep rules, achieved **85.2% precision** and **91.8% recall**, outperforming state-of-the-art tools.
- **Hybrid Intrusion Detection Systems (IDS)** using models like XGBoost and CatBoost on the UNSW-NB15 dataset reached **87% accuracy** with low false positive and negative rates.
- **SOAR platforms** powered by AI reduced incident response times by **70%**, streamlining operations in Security Operations Centers (SOCs).
- **IoT threat detection** systems using AI-based anomaly detection helped manufacturers avoid costly downtime by identifying abnormal device behavior.

Cross-Domain Validation

Hybrid AI's effectiveness is further validated in other domains:

- In **finance**, hybrid models combine ML with compliance rules to detect fraud while minimizing false positives.
- In **healthcare**, AI-driven email filters successfully blocked spear phishing attacks.
- In **medical diagnostics**, Logical Neural Networks (LNNs) demonstrated improved accuracy and explainable decision pathways—paralleling the needs of cybersecurity systems.

These results affirm that hybrid AI systems not only enhance detection performance but also restore the **interpretability** and **trust** that are critical in cybersecurity environments.

Table 5: Summary of Empirical Results and Case Studies in AI-driven Cybersecurity Threat Detection

Application Area	AI Approach	Key Performance Indicators / Results	Source
Overall Threat Detection Efficiency	AI-based systems (general)	Identify cyberattacks 85% faster than traditional tools.	[72]
Cost Savings	AI-driven cybersecurity solutions (general)	Expected to save organizations over $150 billion annually by 2025.	[44]
Fraud Detection (Finance)	Hybrid AI (ML + Symbolic Rules)	Combines ML for anomaly detection with symbolic rules for compliance; prevents fraud and reduces false positives.	[25]
Phishing Detection (Healthcare)	AI-driven email filtering (ML, NLP)	Successfully blocked spear phishing emails impersonating executives; prevented credential theft and ransomware.	[73]
Malware Rule Generation	RuleLLM (LLM-based for YARA/Semgrep rules)	Generated 763 rules (452 YARA, 311 Semgrep) with 85.2% precision and 91.8% recall; outperformed SOTA tools.	[71]
Malware Detection (Deep Learning)	Dynamic deep learning + heuristic approaches	Outperformed static deep learning methods in classifying and detecting	[49]

		modern malware families.	
Intrusion Detection (ML-based IDS)	XGBoost and CatBoost on UNSW-NB15 dataset	Achieved 87% accuracy, 0.07 false positive rate, 0.12 false negative rate; highlighted key features for interpretability.	[64]
SOC Automation (SOAR platforms)	AI-driven SOAR	Led to a 70% reduction in average incident response time.	[54]
IoT Threat Detection (Manufacturing)	AI-based anomaly detection	Learned normal device communication patterns and flagged abnormal activity; avoided costly downtime.	[34]
Medical Diagnosis (Neuro-Symbolic AI)	Logical Neural Networks (LNNs)	Enhanced diagnostic accuracy and provided explainable diagnostic pathways; bridged accuracy and interpretability.	[71]

Discussion

The findings of this research highlight a significant transformation in cybersecurity threat detection, reflecting the broader evolution of artificial intelligence from rule-based systems to hybrid neuro-symbolic architectures. This progression is not merely technological—it represents a strategic response to the increasing complexity, dynamism, and unpredictability of cyber threats.

From Transparency to Adaptability: The RBES Legacy

Rule-Based Expert Systems (RBES) laid the groundwork for intelligent cybersecurity by offering **transparent, logic-driven decision-making**. Their architecture, centered around human-defined rules and explanation modules, enabled clear traceability of decisions—an essential feature for trust, accountability, and forensic analysis. However, their **static nature** and reliance on manual rule updates rendered them ineffective against **novel and polymorphic threats**, leading to scalability and maintenance challenges.

The **knowledge acquisition bottleneck**—the difficulty of extracting and formalizing expert knowledge—further limited RBES in rapidly evolving threat environments. These limitations underscored the need for systems capable of **autonomous learning and adaptation**.

