An analytical framework for an ethics of AI in Saudi Arabia beyond individual rights

Albara Abulaban[1] and Abderrazak Belabes[2]

[1] Assistant Professor of International Business Law, Islamic Economics Institute, King Abdulaziz University,
aabulaban@kau.edu.sa
[2] Professor in Islamic Economics, Islamic Economics Institute, King Abdulaziz University, abelabes@kau.edu.sa

Abstract. The aim of the research is to propose an analytical framework for the ethics of AI in Saudi Arabia, called the Pentagon SMLET (*Sharīʿah*, Morality, Law, Ethics, Technical aspect), to guide the development and deployment of AI, maximizing its benefits and minimizing its harms. Based on the distinction between unforeseen events, hazards, and risks, it is important to distinguish collective disaster from individual risk in the context of AI systems and advocates a holistic governance approach that prioritizes multi-stakeholder collaboration and proactive policies. In light of this framework, the research analyzes Saudi Arabia's progress and regulatory framework for AI systems to promote trustworthy AI. Furthermore, it compares Saudi Arabia's ethical principles for AI with UNESCO's recommendations, illustrating how the kingdom not only adopts these guidelines but also develops them by integrating its own cultural and strategic considerations, with the aim of becoming a global leader in responsible AI development and policies. In terms of recommendation, the analytical framework based on the Pentagon SMLET could be useful in the enrichment of analytical foundations of the Riyadh Charter on AI ethics launched by ICESCO and SDAIA, just as much as for the Charter of Ethics of Artificial Intelligence developed by the ALECSO, for developing AI systems in line with Islamic values and principles.

Keywords: AI, Ethics, Law, Regulation, Governance

1. Introduction

Our personal data are collected by AI without national and international legislation being able to claim to protect privacy in such a situation. AI has a serious impact on the right to privacy and the protection of personal data. Protecting personal information is crucial for maintaining privacy and security (King & Meinhardt, 2024). How many times have we heard or read such statements that seem to be common sense? Yet, as important as they are, they are not based on a analytical framework to describe the impact of AI on the social structures of society beyond the rights of the individual (Brandtzaeg et al., 2025) and, consequently, to define clear priorities regarding the values that should govern the use of AI technologies (Stahl et al., 2023). Hence the importance of the legal aspects of AI beyond ethics strictly speaking (Carrillo, 2020), and the importance of the legal perspective to the development of an AI governance regime (Walz & Firth-Butterfield, 2019).

In view of this epistemological position underlying the mainstream literature on ethics of AI which favors the protection of individual rights and privacy (Burgess et al., 2018), it is necessary to work towards establishing a analytical framework allowing us to describe the impact of AI on the social structures of society beyond principles underpinning privacy and the protection of personal data (UN, 2022). This is all the more important given that there are several forms of AI ethics, which differ according to the operators, the type of tasks carried out, the main ethical effects

targeted, and the first cardinal virtues (Ménissier, 2023). This research aims to contribute to the development of such a analytical framework in light of Saudi Vision 2030 where AI ethics must be conceptualized within the framework of societal values shaped by family identity, which guide actions and relationships, and contribute to the well-being and development of each member, beyond purely economic considerations (Belabes, 2025a).

After defining the main technical terms with regard to the Saudi authority in charge of data and AI, we explore the relationship between ethics, individualism and risk, present the analytical model on which the study is based and which distinguishes between Sharia, morality, law, ethics and technical aspects of AI, address the Saudi AI ethics, institutional, legal, regulatory, and governance, analyze the UNESCO Recommendation on the Ethics of AI in light of the proposed analytical framework, and explore the challenges related to the implementation of the UNESCO Recommendation on the Ethics of AI in Saudi Arabia. The restate our main points and provide closure.

2. Definition of the main technical words based on the reports of the Saudi Data and AI Authority

Before getting to the heart of the matter, it is necessary to define the main technical terms relating to the world of data and AI and on which this research is based before addressing the key notions on which the analytical framework proposed in the research is based:

2.1. Law

"*The Personal Data Protection Law*" (SDAIA, 2023b: 2).

2.2. Regulations

"*The Implementing Regulations of the Law*" (SDAIA, 2023b: 2). This regulation primarily concerns the supervision of transfers of personal data outside Saudi Arabia. These transfers are subject to strict conditions to guarantee an adequate level of data protection, similar to that offered within the Kingdom. (SDAIA, 2024f: 3).

