International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Artificial Intelligence and Its Impact on Building an Effective Cybersecurity Protection Framework

Dr. Yussri Awad Abdalla

Director of the Legal Department, Al Fozan Holding Group, AL Khobar, Kingdom of Saudi Arabia.

Abstract

This study examines the impact of artificial intelligence on building an effective cybersecurity protection framework from a legal and analytical perspective in light of Saudi legislation, international standards, and risk governance frameworks. The research addresses the central question of whether the existing legal and regulatory frameworks are sufficient to govern the use of artificial intelligence in cybersecurity while maintaining an appropriate balance between cybersecurity requirements, personal data protection, transparency, and legal accountability.

The study adopts both the descriptive-analytical and comparative approaches by analyzing the Saudi regulatory framework and comparing it with leading international standards, particularly ISO/IEC 27001:2022, ISO/IEC 27005:2022, ISO/IEC 42001:2023, the NIST Cybersecurity Framework (CSF 2.0), the NIST AI Risk Management Framework (AI RMF 1.0), and the European Union Artificial Intelligence Act (AI Act).

The findings demonstrate that artificial intelligence has become a fundamental component of modern cybersecurity systems, significantly enhancing threat prevention, early threat detection, incident response, cyber risk management, and business continuity. The study further concludes that the effectiveness of AI-enabled cybersecurity does not depend solely on technological advancement but rather on the existence of an integrated governance framework encompassing legal compliance, institutional governance, and comprehensive risk management. Moreover, the research reveals that Saudi Arabia has established an advanced regulatory environment for cybersecurity and personal data protection that is broadly consistent with contemporary international approaches. Nevertheless, the rapid evolution of AI applications necessitates a more specialized legal framework addressing AI governance, legal liability, algorithmic transparency, human oversight, and AI risk assessment.

The study concludes that establishing an effective cybersecurity framework in the era of artificial intelligence requires the integration of advanced technologies, sound legal regulation, and robust risk governance. Such integration ensures the responsible use of artificial intelligence, strengthens cybersecurity resilience, safeguards fundamental rights and freedoms, and supports the objectives of digital transformation and Saudi Vision 2030.

Keywords: Artificial Intelligence; Cybersecurity; Risk Governance; Saudi Legislation; Personal Data Protection; AI Governance; Cyber Risk Management.

Introduction

Artificial intelligence has become one of the most influential variables shaping the structure of contemporary cybersecurity. Its role is no longer confined to improving the technical performance of digital systems; it has extended to reshaping the very philosophy of protection itself — from traditional protection based on reacting after an attack has occurred, to proactive protection grounded in analysis, prediction, pattern learning, and the detection of anomalous behavior before it develops into a fully-fledged cyber incident. It is against this backdrop that the importance of studying artificial intelligence emerges — not as an abstract technical tool, but as an influential element in building a legal, regulatory, and institutional system capable of protecting cyberspace, managing its risks, and ensuring a balance between digital innovation and the protection of public and private rights and interests.

The importance of this subject is further heightened by the accelerating pace of digital transformation in the Kingdom of Saudi Arabia, accompanied by an expansion in digital government services, banking applications, cloud platforms, big data, and AI-enabled services. The Saudi legislator has kept pace with this transformation through a multi-tiered regulatory system, most notably the Essential Cybersecurity Controls issued by the National Cybersecurity Authority, which set the minimum cybersecurity requirements for protecting the informational and technical assets of national entities against internal and external risks — reflecting the shift of cybersecurity from being a purely internal technical matter to becoming a general governance and regulatory obligation (National Cybersecurity Authority, 2018).

This research is likewise connected to the Personal Data Protection Law in the Kingdom, given that many AI applications rely on the collection, analysis, and processing of data — making the lawfulness of processing, privacy protection, data security, and transparency fundamental issues when AI is used to protect cybersecurity. The Personal Data Protection Law was issued pursuant to Royal Decree No. M/19 of 2021, amended by Royal Decree No. M/148 of 2023, and entered into force on 14 September 2023, confirming that data protection has become a legal pillar inseparable from cybersecurity and artificial intelligence (Kingdom of Saudi Arabia, 2021, as amended 2023).

At the international level, the subject of this research intersects with recent frameworks and standards, most notably the NIST AI Risk Management Framework issued in 2023, which adopts an approach founded on risk management, trust, and governance, and ISO/IEC 42001:2023 on AI management systems, which sets out requirements for establishing, implementing, and improving a responsible management system for the use of artificial intelligence within organizations (NIST, 2023; ISO, 2023).

Accordingly, the research problem does not lie merely in demonstrating the capacity of artificial intelligence to detect or respond to cyberattacks, but rather in determining its legal and regulatory impact on building an effective cybersecurity system, and the extent to which Saudi legislation and international standards are adequate to regulate its use and ensure that it does not turn from a protective tool into a new source of risk — whether through algorithmic bias, privacy violations, weak transparency, or the difficulty of establishing legal liability when harm occurs.

Accordingly, and before turning to the substantive analysis, this research addresses its subject through five main parts: a review of the relevant literature and the gap this study addresses; a statement of the research questions, objectives, and methodology; the conceptual framework of artificial intelligence and cybersecurity; the role of artificial intelligence in building an effective protection system and managing cyber risk; and the legal framework and governance of AI use in light of Saudi legislation and international standards.

Literature Review

Scholarship on the intersection of artificial intelligence and cybersecurity has developed along two largely separate tracks that this study seeks to bring together. The first track is technical and managerial, concerned with how AI improves threat detection and response. Kshetri (2023) surveys the operational opportunities and limitations of AI in cybersecurity, concluding that detection gains are consistently offset by new attack surfaces created by the AI systems themselves. Foundational technical treatments of the underlying methods — deep learning (Goodfellow, Bengio, & Courville, 2016) and the broader discipline of computer security (Bishop, 2019; Stallings, 2024; Whitman & Mattord, 2022) — establish that AI-based defensive tools are not categorically different from other software artifacts for purposes of security assurance, testing, and governance, a premise this study adopts in treating AI as a governed technical system rather than an autonomous actor.

The second track is legal and regulatory, concerned with how AI should be governed once deployed. Cath (2018) frames AI governance as a problem of reconciling ethical, legal, and technical demands that existing regulatory instruments were not designed to address jointly. Brundage et al. (2020) argue for verifiable claims and independent auditing as mechanisms of AI trustworthiness — a position this study draws on directly in Section Three's discussion of independent review mechanisms. At the level of public international law, Schmitt's Tallinn Manual 2.0 (2017) remains the leading account of how existing rules of international law apply to cyber operations, though, as this study notes in Section Three, it does not extend to the AI-specific questions of algorithmic liability and explainability that have since become central to the field.

Within the Saudi and Arabic-language literature specifically, several studies bear directly on this research. Alzahrani (2024) examines AI governance and cybersecurity regulation in Saudi Arabia and, like this study, finds an advanced institutional foundation; that work, however, is framed primarily as a policy and governance survey and does not undertake a systematic doctrinal comparison against ISO, NIST, and EU instruments, nor does it address the allocation of civil liability for automated security decisions — the comparative and liability-focused analysis that the present study contributes. Al-Awadi (2024) offers a comparative legal study of AI and cybersecurity centred on Kuwaiti and regional frameworks, providing a useful external comparator but not a Saudi-specific regulatory reading. Al-Shammari (2023) analyses cyber risk management under the Saudi Essential Cybersecurity Controls without addressing artificial intelligence, while Al-Sayed (2023) and Al-Otaibi (2024) address, respectively, the general legal protection of cybersecurity and the governance of artificial intelligence in the Kingdom as separate subjects rather than as an integrated legal problem.

Read together, this literature establishes three things clearly: that AI meaningfully strengthens technical cyber-defence capability; that AI governance has become a distinct and active field of legal scholarship; and that Saudi Arabia possesses a developing regulatory base for both cybersecurity and AI. What the existing literature has not yet done — and what this study undertakes — is to (i) systematically map Saudi cybersecurity and data-protection instruments against ISO/IEC 27001:2022, ISO/IEC 27005:2022, ISO/IEC 42001:2023, NIST CSF 2.0, the NIST AI RMF, and the EU AI Act on a common set of governance functions; (ii) trace how each stage of the AI-enabled cybersecurity lifecycle (prevention, detection, response, recovery) generates a distinct legal question of liability, transparency, or data protection; and (iii) translate that comparative reading into concrete, sequenced recommendations for Saudi legislative development. This is the gap the present study addresses.

Research Questions, Objectives, and Methodology

Research Questions

Building on the problem identified above, this study is organised around one central question and three subsidiary questions:

1. Are the existing Saudi legal and regulatory frameworks sufficient to govern the use of artificial intelligence in cybersecurity while maintaining an appropriate balance between security requirements, personal data protection, transparency, and legal accountability?
2. What legal and governance functions does artificial intelligence perform, and what new categories of risk does it introduce, across the prevention, detection, response, and recovery stages of the cybersecurity cycle?
3. To what extent does Saudi regulation converge with, or diverge from, the leading international standards and frameworks governing AI-enabled cybersecurity?
4. Where the current framework is insufficient, what specific legislative and institutional measures would close the gap without unduly constraining innovation?

Research Objectives

Correspondingly, the study aims to: (i) clarify the legal and functional nature of artificial intelligence as distinct from its technical description; (ii) analyse the practical contribution of AI across each stage of the cybersecurity cycle and the legal questions that contribution raises; (iii) map the Saudi regulatory framework against the relevant international standards on a shared set of governance criteria; and (iv) formulate specific, sequenced recommendations for developing Saudi AI-governance legislation.

Methodology

This study adopts a doctrinal (black-letter) legal method combined with a comparative-legal method, consistent with standard approaches in legal and regulatory scholarship on emerging technology. The doctrinal component analyses the text and stated purpose of the relevant Saudi legal instruments — principally the Essential Cybersecurity Controls (ECC-1:2018 and ECC-2:2024), the Personal Data Protection Law (Royal Decree No. M/19 of 2021, as amended by Royal Decree No. M/148 of 2023), and the AI Ethics Principles issued by SDAIA — to determine the scope of obligations they impose on the use of artificial intelligence in cybersecurity. The comparative component places these instruments alongside six international reference frameworks selected on two criteria: (a) direct relevance to either information-security management or AI-system governance, and (b) current force or standing as of the time of writing. The frameworks compared are ISO/IEC 27001:2022, ISO/IEC 27005:2022, ISO/IEC 42001:2023, the NIST Cybersecurity Framework 2.0 (2024), the NIST AI Risk Management Framework 1.0 (2023), and Regulation (EU) 2024/1689 (the EU AI Act). These six were selected because each is either the leading international management-system standard in its domain (ISO), the most widely referenced national cybersecurity or AI risk framework outside the EU (NIST), or the first comprehensive statutory AI regime with extraterritorial relevance (the EU AI Act); frameworks with narrower sectoral scope or non-binding status (for example, the OECD AI Principles or ISO/IEC 23894 on AI risk management) were treated as secondary and referenced only where they add a distinct point not already covered by the primary six.

The comparison is organised around five governance functions common to the frameworks reviewed: (1) risk identification and assessment, (2) governance and accountability, (3) transparency and explainability, (4) human oversight, and (5) incident response and continuous improvement. This five-function structure,

applied consistently in Section Three, is what allows the frameworks to be compared on a like-for-like basis despite their different legal status and drafting style.

The analysis is bounded to legal and regulatory instruments in force, or formally adopted and scheduled to enter into force, as of January 2026; subsequent regulatory developments, including any further amendment to the ECC or supplementary EU AI Act implementing measures, fall outside this study's scope. The study is doctrinal and comparative in nature and does not include empirical fieldwork, structured interviews with regulators or practitioners, or a systematic review of enforcement decisions or case law, since no body of published Saudi judicial or administrative decisions specifically addressing AI-related cybersecurity liability was identified at the time of writing. This is a genuine limitation, discussed further below, and is one of the principal directions this study proposes for future empirical research.

Section One

The Conceptual Framework of Artificial Intelligence and Cybersecurity

The conceptual foundation constitutes the necessary starting point for understanding the relationship between artificial intelligence and cybersecurity, since conflating technical and legal concepts may disturb the analysis and lead to portraying artificial intelligence either as an absolute remedy for all cyber risks or as a threat in its own right. In truth, artificial intelligence possesses a dual nature: on the one hand, it is an advanced tool for enhancing protection, while on the other hand it may become a source of new risks when used without governance, without legal controls, or without effective human oversight.

This section therefore does not aim to provide a detailed technical explanation of artificial intelligence, but rather to clarify its legal and functional nature within the digital environment, then to define the concept of cybersecurity and its dimensions, and finally to analyze the integrative relationship between the two as a relationship grounded in risk management, the protection of digital assets, the assurance of service continuity, and the preservation of trust in electronic transactions.

1.1 The Nature of Artificial Intelligence, Its Evolution, Types, and Role in the Digital Environment

Artificial Intelligence, broadly speaking, is the ability of computer systems to execute activities that previously needed some form of cognition, analysis, learning, prediction of behavior or conclusion drawing, including pattern recognition, understanding of language, classification of data, prediction of behavior, and drawing conclusions from large amounts of data. But this general definition is not adequate for the legal context, since the law is not about technical capacity per se, but about its impact, who is responsible for it, how much it is legally permissible, and who controls its design, operation and output.

In this view, artificial intelligence is a technical system that has legal consequences; it is not restricted to following a set of instructions, but can learn from data, build predictive models, and suggest or make decisions on its own in some of its forms. This is the thing that sets it apart from conventional software systems: A traditional system will function based on set rules, while artificial intelligence — especially machine-learning systems — will work by analyzing the patterns within data and creating indicators based on them. Its legal value is not only as a digital tool, however, but as a tool that can affect choices, danger assessments, and liability allocation.

This understanding is reflected in recent international standards: ISO/IEC 42001:2023, which focuses on the managerial and governance aspects of AI, and not just the technical aspects, outlines requirements to establish, implement, maintain and improve an AI management system in organizations, ensuring that AI-based systems are used responsibly (ISO, 2023). This is important for the current research because it validates the view that AI cannot be left to a standalone technical team, but needs to be integrated within

an institutional structure that establishes policies, responsibilities, supervisory frameworks, and risk management.

This approach aligns with the NIST AI Risk Management Framework, published on 26 January 2023, which is a trust- and risk-based framework based on core functions that include governance, mapping, measurement, and management. The significance of this framework is that it does not consider AI to be a completely neutral technology, but rather a tool that can have social, legal, and regulatory impacts that need to be managed and evaluated continuously (NIST, 2023).