Rise of Machine Learning and the "Black Box" Dilemma

Machine learning (ML) and deep learning (DL) addressed the adaptability gap by enabling systems to learn from data and detect previously unseen threats. These models excelled in pattern recognition and anomaly detection, significantly improving threat detection speed and accuracy. However, their **opaque decision-making processes** introduced the "black box" problem, where the rationale behind alerts became difficult to interpret.

This lack of explainability poses serious risks in cybersecurity, where understanding the **"why"** behind a detection is crucial for incident response, regulatory compliance, and ethical accountability. Moreover, ML/DL models are vulnerable to **adversarial manipulation**, further complicating their deployment in high-stakes environments.

Hybrid AI: Bridging Symbolic Logic and Neural Learning

Hybrid AI, particularly **neuro-symbolic systems**, emerges as a compelling solution to reconcile the strengths and weaknesses of both RBES and ML/DL. By integrating **symbolic reasoning** with **sub-symbolic learning**, hybrid systems offer:

- **Explainability** through rule-based logic
- **Adaptability** via neural networks
- **Data efficiency** by leveraging domain knowledge
- **Robustness** against noisy or incomplete inputs

This fusion mirrors **human cognition**, combining intuitive pattern recognition (System 1) with deliberate reasoning (System 2), as conceptualized by Kahneman. Hybrid AI systems can not only detect threats but also **justify their decisions**, enhancing trust and operational effectiveness.

Ethical Dimensions and Responsible AI Deployment

The ethical implications of AI in cybersecurity are profound. **Transparency**, **fairness**, and **accountability** are essential for responsible deployment. The black-box nature of ML models raises concerns about **bias**, **discrimination**, and **liability**—especially when decisions affect access, privacy, or legal outcomes.

Hybrid AI offers a pathway to mitigate these risks by embedding **human-readable logic** and **compliance rules** into the decision-making process. However, the integration of symbolic and neural components introduces new challenges in **bias detection**, **rule validation**, and **system governance**. Ongoing research and regulatory frameworks are needed to ensure that hybrid AI systems uphold ethical standards while maintaining technical excellence.

Strategic Implications for Cyber Defense

The shift toward hybrid AI reflects a **maturing understanding** of cybersecurity's demands. No single paradigm—symbolic or sub-symbolic—is sufficient in isolation. Instead, a **synergistic approach** is required, one that balances **accuracy**, **adaptability**, and **explainability**.

This evolution is driven by the relentless **arms race** between cyber attackers and defenders. As threats become more sophisticated, AI systems must evolve to remain effective, trustworthy, and resilient. Hybrid AI represents not just a technological advancement but a strategic imperative for future-proof cybersecurity.

Conclusion

The evolution of cybersecurity threat detection reflects a broader transformation in artificial intelligence—from static, rule-based systems to dynamic, learning-based models, and ultimately to hybrid AI architectures that integrate the strengths of both. Rule-Based Expert Systems (RBES) provided a foundational framework for early cybersecurity solutions, offering transparency, logical reasoning, and human-understandable decision-making. However, their inability to adapt to novel threats, coupled with the knowledge acquisition bottleneck and scalability limitations, rendered them insufficient in the face of today's rapidly evolving cyber landscape.

The emergence of machine learning (ML) and deep learning (DL) addressed these limitations by enabling systems to autonomously learn from data and detect previously unseen threats. Yet, these approaches introduced new challenges—most notably, the "black box" problem, which undermines interpretability, trust, and accountability in high-stakes cybersecurity environments.

Hybrid AI, particularly neuro-symbolic systems, offers a promising path forward. By combining the explainability and logical rigor of symbolic AI with the adaptability and pattern recognition capabilities of neural networks, hybrid systems deliver enhanced accuracy, robustness, and transparency. Empirical evidence from domains such as fraud detection, malware analysis, and intrusion detection demonstrates the practical effectiveness of these integrated approaches.