2.3. Data

"*A collection of facts in a raw or unorganized form such as numbers, characters, images, video, voice recordings, or symbols*" (SDAIA, 2023a: 7). In this regard, the assessment of data availability and readiness must be consistent with commitment to relevant regulations and policies (SDAIA, 2024a: 7).

2.4. Artificial Intelligence (AI)

"*Artificial intelligence or AI is a collection of technologies that can enable a machine*

or system to sense, comprehend, act, and learn" (SDAIA, 2023a: 6).

2.5. AI Ethics

"*A set of values, principles, and techniques to guide moral conduct in developing and using AI technologies*" (SDAIA, 2023a: 6).

2.6. Data Governance

"*Data governance is the process of managing the availability, usability, integrity, and security of data in organizations and systems, based on data standards and policies that also control data usage*" (SDAIA, 2023a: 7). The national data governance policy is based on the data classification policy, the national data protection policy, the data sharing policy, the freedom of information policy, the open data policy, the personal data protection policy for children and persons in similar conditions, the general rules for the transfer of personal data outside the geographical borders of Saudi Arabia (SDAIA, 2021: 27).

2.7. Data Sovereignty Public Policy

The public policy of data sovereignty is based on four principles: data as a national asset, data protection, data availability, and encouragement of local and foreign investment (SDAIA, 2024e: 3-4).

2.8. AI System

"*A set of predictive models and advanced algorithms that can be used to analyze data*

and predict the future or facilitate decision -making for projected future events" (SDAIA, 2023a: 6).

2.9. AI System Lifecycle

"*The cyclical process that AI projects are expected follow to be able to design, build,*

and produce a robust and safe system that delivers business value and insights through adhering to a standard and structured way of managing AI model delivery and implementation" (SDAIA, 2023a: 6).

2.10. National Strategy for Data and AI

The National Strategy for Data and Artificial Intelligence is a comprehensive, long-term action plan that includes defining objectives, selecting appropriate means to achieve them, developing detailed implementation plans, and monitoring and evaluating performance. This requires unifying efforts of "*the key stakeholders (government, private sector, academic institutions, specialist, general workforce, and the wider public) to contribute to its success*" (NSDAI, 2020: 18). Based on this, SDAIA developed this strategy which was approved by the Supreme Council on 17 July 2020.

AI works according to what it was designed, developed, trained, fine-tuned via supervised learning, and used, and AI ethics is about establishing a set of principles and values throughout the lifecycle of an AI system. Since the basis of ethics is the individual before involving other individuals, organizations, and systems (Goodpaster & Nash, 1985: 509), it is worth exploring the link between ethics, individualism, and risk which has not attracted the attention of researchers on the risks of AI in view of the data consulted so far.

3. Exploration of the relationships between ethics, individualism, and risk

The minimal ethic that originates with John Stuart Mill reduces all questions governing conduct to the principle that everyone can live their life as they wish as long as they do no harm to others (Callahan, 2013). This proposition leads to defending a conception of freedom as the absence of external obstacles to the action of an individual (Berlin, 2002), to rejecting any doctrine that

recommends an entire art of living and not only principles of peaceful coexistence, and to sticking to three principles of renewed minimalist ethics: equal consideration of the voice of each as it expresses it, moral indifference of the relationship to oneself, and non-harm to others (Ogien, 2007). This leads to putting society's resources at the service of the autonomy of individuals. In this regard, "*there's no such thing as society*" (Thatcher, 1987). In other words, "*what is society but an individual?*" (Dazai, 1973: 118).

It should be noted that Article 10 of *The Basic Law of Governance* of Saudi Arabia, enacted on 2st March 1992, stipulates that "*the State shall endeavor to strengthen family bonds and maintain Arab and Islamic family values, and shall strive for the welfare of all its members and for creating the conditions conducive to the development of their talents and abilities*". It refers to the family composed of members as the basis of society. Family is not just a group of individuals united by blood or legal ties, but an intertwining of shared relationships, emotions, and experiences.

In this regard, the family is considered the foundation of society. It plays a crucial role in transmitting values, customs, and traditions, ensuring the continuity of society across generations. The society is one of the three themes on which the Saudi Vision 2030 is built: "*a vibrant society, a thriving economy and an ambitious nation*".