The AI Ethics Principles laid down by the Saudi Data and Artificial Intelligence Authority (SDAIA) are a guiding principle for the responsible use of artificial intelligence in the Saudi context. Among these principles, there are those which are directly applicable to cybersecurity, most notably, privacy, security, reliability, and safety, and transparency, and explainability: An intelligent protection system that is not transparent, is not explainable, is not privacy, is not security, is not reliable, and is not safe is no longer a means of protection but a source of risk (SDAIA, 2023).

From simple rule-based applications in the digital world to more advanced applications using machine learning, deep learning, predictive analytics, natural language processing, computer vision, and self-adaptive systems. The benefits of such advancements are particularly evident in the realm of cybersecurity, where AI can handle vast amounts of security data that humans simply can't process as quickly or effectively. Modern systems produce millions of digital logs, alerts, login attempts, network traffic logs and usage patterns daily, and a traditional security department can't effectively process these without the help of intelligent analytical tools.

But the significance of AI is not just about how fast it can go, it's also about its ability to identify unusual patterns. Modern attacks are not necessarily obvious or direct, but can be subtle actions, unknown logins, an attempt to elevate privileges, or the use of a known account in an unusual way. If used correctly, AI can thus be able to detect normal and abnormal behavior, and trigger an alert before the damage is done.

However, this raises an important legal question: can an intelligent system be able to detect risk, and is this enough for the use of the system to be considered lawful? The answer is that while technical efficiency is important, so are also a number of controls, such as the lawfulness of data collection, the purpose of data processing, the scope of data retention, the limits of monitoring employees or customers, the explainability of automated decisions, and the possibility of human intervention when it is needed. For this reason, the application of artificial intelligence in the field of cybersecurity should be subject to the principle of proportionality, which means that the methods for processing, monitoring and analysis of data must be appropriate to the type of risk that is being prevented, and that cybersecurity protection is not to be used as a justification for illegal data collection or for the surveillance of people.

This issue becomes more significant in the context of the Saudi Personal Data Protection Law which has set the basic rule that the processing of personal data can only be done on a legal basis and the change in the purpose of processing or increase in scope of processing is subject to legislative restrictions. Compliance with data-protection controls is not just a separate and independent obligation, but a fundamental element of the lawfulness of cybersecurity protection itself, as the applications of AI in cybersecurity could involve analyzing user data, login records, behavioral patterns, and connection locations (SDAIA, 2023).

By way of comparison, it may be said that the international trend leans toward incorporating artificial intelligence within the institutional governance system rather than confining it to the technical domain alone. The NIST Cybersecurity Framework 2.0, issued in February 2024, added a "Govern" function to

the cybersecurity functions, reflecting the fact that cybersecurity management is no longer limited to identifying, protecting, detecting, responding, and recovering, but now begins with the determination of policies, responsibilities, third-party risk, supply chains, and senior-management decisions (NIST, 2024). This addition is directly significant to the subject of this research, as it makes artificial intelligence part of enterprise-level cyber risk management, rather than merely an operational tool within the security operations center.

As for the Kingdom of Saudi Arabia, the Essential Cybersecurity Controls issued by the National Cybersecurity Authority establish an institutional conception of cybersecurity through requirements relating to governance, risk management, asset protection, human-resources security, third-party security, and business continuity. The introduction to the Controls states that they set the minimum cybersecurity requirements for protecting informational and technical assets against internal and external threats — a formulation that reveals the Saudi legislator's adoption of the notion of cybersecurity as an integrated system rather than a standalone technical measure (National Cybersecurity Authority, 2018).

On this basis, artificial intelligence may be viewed as a supporting tool within this system: it does not replace legal and regulatory controls, nor does it dispense with policies, procedures, and human oversight, but rather enhances the institution's ability to implement those controls with greater efficiency. Threat prediction, vulnerability analysis, incident classification, malware detection, and user-behavior analysis are all applications that can help fulfil the statutory obligations relating to protection, detection, response, and recovery.

The corresponding risk, however, lies in excessive reliance on artificial intelligence. Intelligent systems may err in classification, generate false alarms, fail to detect a novel attack for which they were not trained, or be affected by biased or incomplete data. Attackers may also deliberately deceive an intelligent system through what are known as adversarial attacks on AI models, whereby the model is fed inputs specifically crafted to confuse it or push it toward an incorrect decision. Here arises the need not to rely on technical trust alone, but to build a legal and governance system that establishes clear responsibilities for the design, testing, operation, and performance monitoring of the system, and for intervention when errors occur.

Artificial intelligence in the field of cybersecurity must therefore be understood in light of three interconnected facts: first, that it is an effective tool for enhancing digital protection; second, that it is a potential source of legal, technical, and ethical risk; and third, that its true value is realized only when it is integrated within an institutional governance framework linking technology, law, and risk management. This is what grounds the research's core hypothesis: that the effectiveness of artificial intelligence in protecting cybersecurity is not measured by its technical capability alone, but by the extent to which it is subject to a legal and regulatory system that guarantees lawful, secure, and responsible use.

It thus becomes clear that artificial intelligence is neither a substitute for cybersecurity nor merely a tool within it, but a transformative factor in its very structure. It shifts cybersecurity from a rigid protection model to an intelligent protection model, from after-the-fact response to proactive prevention, and from limited manual analysis to broad-scale predictive analysis. Yet this transformation is only commendable when coupled with the necessary legal safeguards, risk governance, transparency, accountability, and respect for data protection and digital rights.

This leads us, in the following subsection, to set out the concept of cybersecurity itself, its objectives, principles, and legal dimensions, so as to clarify the environment within which artificial intelligence operates as a tool for protecting that system, rather than as an end independent of it.

1.2 The Concept of Cybersecurity, Its Objectives, Principles, and Legal Dimensions

AI is the intelligent tool that can be used to defend the digital environment and cybersecurity is the content area where the tool works and the purpose it is meant to serve. Thus, before studying the impact of AI on cybersecurity protection, it is necessary to define what is meant by cybersecurity and outline its goals, principles, and legal aspects; Cybersecurity is no longer a purely technical term related to the protection of devices and networks, but rather a composite term that is related to the protection of the state, the digital economy, personal data, critical infrastructure, business continuity, and trust in electronic transactions.

The general sense of Cybersecurity is defined as the protection of systems, networks, software, and data from attacks, breaches, un-authorized access, disruption, damage or misuse. However, this technical definition is not enough for legal analysis because cybersecurity in today's context does not limit itself to the prevention of intrusion, but also encompasses the establishment of an integrated system of risk management, the delineation of responsibilities, adherence with regulatory controls, the protection of digital rights and the maintenance of the continuity of essential services.

From this comes the difference between information security and cybersecurity. While information security is concerned with the protection of information in terms of confidentiality, integrity and availability, cybersecurity is a much broader concept because it deals with the protection of the whole digital space, from networks and infrastructure to operating systems, applications, connected devices, digital services, technical supply chains, and the data moving through it. Cybersecurity, therefore, is a higher level of security than information security, as the threat is a variable, continuously renewed and trans-border threat.

The National Cybersecurity Authority (NCA) has confirmed that the main purpose of these Controls is to establish the minimum cybersecurity requirements for protecting the informational and technical assets of national entities and to reduce the risks of cyber threats that may result from internal and external threats, in the Kingdom of Saudi Arabia (KSA) (National Cybersecurity Authority, 2018). This approach shows that cybersecurity in the Saudi system is not considered a technical process but rather an institutional obligation that is organized and has the objective of protecting assets, minimizing risk, and ensuring compliance.

This general dimension is confirmed by the official website of the National Cybersecurity Authority which indicates that the Authority is the national reference for cybersecurity affairs and that it is tasked with enhancing cybersecurity to safeguard the vital interests of the state, national security, sensitive infrastructure, priority sectors, and government services and activities (National Cybersecurity Authority, 2024b). The statement shows that cybersecurity is not just a matter for one institution anymore, but it is one of the most important ways to protect the supreme interests of the state, the stability of public services, and the integrity of the digital economy.

The NIST Cybersecurity Framework 2.0 has taken a forward-looking approach to the concept of cybersecurity as organized risk management at the international level. The Framework provides a structure for cybersecurity responsibilities and activities in six core functions: Govern, Identify, Protect, Detect, Respond, and Recover (NIST, 2024). The "Govern" function added to the second version marks the start of cybersecurity not with technology, but with the establishment of policies and responsibilities, oversight and the involvement of senior management in managing cyber risk.

The findings from the comparison of the Saudi and international conceptions show a significant convergence in terms of the view of cybersecurity as a governance and risk system. The Saudi Essential Cybersecurity Controls emphasize minimum requirements, asset protection, risk reduction, and

institutional compliance, whereas the NIST CSF 2.0 emphasizes organizing security outcomes in interconnected functions starting with governance and ending with recovery. The distinction between the two is that the Saudi Controls are applicable to entities subject to them in the Kingdom and the NIST Framework is a widely used international guidance framework. Both methods, however, share a core concept: the need to manage cybersecurity, not reactively, but as an ongoing institutional system.

The first goal of cybersecurity is confidentiality, which means that data, systems or information should not be made available to unauthorized persons. In the era of Artificial Intelligence, this principle becomes even more significant, as data is the indispensable raw material for training, operating and optimizing intelligent models. A breach in data security, data use without lawful purpose and data access without permission may not only result in a privacy violation but also in data being fed into AI models that is stolen, misleading or illegal, which could compromise future security decisions.

The second goal is data and system integrity, which means that data is not modified, lost or corrupted without permission. This objective has a strong legal aspect since data integrity is related to the validity of digital evidence, the soundness of electronic records, the probative value of electronic transactions and the reliability of decisions made by intelligent systems. When data is altered, an intelligent system can make decisions on a false premise, falsely label something as a threat or miss a threat, or make a wrong recommendation, resulting in material or legal damage.

The third goal is availability, meaning that systems, services and data are available when required. It is no longer an operational issue but one of law and regulation because the failure of digital services in critical areas (banking, healthcare, energy, telecommunication, government services) can have a widespread impact starting with the institution and reaching the society and the economy. Cybersecurity is therefore tightly integrated with business continuity, disaster-recovery planning, incident management, and the responsibility to report breaches, if required.

ISO/IEC 27001 embodies this conception by presenting an information security management system that sets requirements for establishing, implementing, maintaining, and improving an information security management system within organizations. The significance of this standard lies in the fact that it does not treat security as a mere set of technical tools, but as a management system based on policies, risk assessment, controls, and continual improvement (ISO, 2022). ISO/IEC 27001 thus converges with the modern trend in cybersecurity toward transforming protection from disparate measures into an integrated management system.

From a legal standpoint, it may be said that cybersecurity rests on four principal dimensions. The first is the preventive dimension, consisting of the obligations imposed by legislation, regulations, and standards on entities to protect their systems and data before an incident occurs. This dimension is clearly evident in the Essential Cybersecurity Controls, which do not wait for a breach to occur before responsibility is triggered, but require entities to build advance protective capabilities encompassing governance, risk management, asset protection, identity and access security, systems security, network security, and third-party security.

The second is the oversight dimension, consisting of the existence of a competent national authority that sets policies and controls and monitors compliance. In the Kingdom, the National Cybersecurity Authority performs this role as the competent entity and national reference for cybersecurity affairs. This regulatory centrality gives cybersecurity a sovereign and institutional character, as it prevents the fragmentation of authority and helps unify the minimum national requirements.

The third is the remedial dimension, consisting of incident response, damage containment, recovery from

breaches, and the determination of responsibilities after an incident occurs. This dimension acquires special importance when artificial intelligence is used, since response may take place automatically or semi-automatically — such as isolating a device, suspending an account, halting a process, or raising an alert to the security operations center. An important legal question arises here: if an intelligent system takes an erroneous action that causes harm, who bears liability? Is it the entity that owns the system, the provider, the developer, the party responsible for its operation, or the person who approved the automated decision without review? This question confirms that cybersecurity cannot be separated from the rules of legal liability.

The fourth is the rights-based dimension, consisting of the need to reconcile the protection of digital systems with the protection of individual rights, particularly the right to privacy and personal data protection. Cybersecurity may require monitoring network traffic, analyzing login records, tracking usage patterns, and possibly using advanced behavioral-analysis tools. These means, however, must not turn into absolute surveillance or disproportionate processing of personal data. Accordingly, the lawfulness of cybersecurity protection is not measured solely by its effectiveness in preventing attacks, but also by the extent to which it respects the principles of necessity, proportionality, lawfulness, and transparency.

On this basis, it may be said that cybersecurity in the modern digital environment is no longer built on the idea of the "protective wall" alone, but on a multi-layered system. Cyber threats have become more complex, potentially originating from outside or inside, from individuals or organized groups, from contracting parties or technical supply chains, and may target data, systems, reputation, or service continuity. Effective cybersecurity therefore requires combining technical controls, legal policies, institutional governance, training, awareness, auditing, and risk management.

From a comparative perspective, it is observed that the Saudi framework tends toward detailed regulation through mandatory controls directed at national entities and vital sectors, whereas international standards, such as NIST and ISO, tend to offer flexible normative frameworks applicable across different environments. This flexibility, however, does not diminish their importance, as it allows organizations to build cybersecurity programs suited to their size, the nature of their risks, and their level of maturity. The Saudi Controls, by contrast, are distinguished by their direct connection to the national environment and the requirements of protecting vital interests and sensitive infrastructure.

This difference is reflected in the role of artificial intelligence. In the Saudi environment, artificial intelligence must be used in cybersecurity within the framework of compliance with national controls, including governance controls, risk-management controls, and asset-protection controls. In the international environment, it may be integrated within maturity and management frameworks such as NIST CSF, ISO/IEC 27001, and ISO/IEC 42001. The conclusion is that artificial intelligence is not effective merely by being incorporated into security systems; it becomes effective when it is positioned within a clear governance system that determines: why is it used? what data does it process? who monitors it? who bears responsibility for it? how is its accuracy assessed? and how are its decisions reviewed?

It thus becomes apparent that cybersecurity represents an integrated legal and regulatory environment in which data-protection rules, institutional compliance, civil and criminal liability, governance, risk management, and critical-infrastructure protection all intersect. This makes the incorporation of artificial intelligence into this field a matter of dual nature: on the one hand, a practical necessity for confronting evolving threats, and on the other, a legal challenge requiring the regulation of use, the prevention of abuse, the determination of liability, and the assurance of respect for rights.