Looking ahead, the future of cybersecurity will increasingly depend on intelligent, adaptive, and explainable AI systems. Continued research is essential to refine hybrid architectures, improve knowledge representation techniques, and develop robust evaluation frameworks. Equally important is the ethical deployment of AI—ensuring fairness, mitigating bias, and establishing clear accountability mechanisms.

In an era defined by a persistent arms race between cyber attackers and defenders, hybrid AI represents not just a technological advancement but a strategic necessity—one that empowers organizations to stay ahead of threats while maintaining the trust and confidence of stakeholders.

References

1. Abu-Jamie, T. N., et al. (2021). "Diagnosing Cough Problem Expert System Using CLIPS." International Journal of Academic Information Systems Research (IJAISR) 5(5): 79-90.
2. Abu-Saqer, M. M. and S. S. Abu-Naser (2019). "Developing an Expert System for Papaya Plant Disease Diagnosis." International Journal of Academic Engineering Research (IJAER) 3(4): 14-21.
3. Abu-Saqer, M. M. and S. S. Abu-Naser (2019). "Knowledge Based System for Uveitis Disease Diagnosis." International Journal of Academic Information Systems Research (IJAISR) 3(5): 18-25.
4. Ahmed, A., et al. (2019). "Knowledge-Based Systems Survey." International Journal of Academic Engineering Research (IJAER) 3(7): 1-22.
5. Aish, M. A., et al. (2021). "Lower Back Pain Expert System Using CLIPS." International Journal of Academic Information Systems Research (IJAISR) 5(5): 57-67.
6. Alajrami, M. A. and S. S. Abu-Naser (2019). "Grapes Expert System Diagnosis and Treatment." International Journal of Academic Engineering Research (IJAER) 3(5): 38-46.
7. Albadrasawi, S. J., et al. (2023). "Development and Evaluation of an Expert System for Diagnosing Kidney Diseases." International Journal of Academic Engineering Research (IJAER) 7(6): 16-22.
8. Albanna, R. N., et al. (2023). "Colon Cancer Knowledge-Based System." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 27-36.
9. Albanna, R. N., et al. (2023). "Knowledge-Based System for Diagnosing Colon Cancer." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 27-36.
10. Al-Borno, D. F. and S. S. Abu-Naser (2023). "A Proposed Expert System for Vertigo Diseases Diagnosis." International Journal of Academic Information Systems Research (IJAISR) 7(6): 1-9.
11. Aldaour, A. F. and S. S. Abu-Naser (2019). "An Expert System for Diagnosing Tobacco Diseases Using CLIPS." International Journal of Academic Engineering Research (IJAER) 3(3): 12-18.
12. Aldaour, A. F. and S. S. Abu-Naser (2019). "Anemia Expert System Diagnosis Using Sl5 Object." International Journal of Academic Information Systems Research (IJAISR) 3(5): 9-17.
13. Aldeeb, M. H. and S. S. Abu-Naser (2023). "Breast Cancer Knowledge Based System." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 46-51.
14. Aldeeb, M. H. and S. S. Abu-Naser (2023). "Knowledge Based System for Breast Cancer Diagnosis." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 46-51.
15. Alfarra, A. H., et al. (2021). "An Expert System for Neck Pain Diagnosis." rnal of Academic Information Systems Research (IJAISR) 5(7): 1-8.
16. Al-Ghoul, M. M., et al. (2022). "Knowledge Based System for Diagnosing Custard Apple Diseases and Treatment." International Journal of Academic Engineering Research (IJAER) 6(5): 41-45.
17. Alkahlout, M. A., et al. (2021). "Expert System Diagnosing Facial-Swelling Using CLIPS." International Journal of Academic Information Systems Research (IJAISR) 5(5): 29-36.