Such a perspective sheds light on the invisible connection between the notion of the individual and that of risk in the sense that the discourse on the risks associated with AI defines users—without them realizing it—as individuals. In the context of navigating AI management in an increasingly uncertain economic world (Strategic Gears, 2025), this leads to a fundamental distinction between *ghayr al-mutawaqa'* (unforeseen events), *al-gharar* (hazard), and *al-khatar* (risk) (Belabes, 2017), as summarized in Table 2.

Table 2. Differences between unforeseen events, hazard, and risk

	Unforeseen events	Hazard	Risk
Description	Future event that could impact the project and cannot be anticipated	Future event that can be identified but whose impacts are unknown	Future and uncertain event, with a probability of occurrence, which could impact the project
Features	Unidentifiable, unquantifiable	Identifiable, non-quantifiable	Identifiable, quantifiable
Impacts	Positive or negative, cannot be clearly estimated	Negative, cannot be clearly estimated	Negative, can be estimated, qualified and quantified

Source: Authors' own

A risk is traditionally predictable, quantifiable, and compensable. It fits into a logic that insurance companies can assess and cover (Knight, 1921). Insurance is based on the concept of risk, which must be assessed using mathematical and statistical techniques through the law of large numbers. Such a statistical law states that the random variable M_n satisfies, for any strictly positive real number a:

$$\lim_{n \to \infty} P(\mid M_n - E(X) \mid \geqslant a) = 0$$

This equation demonstrates that the empirical mean is a legitimate estimator of an unknown proportion and allows us to estimate a probability or proportion using numerical simulations.

By contrast, a disaster is an event of such disproportionate magnitude that it is immeasurable, both in its immediate consequences and in its long-term impacts. It is a possible and identifiable future event, but one that cannot be quantified, described in risk management as a hazard. A disaster involves widespread destruction in time and space, requiring an approach far beyond private risk management mechanisms.

It is crucial to distinguish events that bring great damage, loss, or destruction, as well as their impacts on individuals, society, and the planet, because this distinction allows for the development of a structured approach that includes the analysis of the situation, the definition of clear objectives, the development of specific action plans, and the establishment of monitoring and evaluation mechanisms:

• For individuals, the priority is to protect themselves against what directly threatens them.

• For societies, it is important that they take care to avoid being drawn into closed systems that could cause their self-destruction.

• For the planet, it is essential to develop solutions that address the root causes of the problem and strengthen the resilience of ecosystems. This involves a combination of climate change mitigation, disaster prevention, and effective preparedness and response.

This typology reveals the limitations of defining AI-related ethical risk "*as any risk associated with AI that may cause stakeholders to fail in one or more of their ethical responsibilities to other stakeholders*" (Douglas et al., 2025: 2189). While it has the advantage of distinguishing it from other forms of risk, such as social, reputational or legal risk, it focuses on the ability of an individual or organization to make decisions and act in accordance with ethical principles, in a given context. It refers to recognizing the ethical implications of one's actions, being accountable for one's actions and assuming the consequences.

It also reveals the importance of sharing values in relation to the possible menace, more or less predictable, inherent in a situation, an event, or a technology. Solidarity is needed at the family, social, and planetary levels to overcome risks, disasters, and vulnerabilities beyond the individual (Smuha, 2021). This analytical perspective requires a broader framework than ethics, which focuses on behaviors, often at the individual or organizational level. Such a framework encompasses broader aspects that refer to religious, moral, and legal values that influence human behavior and social relationships.

4. The analytical framework of AI ethics in Saudi Arabia through the Pentagon SMLET

Sharīʿah (S), moral (M), law (L), ethics (E), and technical (T) aspects refer to five distinct but often related areas of regulating human behavior in the use of AI systems, which require specific mathematics (Belabes, 2024), allowing to link meanings beyond those often associated with classical statistical and data analysis methods (Belabes, 2025d). While AI operates in complex business ecosystems (Belabes, 2025b), the abuses inherent in the development of AI systems should not be limited to ethical considerations, which have thus far been the focus of a discourse of legitimization centered on the individual (Martin, 2021). Hence the importance of the Pentagon SMLET, which is all the more useful when morality and law are often associated with *Sharīʿah* and when the words moral and ethics are often considered synonymous (SDAIA, 2024a: 3). Such confusion invites us to clearly define the five components of the pentagon as follows:

● *Sharīʿah* —literally 'the correct path'— refers to what Allah has legislated through His Prophet Muhammad —peace and blessings of Allah be upon him— (Ibn Hazm, 1983: 46). This definition includes belief, worship, morals, and transactions to regulate people's relationship with their Lord and with each other.