Accordingly, a correct understanding of cybersecurity leads to a fundamental conclusion: artificial intelligence builds an effective cybersecurity protection system only when it forms part of a broader structure founded on governance, risk management, compliance, the protection of rights, and continuous improvement. Its use outside this framework may instead produce the opposite result, potentially creating excessive trust in automated systems, expanding the scope of unlawful data processing, weakening human oversight, or rendering the institution unable to explain its security decisions in the event of a dispute or investigation.

In light of the foregoing, cybersecurity may be defined for the purposes of this research as: a legal, regulatory, technical, and administrative system aimed at protecting systems, networks, data, and digital services from cyber risks and threats, ensuring their confidentiality, integrity, and availability, achieving compliance, governance, and business continuity, while respecting digital rights and protecting personal data.

This definition is the most appropriate for the subject of this research, as it does not confine cybersecurity to its technical dimension, nor does it separate it from artificial intelligence, nor does it overlook its legal dimension. It also allows the following subsection to analyze the integrative relationship between artificial intelligence and cybersecurity, and to show how artificial intelligence can move cybersecurity from traditional protection to intelligent protection, while the need for legal controls and institutional governance remains.

The development artificial intelligence has undergone is not limited to increases in computational power or data-processing speed, but extends to a fundamental change in its underlying philosophy of operation. Intelligent systems have moved from the stage of traditional software based on fixed logical rules (rule-based systems) to systems that learn continuously from data and rebuild their statistical and algorithmic models as the volume of data they process increases. This transformation is what has enabled artificial intelligence to cope with complex and changing environments such as the cybersecurity environment, characterized by the rapid evolution of attack methods, the diversity of threat sources, and the constant daily emergence of new vulnerabilities.

From a technical standpoint, scholarship distinguishes between several types of artificial intelligence, each differing in the scope of its capabilities and the legal functions attaching to it. The first type is Narrow Artificial Intelligence, the most widespread at present, designed to perform a specific task with high efficiency, such as detecting malware, analyzing network traffic, recognizing faces, or classifying malicious emails. This type cannot transfer to tasks outside the scope of the programming or training for which it was designed.

The second type is Artificial General Intelligence (AGI), a theoretical model at present, which aims to comprehensively simulate human cognitive abilities such that it can learn, reason, and adapt to various situations without needing to be reprogrammed for each new task. Should this type materialize in the future, it would raise unprecedented legal issues relating to civil liability, criminal liability, the evidentiary weight of its decisions, and the limits of its autonomy in decision-making (Russell & Norvig, 2021).

To this must be added what has become known in recent years as Generative Artificial Intelligence, which relies on large language models and generative models to produce text, images, program code, and synthetic data. This type has opened wide horizons in the field of cybersecurity, being used to generate attack scenarios, simulate breaches, prepare security reports, and analyze vulnerabilities. At the same time, however, it has also come to be exploited by attackers to craft more convincing phishing messages,

develop malware in a semi-automated manner, or produce fake content that is difficult to detect — making artificial intelligence itself simultaneously a tool of defense and a tool of attack (NIST, 2023).

The legal value of artificial intelligence is therefore not tied to its type alone, but to the manner of its use, its degree of autonomy, and the nature of the decision it produces. The greater a system's reliance on self-learning and the lesser the human intervention in its outputs, the greater the need for legal controls ensuring that those outputs can be reviewed, that responsibility for them can be determined, and that they do not become decisions incapable of explanation or challenge.

Among the most notable features distinguishing artificial intelligence from other traditional software is its reliance on data as the principal source of learning. No matter how sophisticated an algorithm may be, it cannot produce accurate results unless it relies on data that is correct, sufficient, and representative of reality. Technical scholarship accordingly holds that the quality of artificial intelligence is measured not only by the efficiency of the algorithm, but also by the quality of the data on which it was trained. This carries a highly significant legal consequence, as any defect in the lawfulness, integrity, accuracy, or manner of collection of the data may lead to unlawful outcomes, even where the algorithm itself is technically sound.

For this reason, data governance has become an integral part of AI governance. It is not enough for an institution to comply with cybersecurity controls; it must also ensure that the data used to train intelligent systems has been collected lawfully, that it is accurate and up to date, and that its use is consistent with the purpose for which it was collected — consistent with the principles established by the Saudi Personal Data Protection Law (Kingdom of Saudi Arabia, 2021, as amended 2023) and with the AI Ethics Principles issued by the Saudi Data and Artificial Intelligence Authority (SDAIA, 2023).

From another perspective, artificial intelligence relies on algorithms, which constitute the beating heart of the intelligent system. An algorithm is not merely a set of programming instructions, but the mathematical or logical model that transforms data into results, predictions, or decisions. This has given rise to a legal issue known in recent literature as algorithmic transparency, the essence of which is that many AI systems — particularly deep-learning models — produce accurate decisions without offering a clear explanation of how those decisions were reached, a phenomenon known as the "black box problem." This problem poses a genuine challenge in the field of cybersecurity, since decisions to block a user, suspend a service, or classify an activity as a cyberattack may carry serious legal and financial consequences and must therefore be capable of explanation and review.

The European Union has attached particular importance to this issue in its recent legislation on artificial intelligence, adopting an approach based on classifying AI systems according to their level of risk, while imposing stricter obligations on high-risk systems, including requirements of transparency, data quality, human oversight, risk management, and technical documentation, so as to strike a balance between innovation and the protection of fundamental rights (European Union, 2024).

The growing role of artificial intelligence in the digital environment is evident in its contribution to supporting the digital-transformation processes undertaken by both government and private institutions alike. Artificial intelligence has become an essential element in the management of smart cities, e-government services, digital commerce, financial technology (FinTech), smart health services, cloud computing, the Internet of Things, and digital supply chains. As a consequence of this expansion, any malfunction in intelligent systems may extend its impact across multiple sectors simultaneously, which redoubles the importance of subjecting these systems to legal oversight and institutional governance.

In the Kingdom of Saudi Arabia, this trend is consistent with the objectives of Saudi Vision 2030, which has made digital transformation and a data- and AI-based economy central pillars of national development. The Saudi Data and Artificial Intelligence Authority (SDAIA) was established as the national entity responsible for leading this transformation and setting national policies relating to data and artificial intelligence, so as to ensure the responsible and secure use of these technologies (SDAIA, 2023).

The impact of artificial intelligence on the digital environment is not confined to enhancing operational efficiency, but extends to reshaping the very concept of risk management. Under traditional systems, risk management relied heavily on periodic human analysis, historical reports, and assessments conducted at widely spaced intervals. Under artificial intelligence, however, risk management has become a continuous, dynamic process based on real-time data analysis, the prediction of risks before they materialize, and the instantaneous updating of risk levels in line with emerging variables. This transformation is one of the most important reasons for the growing interest in artificial intelligence in the field of cybersecurity, as it shifts protection from post-incident response to proactive prevention.

A comparison between the Saudi approach and international standards reveals a clear convergence in regarding artificial intelligence as a technology that must be subject to governance and risk management, rather than operating in isolation from them. The Saudi principles for responsible artificial intelligence emphasize the values of fairness, transparency, security, and privacy, while ISO/IEC 42001:2023 and the NIST AI RMF focus on integrating risk management and institutional oversight throughout all stages of the AI system's lifecycle. This convergence reveals a growing global trend toward subjecting artificial intelligence to institutional governance standards, given its far-reaching legal, economic, and social impact.

The researcher takes the view that the development artificial intelligence has undergone necessitates reconsidering many traditional legal concepts, particularly those relating to liability, oversight, evidence, and risk management. Artificial intelligence is no longer merely an assistive program, but has become an active participant in the decision-making process — which requires that legislation not be limited to regulating its use, but must extend to establishing an integrated governance framework that determines responsibilities, ensures transparency, and achieves a balance between encouraging innovation and protecting cybersecurity and digital rights. The true value of artificial intelligence, therefore, lies not in its abstract technical capabilities, but in the extent to which it is integrated within a legal and institutional system capable of directing those capabilities toward achieving effective, secure, and sustainable cybersecurity.

1.3 The Integrative Relationship between Artificial Intelligence and Cybersecurity and Its Impact on Building an Effective Digital Protection System

The study of the relationship between artificial intelligence and cybersecurity is no longer confined to demonstrating the use of the former in the service of the latter, but has come to concern the analysis of the structural relationship that has emerged between them amid global digital transformation. Cybersecurity, in its traditional form, relied on fixed defensive means such as firewalls, anti-malware systems, intrusion-detection systems, and signature-based detection rules — tools that proved effective against traditional threats but that now face growing challenges with the evolution of AI-driven cyberattacks, multi-stage attacks, polymorphic malware, zero-day attacks, and advanced persistent threats (APTs). Confronting these threats is therefore no longer possible by relying on traditional protective means alone, but now requires integrating AI technologies into the very core of the security system.

This integration does not mean that artificial intelligence has become a substitute for cybersecurity, but rather that it has become one of its most important strategic components. Cybersecurity remains the overarching framework aimed at protecting digital systems, data, and infrastructure, while artificial intelligence represents the advanced means that enhances this system's ability to achieve its objectives more efficiently, through real-time data analysis, risk prediction, the detection of abnormal patterns, and continuous learning from past incidents. The relationship between them is therefore one of functional complementarity, not one of substitution or replacement.

This understanding is reflected in recent literature issued by the (US) National Institute of Standards and Technology (NIST), which regards artificial intelligence as a key element in enhancing cyber risk management, while emphasizing the need to subject it to governance and risk management throughout its entire lifecycle, rather than merely assessing its technical performance (NIST, 2023). This approach confirms that the success of artificial intelligence in the security field is measured not only by its ability to detect attacks, but by its ability to do so reliably, in an explainable manner, and subject to legal accountability.

The integrative relationship between artificial intelligence and cybersecurity is manifested in the fact that each depends on the other to achieve its objectives. On the one hand, artificial intelligence needs a secure cyber environment that protects the data from which it learns, preserves the integrity of its algorithms, and prevents tampering with intelligent models or data poisoning. On the other hand, cybersecurity needs the capabilities of artificial intelligence in order to cope with the enormous volume of digital data and to analyze the millions of security events generated daily within organizations, which the human element cannot analyze within an appropriate time frame.

This relationship is most clearly evident in Security Operations Centers (SOCs), where intelligent systems have come to undertake the triage of security alerts, their classification by severity, the correlation of similar incidents, the analysis of abnormal user behavior, and the provision of real-time recommendations to incident-response teams. The human role in this environment is no longer confined to executing manual procedures, but has shifted to overseeing the decisions proposed by intelligent systems, verifying their soundness, and making the final decision in cases requiring human intervention.

This integration produces a fundamental shift in the philosophy of cybersecurity. Under the traditional model, the protection process began after an attack was discovered, followed by analysis of its causes and then attempts at containment. Under artificial intelligence, however, the security system has come to rely on predictive analytics, which aims to detect the initial indicators of an attack before it is completed, through the analysis of user behavior, communication patterns, data traffic, and the relationships between different security events. Cybersecurity thus shifts from a stage of after-the-fact response to a stage of proactive prevention — a transformation representing one of the most significant developments in this field.

From a legal standpoint, this transformation requires reconsidering the concept of due care and due diligence in the field of cybersecurity. It is no longer sufficient for an institution to rely on the minimum traditional means of protection where the nature of its activity and the scale of its risks require the use of more advanced tools. Institutions that manage critical infrastructure, process personal data on a large scale, or provide digital financial services may become legally required to adopt technical measures proportionate to their level of risk, including the use of AI technologies where these have been established as best practices within the sector.

This approach is consistent with the philosophy adopted by the Essential Cybersecurity Controls issued by the National Cybersecurity Authority in the Kingdom, which are based on risk management and the application of appropriate controls according to the nature of assets and threats, rather than the mere application of uniform procedures across all entities (National Cybersecurity Authority, 2018). The legal standard here is not the use of any particular technology, but the achievement of a level of protection commensurate with the nature of the risks involved.

The integrative relationship between artificial intelligence and cybersecurity is also manifested in the field of Enterprise Risk Management. Cybersecurity was previously regarded as the responsibility of the information-technology department alone, but today it has become one of the strategic risks subject to the oversight of the board of directors and audit and risk committees. Within this framework, artificial intelligence has become a means of supporting strategic decision-making, by providing real-time analyses of risk levels, measuring the probability of their occurrence, and assessing their potential impact on business continuity and regulatory compliance.

By way of comparison, it is observed that international standards — particularly ISO/IEC 42001:2023 — focus on building an AI management system directly linked to risk governance, whereas ISO/IEC 27001:2022 focuses on an information security management system, while the NIST Cybersecurity Framework 2.0 combines governance, risk management, and cybersecurity within a single framework. Linking these standards together reveals that the international trend is moving toward integrating artificial intelligence into the institutional governance system, so that it becomes one of the elements of risk management rather than merely an independent technical tool (ISO, 2022; ISO, 2023; NIST, 2024).

In the Kingdom of Saudi Arabia, this integration is consistent with the national orientation toward building a digital economy based on data and artificial intelligence, while simultaneously preserving the security of cyberspace and the protection of personal data. This is evident in the institutional coordination between the mandates of the National Cybersecurity Authority and the Saudi Data and Artificial Intelligence Authority, the former being responsible for setting national cybersecurity controls, and the latter for leading the national transformation in the field of data and artificial intelligence and establishing the principles governing their responsible use. This institutional integration reflects the Saudi legislator's recognition that protecting the digital environment cannot be achieved through cybersecurity alone, nor through artificial intelligence alone, but through coordination between the two within a unified legal and regulatory framework.

The relationship between artificial intelligence and cybersecurity, however, is not solely one of positive integration; it also involves a relationship of mutual influence on the risk side. Artificial intelligence used to strengthen cybersecurity may itself become a target of attacks, just as attackers have come to use AI technologies to develop more sophisticated means of intrusion, such as producing phishing messages that are difficult to distinguish from genuine ones, developing malware capable of automatically altering its behavior, or executing attacks based on the analysis of victim behavior. Artificial intelligence has therefore come to represent, simultaneously, both a means of protection and a means of attack, which necessitates building a legal framework that keeps pace with this development and addresses its emerging risks.

Consequently, the relationship between artificial intelligence and cybersecurity cannot be reduced to the former merely serving the latter; rather, it is a relationship of continuous interaction in which each technology influences the other. Every advance in artificial intelligence leads to the development of cyber-defense tools, but at the same time provides attackers with more sophisticated means of attack, producing a permanent technical and legal race between the means of protection and the means of intrusion.