18. Alkahlout, M. A., et al. (2021). "Expert System for Throat Problems Using SL5 Object." International Journal of Academic Information Systems Research (IJAISR) 5(5): 68-78.
19. Alkahlout, M. A., et al. (2021). "Knowledge Based System for Diagnosing Throat Problem CLIPS and Delphi languages." International Journal of Academic Engineering Research (IJAER) 5(6): 7-12.
20. Alkurd, Y. M. and S. S. Abu-Naser (2025). "A proposed accurate system for diagnosing hypertension."
21. Almadhoun, H. R. and S. S. Abu-Naser (2020). "An Expert System for Diagnosing Coronavirus (COVID-19) Using SL5." International Journal of Academic Engineering Research (IJAER) 4(4): 1-9.
22. Al-Masawabe, M. M. and S. S. Abu-Naser (2021). "Expert System for Short-term Abdominal Pain (Stomach Pain) Diagnosis and Treatment." International Journal of Academic Information Systems Research (IJAISR) 5(5): 37-56.
23. Almzainy, M. M., et al. (2023). "Development and Evaluation of an Expert System for Diagnosing Tinnitus Disease." International Journal of Academic Information Systems Research (IJAISR) 7(6): 46-52.
24. Al-Qadi, M. H., et al. (2022). "Developing an Expert System to Diagnose Tomato Diseases." International Journal of Academic Engineering Research (IJAER) 6(5): 34-40.
25. AlQatrawi, M. J., et al. (2022). "Rule Based System for Diagnosing Lablab Problems." International Journal of Academic and Applied Research (IJAAR) 6(5): 249-256.
26. Al-Qumboz, M. N. A. and S. S. Abu-Naser (2019). "Spinach Expert System: Diseases and Symptoms." International Journal of Academic Information Systems Research (IJAISR) 3(3): 16-22.
27. Al-Qumboz, M. N. A., et al. (2019). "Kidney Expert System Diseases and Symptoms." International Journal of Academic Engineering Research (IJAER) 3(5): 1-10.
28. Al-Saloul, N. J., et al. (2022). "A Knowledge Based System for Cucumber Diseases Diagnosis." International Journal of Academic Information Systems Research (IJAISR) 6(5): 29-45.
29. Alsaqqa, A. H., et al. (2021). "Knowledge Based for Tooth Problems." International Journal of Academic Information Systems Research (IJAISR) 5(5).
30. Alshawwa, I. A., et al. (2019). "An Expert System for Coconut Diseases Diagnosis." International Journal of Academic Engineering Research (IJAER) 3(4): 8-13.
31. Alshawwa, I. A., et al. (2019). "An Expert System for Depression Diagnosis." International Journal of Academic Health and Medical Research (IJAHMR) 3(4): 20-27.
32. Al-Shawwa, M. and S. S. Abu-Naser (2019). "Knowledge Based System for Apple Problems Using CLIPS." International Journal of Academic Engineering Research (IJAER) 3(3): 1-11.
33. Al-Shawwa, M. O. and S. S. Abu-Naser (2019). "A Proposed Expert System for Diagnosing Skin Cancer Using SL5 Object." International Journal of Academic Information Systems Research (IJAISR) 3(4): 1-9.
34. Altarazi, R. E., et al. (2023). "A CLIPS-Based Expert System for Brain Tumor Diagnosis." International Journal of Academic Engineering Research (IJAER) 7(6): 9-15.
35. Altayeb, J. M., et al. (2023). "Mango Pests Identification Expert System." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 17-26.
36. Aslem, Y. and S. S. Abu-Naser (2022). "CLIPS-Expert System to Predict Coriander Diseases." International Journal of Engineering and Information Systems (IJEAIS) 6(6): 89-95.
37. Dheir, I. and S. S. Abu-Naser (2019). "Knowledge Based System for Diagnosing Guava Problems." International Journal of Academic Information Systems Research (IJAISR) 3(3): 9-15.