The *Sharīʿah* refers to *aḥkām* (injunctions), *qawāʿid* (maxims), and *maqāṣid* (purposes) to promote *masāliḥ* (benefits) and prevent *mafāsid* (harms), knowing that repelling harm takes precedence over procuring benefit (Al-Suyūtī, 1983: 87). The reference to *Sharīʿah* highlights the importance of both establishing regulations, standards, and rules, and creating frameworks, policies, and practices to effectively address the challenges arising from the development, deployment, and use of AI.

● Moral refers to the customs, traditions and lifestyles specific to the society.

● Law "*is a set of rules that regulate social relations and which the state compels people to follow, even by force when necessary*" (Al-Sanhuri, 1936: 4).

● Ethics finds its source in the notion of *pharmakon* (Plato, 1868: 134-139), which is at the heart of modern reflections on the relationship between humans and technology (Derrida, 1972; Stiegler, 2007). The Greek term *pharmakon* means both poison and remedy, highlighting the dual potential of any invention: beneficial if thought through with discernment, harmful when adopted thoughtlessly. This ambivalence is fully reflected in the field of AI, whose rapid progress raises ethical challenges.

Compared to the previous definition proposed by SDAIA (2023a: 6), it is preferable to consider ethics as a set of values and principles to guide conduct in the development and use of AI systems beyond what is strictly moral and legal (Foster, Miola, 2015)

● AI technical aspect refers to the technical details of the various methods and technologies of AI and machine learning, which focuses on the development of computer algorithms that automatically improve themselves through experience and the use of data.

In this regard, *Sharīʿah* is the complete code of Muslim life. Morality is a set of rules of conduct, law is a formal system of rules established by an authority, and ethics is a reflection on these rules. The problem usually lies not in ethical principles but in their confrontation with today's complex world, and in their adoption separately, rather than considering the whole system in which they are embedded (Lauer, 2021). The study of some regulatory issues reveals difficulties in moving from principles to more concrete measures and problems in implementing the ethics of AI (Coeckelbergh, 2019). This calls for the implementation of proactive policies based on dialogue with stakeholders and adequate regulations that should not be rigid.
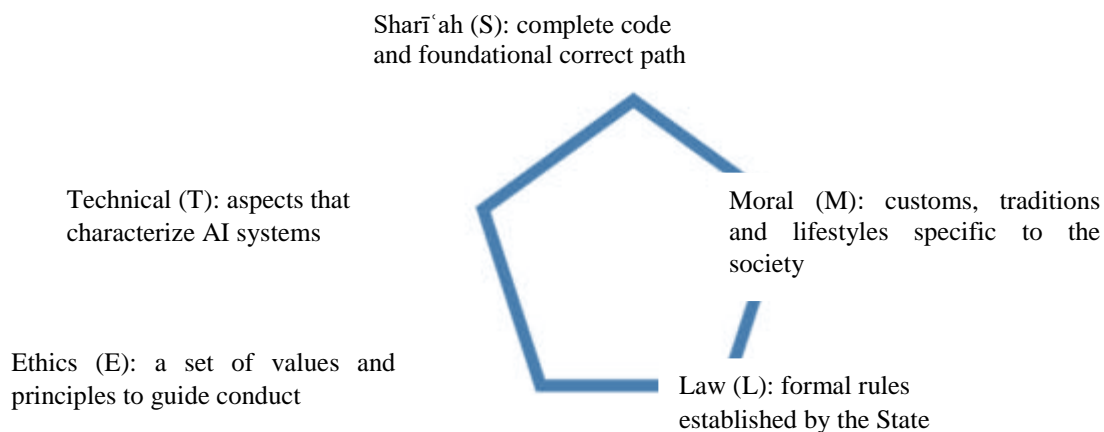
If principles alone cannot guarantee ethical AI (Mittelstadt, 2019), the ethics of AI reflects an ever-growing need to strengthen social control of technology (Héder, 2020). The term 'ethics of AI' tends to encompass all non-technical aspects of algorithmic systems: questions about their uses, their social consequences, or their legal regulations (Beaudouin, Velkovska, 2023: 12). Ethical guidelines are immediately confronted with the fact that the internal processes of AI systems remain hidden from users, which makes it difficult to understand how decisions are made. This poses challenges not only in terms of transparency of AI in its development, practical applicability, but also active engagement of diverse, independent and trustworthy stakeholders (Bélisle-Pipon et al., 2023: 1507).