This study argues that this continuous interaction necessitates a shift from the concept of "technical cybersecurity" to the concept of "intelligent cybersecurity" — a concept founded on integrating artificial intelligence, governance, risk management, compliance, data protection, and legal liability within a single, interconnected system. It is not enough for an institution to possess advanced AI tools if those tools are not subject to clear policies, oversight controls, periodic risk assessment, and effective human supervision. Nor is it sufficient for cybersecurity legislation to exist if it fails to accommodate the novel legal characteristics of intelligent systems.

Taken together, this shows that the relationship between artificial intelligence and cybersecurity represents a relationship of strategic integration aimed at building a digital protection system capable of predicting risks, responding to them, learning from them, and continuously improving, while simultaneously observing the requirements of governance, transparency, accountability, and the protection of rights. This conclusion forms the foundation from which the research will proceed in Section Two, moving from conceptual grounding to an analytical study of the practical applications of artificial intelligence in protecting cybersecurity and managing its risks, with a more detailed account of their legal and regulatory implications.

Section Two

The Role of Artificial Intelligence in Building an Effective Cybersecurity Protection System

Preamble

Having set out the conceptual foundation of artificial intelligence and cybersecurity, and having analyzed the integrative relationship between them in the preceding section, methodological sequence requires moving on to examine the practical impact of artificial intelligence in building an effective cybersecurity protection system. The development that AI technologies have undergone over the past decade has not been limited to improving data-processing efficiency or automating procedures, but has brought about a fundamental transformation in the very philosophy of cybersecurity; security systems have moved from a model relying on human intervention and reaction after an incident occurs, to an intelligent model based on risk prediction, continuous data analysis, and proactive preventive action.

This transformation responds to the changing nature of cyber threats, which have become characterized by rapid evolution, complexity, globality, and the ability to bypass traditional means of protection. Attacks no longer rely on standard methods that can easily be detected through fixed signature rules, but now employ advanced techniques such as multi-stage attacks, polymorphic malware, AI-powered attacks, targeted phishing, and attacks on digital supply chains. In confronting this changing environment, artificial intelligence has emerged as one of the most important tools capable of analyzing the enormous volume of security data, detecting abnormal patterns, predicting malicious behavior, and responding to incidents at a speed exceeding human capability.

The importance of artificial intelligence in this field, however, is not measured merely by its technical capacity to detect or contain attacks, but by the extent of its contribution to building a cybersecurity system characterized by efficiency, resilience, sustainability, and compliance with legal and regulatory controls. An effective security system is not built on technology alone, but on the integration of technology, governance, risk management, regulatory compliance, personal data protection, and institutional oversight — a point emphasized by the Essential Cybersecurity Controls issued by the National Cybersecurity Authority, the NIST Cybersecurity Framework 2.0, and ISO/IEC 27001 and ISO/IEC 42001 (National Cybersecurity Authority, 2018; NIST, 2024; ISO, 2022; ISO, 2023).

This section accordingly addresses the role of artificial intelligence in building an effective cybersecurity protection system through three interconnected subsections: the first is devoted to examining the role of artificial intelligence in the prevention of cyber threats and their early detection; the second addresses its role in incident response, risk management, and business continuity; and the section concludes with a third subsection addressing the technical, legal, and ethical challenges of using artificial intelligence in this field.

2.1 The Role of Artificial Intelligence in the Prevention of Cyber Threats and Attacks and Their Early Detection

Prevention of cyber threats is the first step in the cybersecurity cycle, designed to stop an attack or make it less likely to be successful before it happens and before it becomes a cyber incident. This stage was mostly based on conventional protection mechanisms like firewalls, anti-virus software, signature-based intrusion detection systems, and identity authentication and privilege management processes until recently. However, these tools tend to be limited to known attacks or malware and are less effective for new or mutated attacks, or attacks using new or unobserved methods.

This limitation has spurred the quest for technologies that can more effectively handle the ever-changing landscape of cyber threats, and AI has come to the fore as a technology that can transform the "PRIOR KNOWLEDGE" detection concept into a "BEHAVIORAL ANALYSIS/ RISK PREDICTION" approach to detection. Instead of looking for a particular digital signature of malware, it is now possible to study the actions of users, devices and networks and identify any unusual or abnormal activity, including activity that has not been previously seen. This is a qualitative change in the concept of cyber prevention, allowing detection of a threat at an early stage, before its impact is spreading or its consequences are felt.

In this area, AI uses several methods, primarily machine learning, deep learning, big-data analytics and behavior analytics, which allow security systems to process millions of security events in real time, correlate them with each other and provide risk indicators based on them, which is hard to do with traditional human methods. NIST AI Risk Management Framework reaffirms the need to ensure that the use of AI in security contexts is also tied to risk management and that it is continually evaluated for reliability and explainability of outcomes and bias (NIST, 2023).

The application of these technologies, however, is clear as they can identify known and unknown attacks. Traditional systems are based on rules and if they are faced with an attack that is not in the rules, there is only a small chance of it being detected. AI-powered systems, on the other hand, don't just look for the digital signature of an attack, but the behavior that's occurring with it — like an unusual surge in data traffic, multiple login attempts, abrupt shifts in user privileges, or unexpected communication between devices in the network. These indicators can be aggregated and analyzed to determine the probability of a threat before a specific determination of the threat is made.

This is one of the biggest examples of proactive cyber security, which is the dominant theme in the current literature. The security system doesn't wait for an attack to happen and then try to stop it, it tries to foresee the attack and reduce its potential. This is in line with the philosophy of the risk management approach behind the Essential Cybersecurity Controls in the Kingdom of Saudi Arabia, which emphasizes that entities must identify and evaluate risks and implement preventative controls before they occur (National Cybersecurity Authority, 2018).

One of the most notable examples of this change is what's called User and Entity Behavior Analytics (UEBA), one of the most prevalent AI applications in today's security landscape. This system creates a model of the typical behavior of each user, device or application in the organization, based on a wealth of

historical data, including typical login times, devices used, geographical locations, types of applications accessed, amount of data exchanged and patterns of connection to internal and external networks.

The moment this reference model is created, the system can compare any new activity with the expected normal activity, calculate the deviation in behavior and evaluate the probability of the activity being related to a cyber threat. If, for instance, an employee logs on to the system from within the Kingdom of Saudi Arabia during working hours, and the system subsequently records a login attempt from another country at night, and then attempts to access databases the employee had never before accessed, the intelligent system doesn't look at each event individually, but rather at the entire sequence of events as a series of indicators that could be a sign of account compromise or misuse.

The legal significance of this technology is that it reaches a greater level of the care that the legislator of institutions which process sensitive data or manage critical infrastructure expects. A sophisticated behavioral analysis system can more readily identify illegal use in its early stages, which decreases the chance of harm occurring or it escalating. The use of these technologies in the security system can then become a way of meeting the statutory obligations in terms of data protection and risk management, instead of being a technical possibility.

The role of artificial intelligence is not confined to analyzing user behavior, but extends to detecting zero-day attacks, among the most dangerous types of cyberattacks, as they exploit vulnerabilities unknown to the system developer or security-software providers, and consequently have no digital signatures that traditional protection systems can recognize. For this reason, such attacks have posed an extremely serious challenge to traditional security systems.

Artificial intelligence, however, offers a different approach to this problem: rather than searching for the digital signature of the attack, it focuses on the behavior resulting from it. If a piece of software begins to perform unusual operations, attempts to access files not normally accessed, or establishes unjustified external connections, the intelligent system can detect these patterns and classify them as risk indicators, even if it has never previously encountered the software itself. AI-based systems have thus become better able to deal with novel attacks compared with systems relying on signature rules alone.

The European Union Agency for Cybersecurity (ENISA) has confirmed that the use of artificial intelligence in threat detection has become one of the main trends in the development of defensive capabilities, given the ability it provides to analyze big data, detect abnormal patterns early, and reduce the time required to detect cyber incidents (ENISA, 2023).

The impact of artificial intelligence also extends to the development of Security Information and Event Management (SIEM) systems, which increasingly rely on machine-learning techniques to analyze the millions of security logs arriving from the various components of the network. Rather than merely collecting logs and presenting them to analysts, modern systems have become capable of prioritizing events, excluding false positives, correlating scattered events into a single scenario, and suggesting the most probable explanation of an incident.

The importance of this development is evident when one considers that the security operations centers of major institutions may receive hundreds of thousands of security alerts daily, making complete reliance on the human element impractical. Studies indicate that a large proportion of these alerts are of no significance or represent false alarms, leading to the depletion of security analysts' time and delays in responding to genuine incidents. Artificial intelligence has therefore contributed to raising the efficiency of these systems by classifying alerts according to their likely severity and directing analysts' attention to priority incidents.

The development of artificial intelligence has likewise been linked to the emergence of Security Orchestration, Automation and Response (SOAR) platforms, which are no longer confined to issuing alerts, but have become capable of executing initial actions automatically, such as isolating a suspected device, blocking a connection to an external server, temporarily suspending a user account, opening a ticket for the security incident, or notifying the response team. These procedures shorten the time between the detection of a threat and its containment, known in the security literature as reducing the Mean Time to Detect (MTTD) and the Mean Time to Respond (MTTR).

Among the recent applications is also what is known as Extended Detection and Response (XDR) platforms, which gather cybersecurity data from various sources — such as endpoints, networks, email, and cloud services — and then use AI techniques to merge and analyze this data as a single security environment. This improves the ability to detect composite attacks that move across more than one technical environment simultaneously, among the most prevalent types of attacks at present.

From a legal perspective, this development raises a question concerning the lawfulness of taking automated security actions. If an intelligent system suspends an employee's account or disables a particular service based on automated analysis, and it later transpires that the decision was based on a false alarm, who bears liability for the resulting harm? Is the decision issued by the intelligent system a purely technical decision, or an administrative decision for which the institution bears full responsibility?

A section of legal scholarship takes the view that liability remains vested in the entity owning the system, since artificial intelligence does not enjoy an independent legal personality but is regarded as a tool used by the institution in performing its obligations. The use of artificial intelligence therefore does not relieve the institution of its duty of oversight, nor does it remove its responsibility to verify the soundness of automated decisions; rather, it requires the institution to establish review mechanisms and human intervention where necessary — a point emphasized by modern AI governance principles (ISO, 2023; NIST, 2023).

This conforms with the philosophy of the Kingdom of Saudi Arabia which emphasizes on governance, risk management, separation of responsibilities, ongoing monitoring, and that technical solutions are not the only solution. The Controls do not mandate any given technology; instead, they set out to ensure that entities reach a certain level of effectiveness in meeting the statutory requirements for asset protection, risk management, incident response, and business continuity, with the flexibility to use any technology that allows that to be accomplished.

A comparison between Saudi system and international standards reveals that both systems consider artificial intelligence as a tool to improve cybersecurity, not as an objective, confirming that both view AI as a technology that can be leveraged for better cybersecurity, not as an end goal. However, Saudi legislation has a stronger emphasis on institutional compliance and ensuring the minimum national controls, while international standards, such as the NIST AI RMF and ISO/IEC 42001, tend to focus on the lifecycle of the AI system, risk management, the transparency and explainability of the AI system's outputs and the presence of effective human oversight during the AI system's operational stages. The two approaches do not compete with each other, but rather complement each other: the Saudi Controls set the regulatory framework and international standards give the optimum technical and managerial approaches to the implementation of the controls.

This analysis leaves the impression that the actual power of artificial intelligence is not to take the place of the human factor, but to redefine the roles in the security system. Although AI has an amazing ability to process data, analyze patterns, and detect early warning signs of risk, it still cannot understand all the

legal, regulatory, and ethical aspects related to a security decision. An effective cybersecurity protection system is one that is smart, human, institutional and legal, with AI used to augment, not replace, the institution's legal obligations in security decision making.

Based on this, the application of AI in threat prevention/early detection marks a paradigm shift in cyber security, from a security model based on fixed rules, to a more dynamic and data-driven one, and from a reactive security model, detecting threats after they have actually happened, to a proactive one, predicting threats before they materialize. But the key to this transformation is the need for an integrated legislative and regulatory framework that will ensure the proper use of these technologies, establish the responsibilities of the parties involved, and create a balance between technical efficiency and the principles of lawfulness, transparency and accountability — and which logically leads to the second subsection on the use of artificial intelligence in cyber incident response, risk management and business continuity.

2.2 The Role of Artificial Intelligence in Cyber Incident Response, Risk Management, and Business Continuity

If prevention and early detection represent the first stage in the cybersecurity cycle, cyber incident response represents the most sensitive stage, as it deals with risk after it has begun to materialize and aims to limit its effects, prevent its expansion, and restore systems and services in the shortest possible time. This stage has undergone radical development thanks to AI technologies, which have shifted response from a model relying primarily on human intervention to a model based on real-time analysis, algorithm-supported decision-making, and the execution of initial actions in a semi-automated manner, while human oversight remains an essential element in decisions of major legal or operational impact.

This transformation is linked to the nature of modern cyber incidents, which now develop within minutes or seconds, such that delay in decision-making may double the extent of harm in a manner difficult to remedy. Ransomware attacks, distributed denial-of-service (DDoS) attacks, attacks on cloud infrastructure, and attacks on digital supply chains all require an immediate response that does not allow complete reliance on traditional manual procedures. Artificial intelligence has accordingly become a pivotal element in building rapid-response capabilities, by analyzing an incident the moment it occurs, determining its scope, estimating its severity, and proposing appropriate containment measures.

Among the most prominent manifestations of this development is the reliance of security operations centers on Automated Incident Response, which allows a set of preventive actions to be executed automatically once predetermined conditions are met. When the intelligent system detects unlawful activity with a high degree of confidence, it can isolate the infected device from the network, disable the suspected account, block communication with a specific electronic address, or halt a data-transfer process — all before the response team reaches the scene of the incident. This action reduces the time required to contain the threat, one of the most important indicators of the efficiency of modern cybersecurity systems (NIST, 2024).

This capacity to take semi-automated action, however, raises a legal issue concerning the lawfulness of an institution's reliance on automated decision-making in cases that may restrict individuals' rights or disrupt their interests. If an intelligent system leads to the suspension of a customer's account or the disruption of a digital service based on an erroneous assessment, the question does not concern only the accuracy of the algorithm, but the institution's liability for this decision, and the extent of its commitment to effective human oversight and subsequent review mechanisms. Recent literature accordingly emphasizes that automation in the field of cybersecurity should be governed automation, not absolute automation that executes decisions without oversight or review (ISO, 2023).