38. Dheir, I. M., et al. (2019). "Knowledge Based System for Diabetes Diagnosis Using SL5 Object." International Journal of Academic Pedagogical Research (IJAPR) 3(4): 1-10.
39. El Kahlout, F. and S. S. Abu-Naser (2023). "Developing an Expert System to Computer Troubleshooting." International Journal of Academic Information Systems Research (IJAISR) 7(6): 16-26.
40. El Kahlout, M. I. and S. S. Abu-Naser (2019). "An Expert System for Citrus Diseases Diagnosis." International Journal of Academic Engineering Research (IJAER) 3(4): 1-7.
41. El Kahlout, M. I., et al. (2019). "Silicosis Expert System Diagnosis and Treatment." International Journal of Academic Information Systems Research (IJAISR) 3(5): 1-8.
42. Eleyan, H. A. R., et al. (2023). "An Expert System for Diagnosing West Nile virus Problem Using CLIPS." International Journal of Academic Information Systems Research (IJAISR) 7(6): 27-37.
43. El-Habibi, M. F., et al. (2022). "A Proposed Expert System for Obstetrics & Gynecology Diseases Diagnosis." International Journal of Academic Multidisciplinary Research (IJAMR) 6(5): 305-321.
44. Elhabil, B. Y. and S. S. Abu-Naser (2021). "An Expert System for Ankle Problems." International Journal of Engineering and Information Systems (IJEAIS) 5(4): 57-66.
45. Elhabil, B. Y. and S. S. Abu-Naser (2021). "An Expert System for Tooth Problems." International Journal of Academic Information Systems Research (IJAISR) 5(4).
46. Elhabil, B. Y. and S. S. Abu-Naser (2021). "Expert System for Hib Problems." International Journal of Academic Information Systems Research (IJAISR) 5(5): 5-16.
47. El-Hamarnah, H. A., et al. (2022). "Proposed Expert System for Pear Fruit Diseases." International Journal of Academic and Applied Research (IJAAR) 6(5): 237-248.
48. Elsharif, A. A. and S. S. Abu-Naser (2019). "An Expert System for Diagnosing Sugarcane Diseases." International Journal of Academic Engineering Research (IJAER) 3(3): 19-27.
49. Hamadaqa, M. H. M. and S. S. Abu-Naser (2021). "Hair Loss Diagnosis Expert System and Treatment Using CLIPS." International of Academic Engineering Research (IJAER) 5(5): 37-42.
50. Hammad, M. S., et al. (2023). "A Proposed Expert System for Diagnosis of Migraine." International Journal of Academic Engineering Research (IJAER) 7(6): 1-8.
51. Harara, F. E., et al. (2022). "Figs Knowledge Based System Disease Diagnosis and Treatment." International Journal of Academic Engineering Research (IJAER) 6(5): 41-45.
52. Harazin, D. and S. S. Abu-Naser (2023). "Developing an Expert System to Warts and Verruca." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 37-45.
53. Jamala, M. N. and S. S. Abu-Naser (2023). "Knowledge Based System for Diagnosing Lung Cancer Diagnosis and Treatment." International Journal of Academic Information Systems Research (IJAISR) 7(6): 38-45.
54. Khalil, A. J., et al. (2019). "Apple Trees Knowledge Based System." International Journal of Academic Engineering Research (IJAER) 3(9): 1-7.
55. LAfi, O. I., et al. (2022). "A Proposed Expert System for Broccoli Diseases Diagnosis." International Journal of Engineering and Information Systems (IJEAIS) 6(5): 43-51.
56. Mahmum, A. S., et al. (2023). "An Expert System for Diagnosing Whooping Cough Using CLIPS." International Journal of Engineering and Information Systems (IJEAIS) 7(6): 1-8.
57. Mansour, A. I. and S. S. Abu-Naser (2019). "Expert System for the Diagnosis of Wheat Diseases." International Journal of Academic Information Systems Research (IJAISR) 3(4): 19-26.
58. Mansour, A. I. and S. S. Abu-Naser (2019). "Knowledge Based System for the Diagnosis of Dengue Disease." International Journal of Academic Health and Medical Research (IJAHMR) 3(4): 12-19.