Considering these five dimensions is all the more important given that digital giants control AI systems and, consequently, information of all kinds, that is, the ways of living and interacting in society and with the world. This has transformed the very fabric of society and the way it interacts with its environment and the outside world. Faced with such a situation, a general awareness is necessary because AI is here to stay and it has not yet reached its end. Ethical principles that guide human conduct can be implemented, but this does not mean that AI can always make correct decisions (Zeller, Atherton, 2025: 10).

Considering these five dimensions requires injunctions, maxims, purposes, rules, values that motivate actions and relationships with others, but also factual data and confrontations between principles and individual, social, and global challenges. In this regard, it is not enough to rely on a single specialty, but it is necessary to combine the knowledge and points of view of different fields to address the issue in depth (Vesnic-Alujevic et al., 2020).

The interest of this Pentagon (Figure 1), which is fundamentally based on the *Sharīʿah*, reflecting the sacred as the foundation of human societies (Godelier, 2010: 123), is to incite a reflection on morality, law, and ethics that goes beyond purely quantitative and commercial considerations to be anchored in human reality, vital needs, and the search for the good for all.

Figure 1. The analytical framework of AI ethics in Saudi Arabia through the Pentagon SMLET

Sharīʿah (S): complete code
and foundational correct path

Technical (T): aspects that
characterize AI systems

Moral (M): customs, traditions
and lifestyles specific to the
society

Ethics (E): a set of values and
principles to guide conduct

Law (L): formal rules
established by the State

Source: Authors' own

This analytical framework could be useful for the Charter that prepares a comprehensive framework for the development of AI technologies in accordance with Islamic values and principles, under the leadership of the Islamic World Educational, Scientific and Cultural Organization (ICESCO), SDAIA, in cooperation with the Saudi National Commission for Education, Culture and Science (ICESCO, 2024). Such a framework is driven by Saudi Arabia's unique positioning in the Arab-Muslim world with over 15 million Hajj and Umrah visitors annually and approximately 420 million Arabic speakers worldwide (NSDAI, 2020: 12).

The Pentagon SMLET invite us to go beyond conceptions that equate morality, law and ethics as a simple matter of transaction or calculation (Weil, 1957: 23), and that hinder societal values such beneficence, nonmaleficence, autonomy, and justice (Pizzolato, 2019: 183; Hagendorf, 2019: 114). This conception of compassion-based justice (Munn, 2023), which contrasts with that of law-

based justice (Bell, 2019), emphasizes duties and attention to others in all aspects of life as a bridge to the beyond (Al-Shatibi, 2009: 2: 63). As the social, environmental, and political impacts of AI increase, AI governance is evolving from values and principles to policies (Schiff et al., 2020: 158).

Furthermore, the Pentagon SMLET differs from the Charter of Ethics of Artificial Intelligence developed by the Arab League Educational, Cultural and Scientific Organization (ALECSO) and adopted on June 17, 2025, which stipulates that "*when developing, distributing, or using an AI system, care should be taken to respect Arab cultural heritage and Islamic beliefs and values, and not to violate, distort, or offend them, including by violating the sacred values of Islam. This principle must also be respected throughout the life cycle of AI systems, including generative AI systems*" (ALECSO, 2025: 18). Beyond respecting Arab cultural heritage and Islamic beliefs and values, the challenge is to design AI systems that embody a holistic approach centered on submission to God, and a life here below that aims for eternal salvation in the hereafter.