The importance of artificial intelligence is also evident in Cyber Incident Management, where the intelligent system can aggregate data arriving from multiple sources, correlate it, and determine the point of origin, the path of the attack's spread, the digital assets affected, and the temporal priority for addressing each part of the incident. This analytical capacity provides the response team with a comprehensive picture of the incident, helping it make more accurate decisions regarding prioritization, resource allocation, and the determination of urgent measures.

This role acquires special importance in environments relying on cloud computing or distributed digital infrastructure, where the impact of an incident may extend to several data centers, applications, or service providers at once. In such cases, manual analysis becomes incapable of grasping the volume of data and the speed of its change, while artificial intelligence can trace the relationship between different events, identify hidden links between them, and propose the most probable scenario for the incident's development.

From the perspective of risk management, artificial intelligence has brought about a shift in the concept of Continuous Risk Assessment. In traditional models, risk management relied on periodic assessments that might be conducted once or twice a year, whereas in modern digital environments risks now change on a daily, indeed instantaneous, basis, as a result of the emergence of new vulnerabilities, software updates, changing threat patterns, or the expanding use of cloud services and the Internet of Things. Artificial intelligence has therefore become capable of continuously reassessing the level of risk, based on actual data arriving from the operational environment, rather than merely on theoretical assumptions.

This approach is consistent with ISO 31000:2018, which views risk management as a continuous process integrated with all of the organization's activities, rather than a separate or seasonal procedure. Artificial intelligence enables the practical application of this concept by monitoring Key Risk Indicators, analyzing them in real time, and issuing alerts when acceptable risk levels are exceeded, thereby enhancing senior management's ability to make decisions based on accurate and up-to-date data (ISO, 2018).

In the Saudi context, these practices are consistent with the methodology of the National Cybersecurity Authority, which makes risk management one of the central pillars of the national Controls and requires entities to identify, analyze, address, and periodically review risks. The use of artificial intelligence, however, adds a new dimension to this methodology, as it transforms risk management from a process relying on periodic reports into a continuous, dynamic process that interacts with the digital environment moment by moment (National Cybersecurity Authority, 2018).

The role of artificial intelligence is not confined to managing current risks, but extends to Predictive Risk Analytics, through the analysis of historical data, the derivation of patterns, and the estimation of the probabilities of particular scenarios materializing. This is one of the most important applications of artificial intelligence in cybersecurity, as it helps institutions move from managing incidents after they occur to managing risks before they materialize, consistent with the modern governance philosophy founded on anticipation rather than reaction.

Among the other important applications is the use of artificial intelligence in Business Continuity and Disaster Recovery. Once an incident has been contained, priority shifts to restoring services to their normal state as quickly as possible, while preserving data integrity and minimizing operational losses. At this stage, artificial intelligence can determine the priorities for restoring systems, analyze the interdependencies between services, estimate the time required to restore each service, and propose the best restart scenario, so as to achieve the least possible impact on the institution's core operations.

This role acquires added importance in vital sectors such as banking, energy, healthcare, and telecommunications, where a service outage of even a few minutes may cause significant financial losses or threaten the safety of individuals. Business continuity has therefore ceased to be a merely operational matter and has become a legal and regulatory obligation linked to the institution's duty to protect the interests of customers and beneficiaries and ensure the stability of essential services.

By way of comparison, it is observed that the NIST Cybersecurity Framework 2.0 treats Respond and Recover as two separate functions within the cybersecurity cycle, whereas the Saudi Controls integrate these two dimensions within the requirements of incident management, business continuity, and risk management. This difference reflects a variation in presentation style rather than substance, as both approaches converge on the necessity of having clear response and recovery plans, determining responsibilities, testing plans periodically, and drawing on lessons learned after each incident.

One aspect worth noting is that artificial intelligence, despite its high efficiency in supporting incident response, cannot replace the human element in all cases. Decisions carrying major legal consequences — such as reporting an incident to regulatory authorities, disclosing a personal-data breach, suspending essential services, or taking actions affecting the rights of customers or employees — continue to require human review and a legal judgment going beyond what an algorithm can analyze. Best practice therefore does not consist in dispensing with the human element, but in building a model of Human–AI Collaboration, whereby artificial intelligence undertakes rapid analysis while the human retains the final decision on matters of legal and strategic significance.

In this study's assessment, the true value of artificial intelligence at the incident-response and risk-management stage lies not merely in the speed of data processing, but in its ability to transform cybersecurity into a system of continuous learning. Every cyber incident becomes a new source of knowledge, every intrusion attempt contributes to improving analytical models, and every response experience increases the accuracy of predicting future risks. This development, however, must not lead to excessive reliance on intelligent systems, since the security decision ultimately remains a legal and administrative decision for which the institution bears responsibility before regulatory and judicial authorities. Artificial intelligence should therefore operate within a clear governance framework founded on transparency, explainability, human oversight, and continuous review, so as to achieve a balance between technical efficiency and the requirements of lawfulness.

The upshot is that artificial intelligence has contributed to a qualitative leap in cyber incident response, risk management, and business continuity, by moving from reactive to proactive management, from partial handling to comprehensive analysis, and from reliance on human expertise alone to the integration of human expertise with intelligent analysis. This leads logically to the third subsection, which addresses the other side of this development — namely, the technical, legal, and ethical challenges of using artificial intelligence in protecting cybersecurity — highlighting the need for a legislative and governance framework that ensures the responsible and secure use of these technologies.

2.3 Technical, Legal, and Ethical Challenges of Using Artificial Intelligence in Cybersecurity Protection

Although artificial intelligence has become one of the most important tools relied upon by institutions and states in strengthening cybersecurity, its use is not without complex challenges extending beyond the technical dimension to legal, regulatory, and ethical ones. This is because artificial intelligence, unlike traditional software, relies on continuous learning from data and the dynamic reshaping of its analytical models, which makes its outputs more difficult to predict, explain, and control. The success of artificial

intelligence in protecting cybersecurity is therefore measured not only by its ability to detect threats or respond to incidents, but also by the extent to which it is subject to an integrated system of governance, compliance, and legal accountability.

The first of these challenges lies in the susceptibility of AI systems themselves to cyberattacks. The traditional conception viewed artificial intelligence as a means of protecting systems, but technical developments have revealed that intelligent models themselves have become targets of attack. A machine-learning model may be subjected to what is known as data poisoning, whereby misleading or manipulated training data is introduced, altering the model's behavior and weakening its ability to distinguish between legitimate and malicious activity. It may also be subject to adversarial attacks, which rely on making slight modifications to input data in order to mislead the model and produce erroneous results, even though the modification may be imperceptible to the human eye (NIST, 2023). ENISA's dedicated analysis of the intersection between cybersecurity and artificial intelligence reaches a similar conclusion, noting that AI systems introduce a distinct attack surface that traditional information-security controls were not designed to address (ENISA, 2024).

These attacks are particularly dangerous in the field of cybersecurity, as they may disable the defensive system's very function. If an attacker succeeds in deceiving a malware-detection model, a behavior-analysis system, or an intelligent response platform, the institution may become more vulnerable to breach despite possessing advanced security tools. Securing digital systems alone is therefore no longer sufficient; it has become necessary also to secure the AI models themselves, verify the integrity of the data on which they were trained, and periodically test their resistance to adversarial attacks.

From a legal standpoint, this challenge raises the question of the duty of care incumbent upon the institution developing or using artificial intelligence. If a system is found to have been manipulated due to weak protective measures or the absence of security tests specific to the model, liability may extend to the operating entity, not to the attacker alone, provided that a failure to apply recognized professional standards is established.

The second challenge lies in the problem of transparency and explainability, one of the most debated issues in contemporary legal scholarship. Many AI systems, particularly deep-learning models, can arrive at highly accurate results without providing a clear explanation of how those results were reached. This phenomenon is termed the "black box" problem, since the decision-making mechanism within the model is unclear, even to its developers in some cases.

The importance of this issue is evident in the field of cybersecurity when an intelligent system takes decisions affecting the legal position of individuals or institutions, such as suspending a user's account, blocking a service, classifying a particular activity as a cyberattack, or filing a report with the competent authorities. In such cases, it is not sufficient for the decision to be technically correct; it must also be explainable and reviewable, so that its lawfulness can be verified and the aggrieved party enabled to object to it where appropriate.

International standards have responded to this issue: ISO/IEC 42001:2023 makes governance, transparency, documentation, and risk management core elements of the AI management system, while the NIST AI Risk Management Framework emphasizes the need for intelligent systems to be trustworthy — a concept encompassing reliability, transparency, explainability, fairness, privacy, security, and accountability (ISO, 2023; NIST, 2023).

The third challenge lies in the protection of personal data and privacy. Artificial intelligence in the field of cybersecurity relies on analyzing enormous quantities of data, including login records, network traffic,

communications, usage patterns, behavioral data, and possibly even geolocation data or biometric data in some applications. Such data may include personal information subject to statutory protection, requiring entities using artificial intelligence to comply with the legal controls relating to the collection, processing, retention, and disclosure of data.

In the Kingdom of Saudi Arabia, the Personal Data Protection Law constitutes the principal regulatory framework governing these matters, providing that the processing of personal data must be based on a lawful statutory basis, must be linked to a specified purpose, and must be limited to what is necessary to achieve that purpose, together with the adoption of measures ensuring the protection of data from unlawful access or unauthorized use (Kingdom of Saudi Arabia, 2021, as amended 2023). The use of artificial intelligence in cybersecurity accordingly does not relieve institutions of compliance with these principles, but rather makes them more important given the volume of data processed by intelligent systems.

The fourth challenge lies in determining legal liability for decisions issued by AI systems. If an intelligent system causes an erroneous decision leading to financial loss, service disruption, data disclosure, or infringement of an individual's rights, who bears liability? Is it the institution owning the system? The developing company? The data provider? The party responsible for operating the system? Or is liability apportioned among them according to the degree of each one's contribution to the error?

This subject remains the object of extensive scholarly debate; nevertheless, the predominant trend in comparative legislation tends not to grant artificial intelligence an independent legal personality, such that liability in principle continues to rest with the natural or legal persons who designed, operated, or used the system. This means that reliance on artificial intelligence is not a ground for exemption from liability, but may instead increase the institution's duty to verify the soundness of the system, monitor its performance, document its operations, and enable the review of its decisions.

Here emerges the concept of Algorithmic Accountability, which has become one of the core principles of AI governance. It requires the institution to be able to explain how the system operates, the basis on which it takes decisions, the procedures for verifying its soundness, and the identification of the person or entity responsible for its oversight. This concept represents a natural extension of institutional-governance principles, as it prevents responsibility from being concealed behind the complexity of algorithms or their apparent autonomy.

The fifth challenge lies in Algorithmic Bias. Artificial intelligence learns from the data on which it is trained, and if that data is incomplete, unbalanced, or biased, the model may produce unfair or inaccurate decisions. In the field of cybersecurity, such bias may increase false-alarm rates directed at a particular category of users, or reduce the assessed level of risk in situations requiring urgent intervention, which may affect the integrity of the security system and the fairness of the measures taken.

For this reason, the AI Ethics Principles issued by the Saudi Data and Artificial Intelligence Authority emphasize the need to achieve fairness, reduce bias, and ensure equality in the design and operation of intelligent systems, so as to strengthen societal trust in the use of these technologies (SDAIA, 2023).

Among the challenges of a strategic nature is also Overreliance on AI, as the significant success of intelligent systems may lead to reduced reliance on human expertise, weaken employees' capacity for independent decision-making, or result in the neglect of developing their analytical skills. Consequently, an institution may become exposed to significant risk if the intelligent system malfunctions, is breached, or misjudges a situation. International best practices accordingly emphasize the need to keep humans in the loop (Human-in-the-Loop), particularly in decisions carrying significant legal, financial, or security consequences.

A comparison between Saudi regulation and international standards reveals a clear convergence in general principles, even though the regulatory tools differ. The Kingdom of Saudi Arabia has focused on building an integrated national cybersecurity system, issuing the Essential Cybersecurity Controls, the Personal Data Protection Law, and the principles of responsible artificial intelligence, whereas international standards have tended toward developing specialized normative frameworks, such as ISO/IEC 42001 and the NIST AI RMF, to regulate the lifecycle of AI systems and manage their risks. This convergence reveals a unified global trend founded on the premise that artificial intelligence cannot perform its role in protecting cybersecurity unless it is subject to rigorous legal and institutional governance.

This study's central contention is that the challenges raised by AI applications in the field of cybersecurity should not be regarded as a justification for limiting the use of this technology, but rather as an impetus for developing the legislative and regulatory environment. The greater artificial intelligence's capacity to take consequential decisions, the greater the need for legal rules regulating its design, operation, testing, and review, and determining liability for its outputs. The future success of the cybersecurity system will therefore depend not only on the sophistication of algorithms, but equally on the sophistication of the legal frameworks governing them — which leads logically to Section Three, devoted to examining the legal framework and governance of AI use in protecting cybersecurity, through an analysis of Saudi legislation, international standards, and mechanisms for developing the regulatory environment so as to achieve the safe and responsible use of this technology.

Section Three

The Legal Framework and Governance of AI Use in Cybersecurity Protection

Preamble

Practical applications have demonstrated that the success of artificial intelligence in strengthening cybersecurity does not depend on the efficiency of algorithms or the sophistication of technical infrastructure alone, but relies to a greater extent on the existence of a legal and regulatory framework that defines the limits of using these technologies, sets out the responsibilities of the entities operating them, establishes the controls necessary to ensure their sound use, protects rights and freedoms, and achieves a balance between security requirements and the requirements of data protection, privacy, and transparency. The widespread proliferation of AI applications has shifted the debate from the question of whether this technology can be used, to the question of how its use should be legally regulated. The real challenge is no longer possessing AI technologies, but building a legislative system capable of accommodating their legal implications, controlling their risks, and determining the responsibilities of the parties connected to them.

Within this framework, the Kingdom of Saudi Arabia has moved to build an integrated regulatory system for cybersecurity, data, and artificial intelligence, through the establishment of the National Cybersecurity Authority and the Saudi Data and Artificial Intelligence Authority, the issuance of the Essential Cybersecurity Controls and the Personal Data Protection Law, alongside the adoption of national policies and principles for responsible artificial intelligence. In parallel, the international community has witnessed notable development in issuing standards and regulatory frameworks, such as ISO/IEC 42001:2023, the NIST AI Risk Management Framework, and the European Union AI Act, reflecting a global trend toward subjecting artificial intelligence to specialized legal governance.

Proceeding from this, this section addresses the legal framework and governance of AI use in cybersecurity protection through three subsections: the first is devoted to examining the Saudi statutory regulation; the

second addresses the relevant international standards and agreements; while the third addresses the mechanisms for developing the legislative environment and strengthening the legal governance of artificial intelligence.