59. Mansour, A. I. and S. S. Abu-Naser (2021). "Expert system for the diagnosis of high blood pressure diseases." International Journal of Academic Information Systems Research (IJAISR) 5(5): 23-28.
60. Mansour, A. I., et al. (2021). "An Expert System for Diagnosing Cough Using SL5 Object." International Journal of Academic Engineering Research (IJAER) 5(6): 13-27.
61. Masri, N., et al. (2019). "Survey of Rule-Based Systems." International Journal of Academic Information Systems Research (IJAISR) 3(7): 1-23.
62. Megdad, M. M., et al. (2022). "Mint Expert System Diagnosis and Treatment." International Journal of Academic Information Systems Research (IJAISR) 6(5): 22-28.
63. Mettleq, A. S. A. and S. S. Abu-Naser (2019). "A Rule Based System for the Diagnosis of Coffee Diseases." International Journal of Academic Information Systems Research (IJAISR) 3(3): 1-8.
64. Mettleq, A. S. A., et al. (2019). "Expert System for the Diagnosis of Seventh Nerve Inflammation (Bell's palsy) Disease." International Journal of Academic Information Systems Research (IJAISR) 3(4): 27-35.
65. Mettleq, M. S. A. and S. S. Abu-Naser (2025). "Expert System for the Diagnosis of coffee Diseases."
66. Murad, W. F. and S. S. Abu-Naser (2023). "An Expert System for Diagnosing Mouth Ulcer Disease Using CLIPS." International Journal of Academic Engineering Research (IJAER) 7(6): 30-37.
67. Okasha, S. M., et al. (2022). "A knowledge Based System for Diagnosing Persimmon Diseases." International Journal of Academic and Applied Research (IJAAR) 6(6): 53-60.
68. Qanoo, F. N., et al. (2023). "A CLIPS-Based Expert System for Heart Palpitations Diagnosis." International Journal of Academic Information Systems Research (IJAISR) 7(6): 10-15.
69. Qaoud, A. N. and S. S. Abu-Naser (2023). "Developing an Expert System to Diagnose Malaria." International Journal of Academic Information Systems Research (IJAISR) 7(6): 9-18.
70. Qunoo, M. M. and S. S. Abu-Naser (2025). "A proposed accurate system for diagnosing low vision." International Journal of Engineering and Information Systems (IJEAIS) 9(6): 23-32.
71. Radwan, H. I., et al. (2022). "A Proposed Expert System for Passion Fruit Diseases." International Journal of Academic Engineering Research (IJAER) 6(5): 24-33.
72. Sababa, R. Z., et al. (2022). "A Proposed Expert System for Strawberry Diseases Diagnosis." International Journal of Engineering and Information Systems (IJEAIS) 6(5): 52-66.
73. Salman, F. M. and S. S. Abu-Naser (2019). "Expert System for Castor Diseases and Diagnosis." International Journal of Engineering and Information Systems (IJEAIS) 3(3): 1-10.
74. Salman, F. M. and S. S. Abu-Naser (2019). "Thyroid Knowledge Based System." International Journal of Academic Engineering Research (IJAER) 3(5): 11-20.
75. Salman, F. M. and S. S. Abu-Naser (2020). "Expert System for COVID-19 Diagnosis." International Journal of Academic Information Systems Research (IJAISR) 4(3): 1-13.
76. Samhan, L. F., et al. (2021). "Expert System for Knee Problems Diagnosis." International Journal of Academic Information Systems Research (IJAISR) 5(4): 59-66.
77. Tantawi, J. and S. S. Abu-Naser (2023). "Knowledge-Based System for the Diagnosis of Flatulence." International Journal of Academic Engineering Research (IJAER) 7(6): 23-29.
78. Wishah, N. D., et al. (2023). "Developing a Knowledge-Based System for Diagnosis and Treatment Recommendation of Neonatal Diseases Using CLIPS." International Journal of Academic Engineering Research (IJAER) 7(6): 38-50.