3.1 Saudi Statutory Regulation of AI Use in Cybersecurity Protection

Saudi regulation of cybersecurity is distinguished by the fact that it did not arise as a single piece of legislation gathering all the provisions relating to the protection of cyberspace, but developed gradually through an integrated system of laws, regulations, and regulatory controls that combine to achieve a single objective: protecting national interests, digital assets, and sensitive infrastructure, while enabling digital transformation and strengthening trust in the digital economy.

It is noted that the Saudi legislator has not yet issued an independent law comprehensively regulating artificial intelligence; this, however, does not mean the existence of a legislative vacuum, since the use of artificial intelligence is regulated through an interconnected set of general laws and controls that govern the environment in which artificial intelligence operates, particularly in the field of cybersecurity. Studying Saudi regulation accordingly requires analyzing this system as an integrated legislative framework, rather than a collection of scattered provisions.

At the forefront of this system stands the Royal Order establishing the National Cybersecurity Authority, which granted the Authority legal personality and financial and administrative independence, and made it the entity competent for cybersecurity in the Kingdom, responsible for setting policies, issuing controls, monitoring compliance with them, and raising the level of national readiness to confront cyber threats. This step is of particular importance, as it moved cybersecurity from a technical function within government entities into a framework of national governance linking cybersecurity to national security, economic development, and digital transformation (National Cybersecurity Authority, 2017).

This regulatory structure entails that any use of artificial intelligence in the field of cybersecurity must take place within the framework of the national policies and controls adopted by the Authority, ensuring that technical solutions do not conflict with national-security requirements or with the minimum mandatory controls prescribed for entities subject to the regulation.

The Essential Cybersecurity Controls (ECC) constitute the cornerstone of Saudi cybersecurity regulation, having been set by the National Cybersecurity Authority as the minimum required for the protection of informational and technical assets. The Controls are divided into five main domains, comprising governance, cybersecurity defense, cybersecurity resilience, third-party and cloud-computing security, and industrial control systems security, together with a large number of detailed controls relating to identity management, asset management, vulnerability management, incident management, and business continuity (National Cybersecurity Authority, 2018). The Authority updated the Controls in ECC-2:2024, which retains this structure while strengthening the requirements applicable to emerging technologies, including AI-based security tools, consistent with the analysis developed in this study (National Cybersecurity Authority, 2024a).

From a legal standpoint, these Controls are distinguished by the fact that they do not mandate the use of a particular technology, but adopt an outcome-based regulatory approach, requiring entities to achieve a specified level of protection and risk management while leaving them free to choose the appropriate technical means, including artificial intelligence. This approach is more flexible than legislation mandating specific technical means, as it allows regulation to keep pace with rapid technological development without the need for continual amendment of regulatory texts.

This approach confirms that artificial intelligence is not a regulatory end in itself, but a means that can be deployed to fulfil the statutory obligations relating to risk management, incident detection, asset protection, and business continuity. If an institution can demonstrate that its use of AI technologies has contributed to raising its level of compliance and meeting the requirements of the Essential Cybersecurity Controls, this constitutes a lawful application of technical means in fulfilling its statutory obligations.

Among the other important related laws is the Personal Data Protection Law, which constitutes the legal framework regulating data processing in the Kingdom. This Law takes on added importance in the context of artificial intelligence, since most AI applications rely on the processing of personal data or data that may be linked to individuals directly or indirectly. The use of artificial intelligence in cybersecurity must therefore comply with all the statutory principles relating to the lawfulness of processing, purpose specification, data minimization, ensuring accuracy, securing data, and respecting the rights of data subjects (Kingdom of Saudi Arabia, 2021, as amended 2023).

The close relationship between the two laws is evident in the fact that cybersecurity constitutes the principal means of protecting personal data, while data protection represents one of the objectives that cybersecurity seeks to achieve. Any breach of security controls may therefore simultaneously constitute a violation of the obligations established under the Personal Data Protection Law, reinforcing the notion of legislative integration between the two laws.

From another perspective, the Saudi Data and Artificial Intelligence Authority has issued the AI Ethics Principles, the first national framework establishing the governing principles for the development and use of AI technologies in the Kingdom. These principles include a set of core values, most notably fairness, transparency, privacy, security, accountability, and reliability — principles that intersect directly with cybersecurity requirements, particularly when intelligent systems are used to analyze data, take security decisions, or monitor digital activities (SDAIA, 2023).

The importance of these principles is not confined to their ethical dimension, but extends to their legal dimension, as they constitute a guiding reference that can be relied upon in interpreting statutory obligations or assessing entities' compliance with best practices in the use of artificial intelligence. They also pave the way for the development of more detailed legislation in the future, particularly given the expected expansion in the use of intelligent systems within government and private sectors.

Among the practical applications of this regulatory framework is what is observed in the financial sector in the Kingdom, where the use of artificial intelligence in fraud detection, risk analysis, and transaction monitoring is subject to dual regulatory oversight: on the one hand, compliance with cybersecurity controls, and on the other, compliance with the statutory requirements relating to data protection, confidentiality of information, governance, and risk management issued by the competent regulatory authorities — reflecting the integration between the various regulatory frameworks.

It is argued here that Saudi regulation is distinguished by several merits, the most important being that it has not treated artificial intelligence as a separate technical issue, but has incorporated it within a broader system founded on governance, risk management, compliance, and data protection. Despite its strength, however, this regulatory framework still requires a specialized statutory framework addressing the novel issues raised by AI applications, such as regulating liability for automated decisions, algorithmic-transparency requirements, procedures for approving high-risk systems, and mechanisms for the independent review of intelligent models, consistent with recent international legislative developments.

This confirms that Saudi regulation has established a strong institutional and legislative foundation for cybersecurity protection, and that this foundation allows for the accommodation of AI applications within

an integrated legal framework. The development of this technology, however, requires the continuous updating of the regulatory environment so as to achieve a balance between encouraging innovation and ensuring legal protection — which leads to an examination of the international standards and agreements regulating the use of artificial intelligence in cybersecurity in the following subsection.

3.2 International Standards Governing the Use of Artificial Intelligence in Cybersecurity

Regulating the use of artificial intelligence in the field of cybersecurity is no longer a purely national matter, but has become part of an international legal and technical system characterized by overlap and integration. This is due to the cross-border nature of digital spaces, the fact that cyber threats do not recognize political or geographic boundaries, and the fact that AI technologies themselves are generally developed in a global environment based on cooperation between technology companies, research centers, and international organizations. Accordingly, no national regulation can achieve full effectiveness unless it is consistent with international standards and global best practices in this field.

For this reason, international organizations have, in recent years, moved toward developing specialized normative frameworks that do not merely regulate cybersecurity or artificial intelligence separately, but address the relationship between them within a single framework founded on risk management, governance, transparency, compliance, and accountability. It is noted that not all of these frameworks take the form of binding legal rules; they are predominantly normative and advisory in character. Nevertheless, their practical importance often exceeds that of national legislation, as regulatory authorities, courts, and oversight bodies have come to rely on them in determining the standard of due diligence and best practices.

Table 1 summarises how the Saudi framework and the six international instruments reviewed in this study address the five governance functions set out in the Methodology above. The table is a synthesis produced for this study, not a reproduction of any single source, and is intended to make the comparative argument developed in the remainder of this subsection easier to follow.

Governance function	Saudi (ECC / PDPL / SDAIA)	ISO 27001 / 27005	ISO/IEC 42001	NIST CSF 2.0	NIST AI RMF	EU AI Act
Risk assessment	Mandatory under ECC; PDPL lawful-basis test	ISMS risk process (27001/27005)	AI-specific lifecycle risk assessment	"Identify" function	"Map" / "Measure"	Four-tier risk classification
Governance & accountability	ECC governance domain; NCA oversight; SDAIA mandate	Top-management accountability	Senior-management AI accountability	New "Govern" function	"Govern" function	Provider/deployer duties
Transparency & explainability	Non-binding (SDAIA principles only)	Documentation, not AI-specific	Explicit, documented requirement	Addressed indirectly	Core "trustworthy AI" trait	Binding, high-risk systems

Governance function	Saudi (ECC / PDPL / SDAIA)	ISO 27001 / 27005	ISO/IEC 42001	NIST CSF 2.0	NIST AI RMF	EU AI Act
Human oversight	Implicit in incident/access controls	Not AI-specific	Required oversight mechanisms	Not AI-specific	Explicit "Manage" principle	Mandatory, high-risk systems
Incident response & improvement	ECC resilience & continuity domain	PDCA continuous-improvement cycle	Continuous AI system monitoring	"Respond" / "Recover"	"Manage" function	Post-market monitoring duty

Table 1. Comparative mapping of Saudi and international frameworks against five common AI-cybersecurity governance functions (author's synthesis based on National Cybersecurity Authority, 2018, 2024a; Kingdom of Saudi Arabia, 2021, as amended 2023; SDAIA, 2023; ISO, 2018, 2022, 2023; NIST, 2023, 2024; European Union, 2024).

The pattern that emerges from Table 1 is consistent with the argument developed throughout this study: Saudi regulation is comparatively strong on risk identification, governance, and institutional accountability — areas where the ECC and PDPL already impose binding obligations — but comparatively underdeveloped on transparency, explainability, and human oversight specifically as they apply to AI-driven decisions, where the SDAIA AI Ethics Principles remain non-binding guidance rather than statutory requirements. This is precisely the gap that Section 3.3 and the Recommendations below address.

Foremost among these standards is ISO/IEC 27001:2022, the principal international reference for information security management systems. This standard's role is not confined to establishing technical controls for information protection, but provides an integrated management framework based on the continuous-improvement cycle (Plan–Do–Check–Act), aiming to integrate information security into the institution's management structure so that it becomes part of the governance and compliance system, rather than a mere standalone technical function (ISO, 2022).

The importance of this standard for artificial intelligence lies in the fact that it establishes the regulatory environment within which intelligent systems operate. However advanced artificial intelligence may become, it cannot perform its security function efficiently unless it forms part of a management system that assigns responsibilities, assesses risks, sets controls, monitors compliance, and is subject to continuous review. The relationship between ISO/IEC 27001 and artificial intelligence is therefore one of complementarity: the standard provides the management framework, while artificial intelligence provides the technical means that help implement that framework more efficiently.

Related to this standard is ISO/IEC 27005:2022, which concerns information security risk management. This standard is distinguished by providing a detailed methodology for identifying, analyzing, assessing, treating, and monitoring risks, while emphasizing that risk management is not merely a periodic activity but a continuous process that changes with the operational environment and evolving cyber threats. This reveals the close relationship between this standard and artificial intelligence, since the latter enables the automated implementation of continuous risk assessment, through real-time data analysis and the dynamic updating of risk levels, consistent with the standard's underlying philosophy (ISO, 2022).

Among the most important recent legislative and normative developments is the issuance of ISO/IEC 42001:2023, the first international standard specialized in Artificial Intelligence Management Systems. The importance of this standard lies in the fact that it has moved the regulation of artificial intelligence from the stage of general principles to the stage of building an integrated management system that determines the responsibilities of senior management, governance requirements, risk management, control of the intelligent system's lifecycle, documentation, review, and continuous improvement (ISO, 2023).

Analysis of this standard reveals that it does not regard artificial intelligence as a technical product, but as an institutional system that must be subject to the same rules governing quality, risk-management, and compliance systems. The standard accordingly imposes on the institution an obligation to conduct a prior assessment of the risks associated with the use of artificial intelligence, establish clear policies for its use, determine responsibilities, ensure the existence of human oversight, and document the processes of development, training, testing, and updating.

This approach represents an important shift in legal thinking, as it links the lawfulness of using artificial intelligence to the extent of the institution's commitment to managing its risks, rather than to the mere possession of the technology. Legal liability is therefore no longer linked solely to the result of an error, but may also arise where the institution has neglected to establish an appropriate AI-governance system, even before harm occurs.

In the United States, the National Institute of Standards and Technology (NIST) has issued two frameworks that are among the most important reference documents in this field. The first is the Cybersecurity Framework 2.0, issued in 2024, and the second is the AI Risk Management Framework 1.0, issued in 2023.

The Cybersecurity Framework (CSF 2.0) is distinguished by the fact that it is no longer confined to the technical aspects of protection, but has added a new function — Govern — bringing the core functions to six: Govern, Identify, Protect, Detect, Respond, and Recover (NIST, 2024). This addition is not a merely formal amendment, but reflects a fundamental shift in the philosophy of cybersecurity, as the board of directors and executive management have become directly responsible for managing cyber risk, allocating resources, adopting policies, and monitoring compliance; cybersecurity is no longer the responsibility of the information-technology department alone.

The AI Risk Management Framework (AI RMF), for its part, is founded on four core functions: Govern, Map, Measure, and Manage (NIST, 2023). This framework is among the most influential documents in contemporary legal thought, as it introduced the concept of Trustworthy AI, linking it to a set of principles including safety, security, transparency, explainability, privacy, fairness, accountability, and the capacity to manage risk.

From a legal perspective, this framework represents a shift from regulating the outcomes of AI use to regulating the process of its development and management. It asks not only: is the system safe? but also: how was it trained? who reviewed it? what are its data sources? does human oversight exist? can its decisions be explained? have tests been conducted to verify its resistance to adversarial attacks? These questions reflect a significant development in the philosophy of the legal governance of artificial intelligence.

Among the most important international legislative developments is also the European Union Artificial Intelligence Act (AI Act), the first comprehensive legislation regulating artificial intelligence on the basis of a risk-based approach. This legislation divides AI applications into different tiers according to their degree of risk, beginning with applications of unacceptable risk, whose use is prohibited, then high-risk

applications subject to strict obligations, then limited-risk applications, and finally low-risk applications subject to lighter regulatory requirements.

This legislation is of particular importance to the subject of this research, as many AI applications used in cybersecurity may fall within the category of high-risk systems, entailing an obligation on the institution to conduct risk assessments, ensure data quality, document the system, provide human oversight, log activities, and demonstrate compliance. This reveals a clear trend toward preventive regulation, focusing on preventing risks before they materialize rather than holding those responsible accountable after harm occurs.

As regards combating cybercrime, the Budapest Convention on Cybercrime (Council of Europe, 2001) represents one of the most important international agreements in this field, having established rules for international cooperation in confronting cybercrime, exchanging information, collecting digital evidence, and mutual legal assistance. Although the Convention does not address artificial intelligence directly, it provides a legal framework for dealing with crimes in which AI technologies may be used, whether in carrying out attacks or in collecting related evidence.

A comparison between these international standards and Saudi regulation reveals substantial convergence in fundamental principles, even though the legal tools employed differ. The Saudi Controls are founded on risk management, governance, compliance, and asset protection, consistent with the philosophy of ISO and NIST. Likewise, the principles of responsible artificial intelligence issued by the Saudi Data and Artificial Intelligence Authority intersect with the principles of trustworthy AI adopted by the NIST AI RMF, as well as with the transparency and human-oversight requirements contained in the European AI Act.

Saudi regulation, however, is distinguished by its focus on the institutional and sovereign character of cybersecurity, being directly linked to the protection of national security and critical infrastructure, whereas international standards focus to a greater extent on building general governance models applicable across different legal environments. The relationship between them is therefore not one of conflict but of complementarity: Saudi entities can apply the national controls as the mandatory minimum, while benefiting from international standards to develop internal practices and raise their level of institutional maturity.

This study maintains that the most significant feature of the contemporary international trend is its shift from regulating the technology to regulating the management of the technology. Modern standards no longer ask only about the efficiency of the algorithm, but focus on the entire lifecycle of the intelligent system, from its design, through its training, operation, and monitoring, to its decommissioning or updating. This trend is consistent with the philosophy adopted by the Saudi legislator in the field of cybersecurity, and constitutes a suitable foundation for developing a more specialized Saudi legislative framework for regulating artificial intelligence — not by replicating foreign models, but by drawing on them in a manner consistent with the particularities of the Kingdom's regulatory environment.

It thus becomes clear that international standards are no longer mere technical references, but have come to represent a legal and normative reference guiding national legislation and determining the standard of professional conduct required of institutions — which paves the way for moving to the third subsection, addressing the means of developing the legislative framework and governance of AI use in cybersecurity protection, presenting a forward-looking analytical vision based on the findings of the comparative study.

3.3 Means of Developing the Legislative Framework and Governance of AI Use in Cybersecurity Protection

The preceding study has revealed that the rapid development of AI technologies has, in many cases, outpaced the speed at which regulating legislation develops — a phenomenon explained by the nature of technical innovation, which changes in rapid succession, whereas the legislative process requires a longer time to prepare, discuss, adopt, and then implement statutory texts. This gap between technical development and legal regulation, however, should not lead to mere reliance on existing general frameworks, but calls for developing a flexible legislative environment capable of continuous updating, one that accommodates technical developments without impeding innovation or limiting institutions' ability to benefit from artificial intelligence in protecting cybersecurity.

From this standpoint, developing the legislative framework does not necessarily mean issuing an independent law for every new technology, but means building a legislative governance system capable of regulating the lifecycle of AI systems, linking them to legal liability, risk management, compliance, data protection, and institutional oversight, so that artificial intelligence becomes part of the institution's legal environment, rather than a mere technical tool operating in isolation from it.

This development begins with establishing a specialized national legislative framework for AI governance. Although the Kingdom of Saudi Arabia possesses an advanced cybersecurity system, a developed personal-data-protection law, and national principles for responsible artificial intelligence, these texts remain distributed among several regulatory instruments. Consolidating them within a statutory framework or an implementing regulation specialized in AI governance would achieve greater clarity and consistency, determine the legal obligations of developers, operators, and users, and enhance legal certainty for investors and governmental and private entities.

This does not mean replicating comparative experiences such as the European Union's Artificial Intelligence Act (AI Act), but rather benefiting from its methodology of classifying applications according to their level of risk, while taking into account the particularities of the Saudi regulatory environment. It is not appropriate for all AI applications to be subject to the same regulatory restrictions, since the nature of risk varies according to the field of use. Systems used in entertainment or marketing differ in their legal implications from systems used in protecting critical infrastructure, managing government services, or in the financial sector. Regulation based on risk assessment is therefore more flexible and efficient than regulation based on uniform rules applied to all applications.

Related to this is the necessity of establishing clear legal rules for liability regarding decisions issued by AI systems. The study has shown that the predominant trend does not recognize an independent legal personality for artificial intelligence, meaning that liability rests with the natural or legal persons connected with developing, operating, or using the system. This trend, however, requires further legislative elaboration, so as to determine the standards for apportioning liability among the developer, the supplier, the operator, and the entity owning the system, according to the degree of control over the system, the foreseeability of harm, and the extent of compliance with technical and professional controls. In this context, benefit may be drawn from traditional principles of civil liability, developed to suit the particular characteristics of artificial intelligence. The standard should not be confined to proving fault after harm occurs, but should extend to assessing the extent of the institution's compliance with governance and risk-management procedures before it occurs. If it is established that the institution failed to conduct appropriate security tests, failed to review training data, or failed to provide human oversight over high-risk decisions, its liability may arise even if the immediate error originated from the intelligent system.

Among the other important aspects requiring development is strengthening the institutional governance of artificial intelligence. It is no longer acceptable for decisions on the use of artificial intelligence to remain the exclusive preserve of technical departments, since the effects of these decisions extend to the legal, financial, regulatory, and strategic aspects of the institution. The board of directors and the audit and risk committees must therefore play a direct role in approving AI-use policies, reviewing the associated risk reports, and ensuring that its use is consistent with statutory requirements and institutional objectives.

This trend is consistent with what is established by the NIST Cybersecurity Framework 2.0, which makes governance the starting point in cybersecurity management, and with ISO/IEC 42001:2023, which requires senior management to bear responsibility for the AI management system rather than confining it to technical departments (NIST, 2024; ISO, 2023). This confirms that governance is no longer a supplementary element, but has become a fundamental prerequisite for achieving the responsible and secure use of artificial intelligence.

Among the aspects deserving particular attention is strengthening the requirements of transparency and explainability. The greater the capacity of intelligent systems to make independent decisions, the greater the need to enable regulatory and judicial authorities, and indeed the individuals affected by a decision, to understand the basis on which it was made. This does not mean disclosing trade secrets or algorithmic details, but providing a reasonable degree of explanation allowing the lawfulness of the decision to be verified, reviewed, and challenged where appropriate. These requirements represent one of the most important legal safeguards for protecting digital rights against automated decisions.

Developing the legislative environment also requires integrating cybersecurity requirements into the lifecycle of AI systems from the design stage onward, an approach expressed through the principle of Security by Design, whereby security is not left to be addressed after the system becomes operational, but becomes part of the development process itself. Linked to this principle is the principle of Privacy by Design, which requires that the protection of personal data be taken into account at every stage of the system's development, from data collection, through model training, to its operation and maintenance.

Among the aspects of importance is also the establishment of independent mechanisms for the assessment and review of AI systems, particularly those used in high-risk sectors. Just as financial statements are subject to external audit, and quality systems to periodic auditing, AI systems should be subject to independent reviews to verify the integrity of training data, the accuracy of models, their resistance to adversarial attacks, and their compliance with statutory controls. Such reviews would enhance trust in intelligent systems and reduce the risks of bias and algorithmic error.

Within the same framework, the researcher considers it important to develop national standards for accreditation and professional certification in the field of artificial intelligence and cybersecurity, whereby entities that comply with governance, risk-management, and compliance requirements are granted national accreditation certificates reflecting their level of institutional maturity. The effect of this measure is not confined to enhancing trust, but also contributes to raising the level of competition and encouraging institutions to adopt international best practices.

From a legislative standpoint, future development should move toward adopting adaptive regulation, balancing the protection of society with the encouragement of innovation. Legislative rigidity may hinder the use of modern technologies, while overly lenient regulation may endanger rights and interests. The best approach is therefore legislation based on general principles, supported by implementing regulations and technical standards capable of periodic updating whenever technologies evolve or new risks emerge.

Institutional cooperation is one of the most important elements for the success of this model. Regulating artificial intelligence in the field of cybersecurity cannot be undertaken by a single body, but requires integration between the National Cybersecurity Authority, the Saudi Data and Artificial Intelligence Authority, sectoral regulatory bodies, the judiciary, the private sector, and universities and research centers. This cooperation leads to building an integrated system combining legislation, implementation, oversight, scientific research, and technical development.

A comparison between the Saudi experience and international trends reveals that the Kingdom possesses a strong regulatory foundation qualifying it to move to a more advanced stage in AI governance. The Essential Cybersecurity Controls, the Personal Data Protection Law, and the national principles of responsible artificial intelligence provide a solid foundation on which to build. The next stage, however, requires moving from regulating the environment in which artificial intelligence operates to regulating artificial intelligence itself, while preserving the flexibility that allows it to keep pace with technical development without impeding innovation.

The present analysis holds that the future of cybersecurity will not be determined by the development of AI technologies alone, but by the capacity of legislation to accommodate this development. The greater institutions' reliance on intelligent systems, the more important legal governance becomes relative to technical development itself, since technology can be purchased or developed, whereas legal trust can only be achieved through an integrated system founded on lawfulness, transparency, accountability, risk management, and the protection of rights. Building an effective cybersecurity protection system in the future therefore requires moving from the stage of regulating the use of artificial intelligence to the stage of governing artificial intelligence itself, as the framework that brings together security, innovation, and the rule of law.

This completes Section Three, and with it the substantive structure of the research, following the literature review and methodology, Section One's treatment of the conceptual foundation, Section Two's treatment of the practical applications of artificial intelligence in cybersecurity protection, and Section Three's analysis of the legal framework and governance and its proposed means of development. Before turning to the general conclusion, the following discussion synthesises the study's findings against the literature reviewed above, considers counterarguments to its central claims, and states the study's limitations.

Discussion

The analysis developed across Sections One to Three converges on a single proposition: Saudi regulation is comparatively mature on the governance and risk-identification functions that predate the current wave of AI adoption, but comparatively underdeveloped on the transparency, explainability, and human-oversight functions that are specific to AI-driven decision-making, as Table 1 illustrates directly. This finding both confirms and extends Alzahrani's (2024) conclusion that Saudi Arabia possesses an advanced institutional foundation for AI governance: the present study agrees with that assessment at the level of institutional architecture, while showing, through the function-by-function comparison in Section 3.2, precisely where that architecture stops short of the binding transparency and oversight obligations found in ISO/IEC 42001:2023, the NIST AI RMF, and the EU AI Act. In this respect the finding also qualifies Kshetri's (2023) broader observation that AI-driven detection gains are generally offset by new attack-surface risk: the Saudi case suggests that the offsetting risk is not purely technical (adversarial manipulation, data poisoning) but also regulatory, arising from the absence of binding explainability requirements for the automated security decisions discussed in Sections 2.1 and 2.2.

A reasonable counterargument to this study's central recommendation — that Saudi Arabia should move toward a more specialised AI-governance instrument — is that the outcome-based, technology-neutral structure of the Essential Cybersecurity Controls is itself a deliberate and defensible regulatory choice, valued precisely because it avoids locking the law to a fast-moving technology (a concern also raised, in the EU context, by critics of the AI Act's prescriptive risk tiers). On this view, adding AI-specific statutory requirements risks the same rigidity the outcome-based approach was designed to avoid. This study's response, developed in Section 3.3, is that the two are not mutually exclusive: an outcome-based framework can be preserved at the level of substantive cybersecurity obligations while a narrower set of governance-process requirements — explainability, human oversight for high-risk decisions, and independent model review — is added specifically where automated decisions affect individual rights. This mirrors the approach ISO/IEC 42001:2023 itself takes in relation to ISO/IEC 27001, layering AI-specific management requirements onto an existing outcome-based security management system rather than replacing it.

A second, related counterargument concerns liability: since this study finds that comparative legislation does not currently grant AI systems independent legal personality, one might ask whether new liability rules are necessary at all, given that ordinary rules of civil and administrative liability already attach to the operating institution by default. The study's position, set out in Section 3.3, is that default liability rules are necessary but not sufficient, because they do not, on their own, specify the standard of care expected of an institution deploying AI in cybersecurity (for example, whether adversarial-robustness testing or training-data review is a mandatory component of due diligence). Without that specification, liability outcomes will vary unpredictably from one dispute to the next, which is itself a cost to legal certainty that a more specific framework would reduce.

Limitations

This study is subject to several limitations that should inform how its conclusions are read. First, and most significantly, the method is doctrinal and comparative rather than empirical: the study analyses the text and stated scope of legal instruments and standards, but does not draw on interviews with regulators, security practitioners, or compliance officers, nor on a systematic review of enforcement decisions, administrative sanctions, or case law, because no sufficient body of published Saudi decisions specifically addressing AI-related cybersecurity liability could be identified at the time of writing. The study's conclusions about the practical operation of the Saudi framework should accordingly be read as a doctrinal reading of the law as written, not as a finding about how the law is applied in practice.

Second, the study is bounded to instruments in force, or formally adopted, as of January 2026; both AI regulation and Saudi cybersecurity regulation are evolving rapidly, and specific comparative conclusions — particularly those concerning the EU AI Act's implementing measures and any further revision of the ECC — may require updating as those instruments mature.

Third, the six-framework comparison in Section 3.2, while selected on stated and defensible criteria, is necessarily a selection rather than an exhaustive survey of the international regulatory landscape; sectoral frameworks (for example, financial-sector-specific AI guidance) and non-binding instruments such as the OECD AI Principles were treated as secondary and may repay closer comparative attention in future work.

Fourth, this is a single-jurisdiction study organised around Saudi law as the primary reference point; while the comparative material provides external benchmarks, the recommendations in Section 3.3 are calibrated

to the Saudi regulatory environment and are not presented as a general-purpose model for AI-cybersecurity governance in other jurisdictions.

These limitations point directly to the priorities for future research identified in the Recommendations below, particularly the call for empirical and case-based studies once a sufficient body of Saudi enforcement practice becomes available.

General Conclusion

Over the past two decades, the world has witnessed a radical transformation in the nature of digital threats, as cyber risks have shifted from being technical incidents of limited scope into strategic challenges affecting national security, economic stability, business continuity, and the protection of digital rights and freedoms. This transformation has been accompanied by the rapid development of AI technologies, which are no longer merely tools for improving technical performance, but have become a core element in building modern security systems, given their exceptional capabilities in data analysis, continuous learning, risk prediction, early threat detection, and support for response operations and cyber-crisis management.

It is based on the above premise that this research has aimed at examining the role of artificial intelligence in developing an effective cybersecurity protection framework from a legal and analytical perspective through the lens of the Saudi legislations, international standards, and risk-governance frameworks, knowing that the success of artificial intelligence (AI) in strengthening cybersecurity does not rely solely on its technical efficiency but is closely linked to the existence of a legal, regulatory, and institutional framework that ensures the responsible use of this technology, defines the responsibilities of the parties involved in it, and achieves a balance between protecting digital security and safeguarding fundamental rights and freedoms.

In Section One of the study, the researchers showed that artificial intelligence and cybersecurity are not two separate systems, but rather two joined systems that cannot be separated from one another. Artificial intelligence is not just another additional tool that can be attached to the traditional cybersecurity suite; it's a change in the mindset of how cybersecurity is conducted, moving from reactive to proactive, from responding to threats to anticipating them, and from reacting to the data to learning from it. The study also found that this transformation has necessitated the rethinking of many concepts that have been embedded in the traditional law, especially the concepts of liability, oversight, due care, governance and compliance, which has demanded the development of legal thought in parallel with technical.

In Section Two of the study, researchers found that AI has revolutionized the entire cybersecurity lifecycle, from securing against attacks and detecting them early, to responding to them and managing risks, all the way to ensuring business continuity and disaster recovery. The analysis can be used to prove that AI has an unprecedented ability to process the massive amounts of security data, analyze it in real time, identify abnormal patterns, prioritize responses and shorten the time it takes to contain a cyber incident. Concurrently, the study identified several challenges that underscore the need for an integrated governance framework to ensure the efficient functioning of AI, including adversarial attacks on intelligent models, transparency and explainability, algorithmic bias, personal data protection and legal liability for the consequences of automated decisions.

In Section Three, the team examined the legal landscape for the use of artificial intelligence in cybersecurity protection, finding that the Kingdom of Saudi Arabia had a sophisticated legal framework that includes the National Cybersecurity Authority, the Essential Cybersecurity Controls, the Personal

Data Protection Law, and the national principles of responsible artificial intelligence, which create a firm legal base that can set the stage for AI applications within a governance-based system of risk management, compliance and data protection. Despite the different approaches in terms of legal tools used by each system, the study found significant alignment between the regulation and recent international frameworks, such as ISO/IEC 27001, ISO/IEC 27005, ISO/IEC 27000, ISO/IEC 42001, the NIST Cybersecurity Framework 2.0, the NIST AI Risk Management Framework, and the European Union Artificial Intelligence Act.

Comparative study showed that the international trend is no longer about the regulation of the technology, but the governance of the technology, i.e., technology lifecycle, technology risks, technology responsibilities, transparency, human oversight, explainability and continuous review. This study revealed that this is aligning with the Saudi regulation philosophy which is based on risk management and institutional governance that pave the way for the Kingdom's regulatory environment to adapt to the future legislation in this field without a significant change to the current law.

One of the most significant conclusions that the study drew is that AI is not a replacement for humans, but rather a valuable tool for enhancing security and legal decisions. Overreliance on intelligent systems can create new risks of human oversight being diminished, decisions being hard to explain, and the responsibility for decisions slipping away from human to algorithmic agents. An efficient cybersecurity protection system is thus one in which the tasks of analysis, prediction and technical support are performed by the artificial intelligence and in which the human element is reserved for ensuring the oversight and legal judgment.

The study also pointed out that the future of cybersecurity does not simply rely on the latest technologies, but on the ability of the states and institutions to develop and maintain a flexible legislative toolbox that can keep pace with the ongoing advances in artificial intelligence, while simultaneously balancing the needs of national security, innovation, privacy, transparency, and digital rights. The technical challenge in the next period will therefore be much less than the legislative and regulatory challenge – namely how to design the legal framework to ensure that it will not hinder the development of innovative solutions, and how to balance the protection of society and the limitations for digital development.

The researcher believes that the Kingdom of Saudi Arabia has the opportunity and the ability to be a leading example in the region and internationally in AI and cybersecurity governance, especially given the policy goals of Saudi Vision 2030, the development of digital infrastructure, the reinforcement of governance and the creation of national institutions for cybersecurity, data and AI. This position must be supported continuously by updating the legislative framework, establishing oversight systems, reinforcing the cooperation among various regulatory authorities and implementation of the most modern international standards, considering the specific characteristics of the regulatory system of the Kingdom.

Based on the above, the study supports the integration of three dimensions: technology, legislation and institutional government, which is the only way to build a complete cybersecurity protection framework in the era of artificial intelligence. This integration alone is capable of achieving sustainable digital security, strengthening trust in the digital environment, and ensuring the responsible use of artificial intelligence in a manner that serves economic development, protects national security, and safeguards rights and freedoms at once.

I. Findings

In light of the study's analysis of Saudi legislation, international standards, and risk-governance framewor-

ks, the following findings may be drawn:

1. The study has demonstrated that artificial intelligence is no longer merely a technology supporting cybersecurity, but has become one of the structural components of the modern cybersecurity system, having contributed to shifting it from a reactive protection model to a proactive protection model founded on risk prediction and the continuous analysis of data.
2. The study has revealed that the effectiveness of artificial intelligence in protecting cybersecurity is not measured by its abstract technical capabilities, but by the extent of its integration within an institutional framework founded on governance, risk management, compliance, and continuous oversight — making governance a complement to the technology rather than something separate from it.
3. The study has found that the relationship between artificial intelligence and cybersecurity is one of strategic integration rather than substitution: artificial intelligence enhances the efficiency of the cybersecurity system, while cybersecurity provides the secure environment that ensures the sound operation and continuity of intelligent systems.
4. The study has shown that artificial intelligence has brought about a fundamental transformation across all stages of the cybersecurity cycle, from the prevention and early detection of threats, through incident response and risk management, to supporting business continuity and disaster recovery.
5. The study has demonstrated that the use of behavioral-analysis, machine-learning, and deep-learning techniques, together with SIEM, SOAR, and XDR platforms, has contributed to raising the efficiency of cyber-threat detection and reducing response time, although these applications remain in need of continuous human oversight to ensure the soundness of security decisions.
6. The study has shown that excessive reliance on artificial intelligence may give rise to new legal and technical risks, most notably adversarial attacks on intelligent models, training-data poisoning, algorithmic bias, and weak explainability of automated decisions — requiring these systems to be subject to more rigorous governance controls.
7. The study has found that the protection of personal data represents one of the most important legal pillars for the use of artificial intelligence in cybersecurity, and that the lawfulness of data processing has become a fundamental condition for the lawfulness of using intelligent systems in the security field.
8. The study has revealed that Saudi regulation is distinguished by its institutional integration, linking the National Cybersecurity Authority, the Saudi Data and Artificial Intelligence Authority, the Personal Data Protection Law, and the Essential Cybersecurity Controls, thereby providing an advanced regulatory environment for accommodating AI applications.
9. The study has found that the Essential Cybersecurity Controls in the Kingdom adopt outcome-based, risk-management-oriented regulation, an approach that grants institutions flexibility in choosing the appropriate technical means, including artificial intelligence, to fulfil their statutory obligations.
10. The study has shown substantial convergence between Saudi regulation and recent international standards, particularly ISO/IEC 27001, ISO/IEC 27005, and ISO/IEC 42001, the NIST Cybersecurity Framework 2.0, and the NIST AI Risk Management Framework, in terms of their focus on governance, risk management, transparency, and compliance.
11. The study has revealed that ISO/IEC 42001:2023 represents a qualitative leap in the regulation of artificial intelligence, having moved from regulating technical aspects to building an integrated AI management system linked to institutional governance and legal liability.

12. The study has found that the contemporary international trend adopts a risk-based regulatory approach rather than one focused on regulating the technology itself — a trend consistent with the philosophy of Saudi legislation, from which benefit may be drawn in developing the national regulatory environment.
13. The study has shown that legal liability for decisions issued by AI systems continues to rest on the liability of the natural or legal persons responsible for developing, operating, or using the system, and that artificial intelligence does not enjoy an independent legal personality under comparative legislation.
14. The study has confirmed that the successful use of artificial intelligence in cybersecurity is linked to the existence of effective human oversight (Human-in-the-Loop), particularly in decisions carrying significant legal, financial, or security consequences, so as to achieve a balance between automation and legal liability.
15. The study has concluded that the future of cybersecurity depends on the integration of three core pillars — intelligent technologies, the legal framework, and risk governance — and that any deficiency in one of these pillars weakens the effectiveness of the security system, however technically advanced it may be.

II. Recommendations

Based on the foregoing findings, the researcher recommends the following:

1. Expediting the preparation of a specialized Saudi statutory framework for AI governance, clearly defining the requirements for designing, operating, and reviewing intelligent systems, managing their risks, and determining the liability of the parties connected to them.
2. Developing the Essential Cybersecurity Controls by adding specific requirements for the governance of AI systems, including model security, training-data quality, adversarial-attack resistance testing, and requirements of transparency and explainability.
3. Aligning the Saudi regulatory frameworks with ISO/IEC 42001:2023, so as to ensure the adoption of global best practices in managing AI systems without compromising the particularities of the Kingdom's regulatory environment.
4. Requiring entities using artificial intelligence in cybersecurity to conduct periodic risk assessments of intelligent models, review and document their results, and link them to enterprise risk-management plans.
5. Adopting the principle of mandatory human oversight for all high-risk security decisions relying on artificial intelligence, and not relying solely on automated decisions in matters affecting individuals' rights or legal standing.
6. Establishing statutory rules regulating legal liability for damage arising from the use of artificial intelligence, setting out the limits of liability of the developer, the supplier, the operator, and the entity owning the system.
7. Establishing an independent national mechanism for assessing and accrediting AI systems used in vital sectors, ensuring their compliance with cybersecurity and data-protection requirements.
8. Strengthening institutional integration between the National Cybersecurity Authority, the Saudi Data and Artificial Intelligence Authority, and sectoral regulatory bodies, so as to unify policies relating to artificial intelligence and cybersecurity.

9. Incorporating cybersecurity and privacy requirements from the design stage of AI systems (Security by Design and Privacy by Design) within the statutory requirements for digital projects.
10. Supporting specialized scientific research into the legal aspects of artificial intelligence and cybersecurity, and encouraging universities and research centers to prepare applied studies addressing emerging statutory issues.
11. Developing specialized training programs for judges, public prosecutors, lawyers, and legal advisors, to enhance their competence in handling disputes relating to artificial intelligence and cybersecurity.
12. Encouraging government and private institutions to adopt risk-management-based governance models, and linking the use of artificial intelligence to performance indicators measuring the level of compliance, efficiency, and reliability.
13. Strengthening international cooperation in exchanging expertise and information relating to AI-powered cyber threats, and benefiting from international experience in developing the national legislative environment.
14. Periodically reviewing relevant laws to keep pace with rapid technical developments, so as to ensure the continued effectiveness of the legislative system and prevent it from lagging behind technical innovation.
15. Working to build an integrated Saudi model for the governance of artificial intelligence in cybersecurity, combining the requirements of national security, data protection, compliance, and innovation, and drawing on international best standards while taking into account the particularities of the Kingdom's regulatory system.

References

I. Saudi Laws and Regulations

1. National Cybersecurity Authority. (2017a). Royal Order No. 6801 establishing the National Cybersecurity Authority. Riyadh: Royal Diwan.
2. National Cybersecurity Authority. (2018). Essential Cybersecurity Controls (ECC-1:2018). Riyadh: National Cybersecurity Authority.
3. National Cybersecurity Authority. (2024a). Essential Cybersecurity Controls (ECC-2:2024). Riyadh: National Cybersecurity Authority.
4. National Cybersecurity Authority. (2024b). About the Authority. Retrieved August 4, 2026, from https://nca.gov.sa
5. Kingdom of Saudi Arabia. (2021, as amended 2023). Personal Data Protection Law, issued by Royal Decree No. M/19 (9/2/1443 AH), as amended by Royal Decree No. M/148 (5/9/1444 AH).
6. Saudi Data and Artificial Intelligence Authority (SDAIA). (2023). AI Ethics Principles.

II. International Standards

1. International Organization for Standardization. (2022). ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements. Geneva: ISO.
2. International Organization for Standardization. (2022). ISO/IEC 27002:2022 Information security controls. Geneva: ISO.
3. International Organization for Standardization. (2022). ISO/IEC 27005:2022 Information security risk management. Geneva: ISO.

4. International Organization for Standardization. (2023). ISO/IEC 42001:2023 Artificial intelligence — Management system. Geneva: ISO.
5. International Organization for Standardization. (2018). ISO 31000:2018 Risk Management — Guidelines. Geneva: ISO.

III. International Reports and Frameworks

1. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg, MD: NIST.
2. National Institute of Standards and Technology. (2024). Cybersecurity Framework (CSF 2.0). Gaithersburg, MD: NIST.
3. European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act).
4. European Union Agency for Cybersecurity (ENISA). (2023). ENISA Threat Landscape 2023. Athens: ENISA.
5. European Union Agency for Cybersecurity (ENISA). (2024). Cybersecurity and Artificial Intelligence. Athens: ENISA.
6. Council of Europe. (2001). Convention on Cybercrime (Budapest Convention).

IV. Foreign Books

1. Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
3. Stallings, W. (2024). Effective Cybersecurity: A Guide to Using Best Practices and Standards. Pearson.
4. Whitman, M. E., & Mattord, H. J. (2022). Principles of Information Security (7th ed.). Cengage.
5. Bishop, M. (2019). Computer Security: Art and Science (2nd ed.). Addison-Wesley.

V. Peer-Reviewed Academic Studies

1. Alzahrani, A. (2024). Artificial Intelligence Governance and Cybersecurity Regulation in Saudi Arabia. Journal of Cyber Policy, 9(2), 175–198.
2. Kshetri, N. (2023). Artificial Intelligence in Cybersecurity: Opportunities and Challenges. IT Professional, 25(4), 45–53.
3. Brundage, M., et al. (2020). Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims. AI Magazine, 41(4), 76–92.
4. Cath, C. (2018). Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges. Philosophical Transactions of the Royal Society A, 376(2133).
5. Schmitt, M. N. (Ed.). (2017). Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations. Cambridge University Press.

VI. Arabic-Language References

1. Al-Sayed, A. M. (2023). Legal Protection of Cybersecurity in Light of Modern Legislation [in Arabic]. Cairo: Dar Al-Nahda Al-Arabiya.
2. Al-Awadi, K. (2024). Artificial Intelligence and Cybersecurity: A Comparative Legal Study [in Arabic]. Journal of Law, Kuwait University, 48(2), 115-170.
3. Al-Shammari, A. (2023). Cyber Risk Management in Light of the National Cybersecurity Controls [in Arabic]. Journal of King Saud University for Legal Sciences.
4. Al-Otaibi, M. (2024). Governance of Artificial Intelligence in the Kingdom of Saudi Arabia [in Arabic]. Journal of Legal and Economic Studies.