The fundamental characteristic of the legal regulation of AI is its relativity; it depends on the conditions of existence of a given society (Filipova, 2024). Its universality in principle does not manage to compensate for a particularity in fact, which requires constant evaluation and an agile approach (Rodrigues, 2020). The direct consequence of this particularity is that it has nothing to do with the good; it is a totally independent fact. It only depends on a power that it would be up to me to actualize; holding the right to perform an action does not oblige me to perform it. The possession of a right implies the possibility of making good or bad use of it. The law is therefore foreign to the good, which implies a transcendence, takes into account rights, duties, and responsibilities (Uhumuavbi, 2025), and consequently exceeds the limits of individual experience and of the material reality that refers to the tangible world.
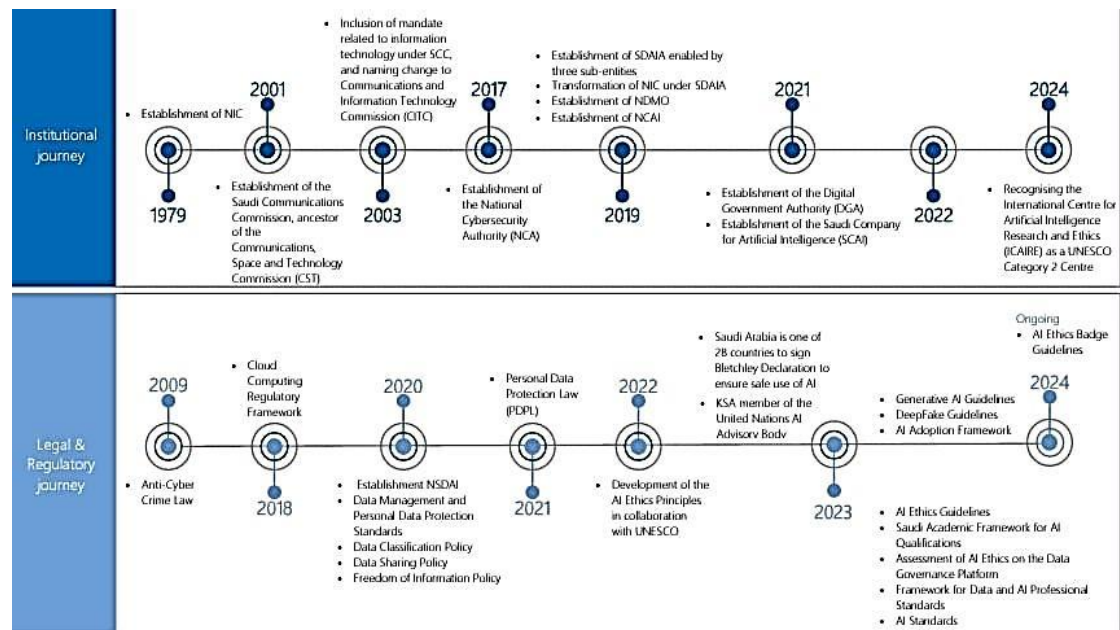
This particularity of the law demonstrates the limitations of AI and data, which focus on fundamental ethical frameworks of AI, such as privacy, accountability, transparency/explainability, fairness, and security, and therefore only cover private sector actors (Kang, 2024). This trend, however laudable, fails to consider the impact of AI on fundamental values and democratic institutions. Indeed, automated drafting, microtargeting, and large-scale social listening techniques can have a harmful effect on the quality of public deliberation (Manheim, Kaplan, 2019). It is not enough to trust companies that use AI to respect ethical principles (Li & Wan, 2025), or to establish complementary legal rules (Hoffmann-Riem, 2020), in the absence of a global framework that defines rights, duties, and responsibilities for a global coordination and global governance of AI (Coeckelbergh, 2025), In this regard, it is necessary to think about the political normativity specific to AI (Murgia, 2025) beyond ethical normativity strictly speaking (Agbavon, 2024). So the question arises: Is UNESCO's global standard on AI ethics, adopted by all 193 of its member states in 23 November 2021, sufficient?

5. Saudi AI ethics, institutional, legal, regulatory, and governance

Since the establishment of the National Information Center (NIC), Saudi Arabia has embraced the world of data, which refers to all structured and unstructured information and knowledge, available and accessible in various formats and sources. With the establishment of the National Data and AI Authority (SDAIA) to drive the ambitious national data and AI agenda and the establishment of the International Centre fir Artificial Research and Ethics (ICAIRE) (Figure 2), Saudi Arabia ranks first regionally and 14th globally in AI, and first globally in the Government AI Strategy, according to the Global AI Index 2024.

Figure 2. The Saudi AI institutional, legal and regulatory journey
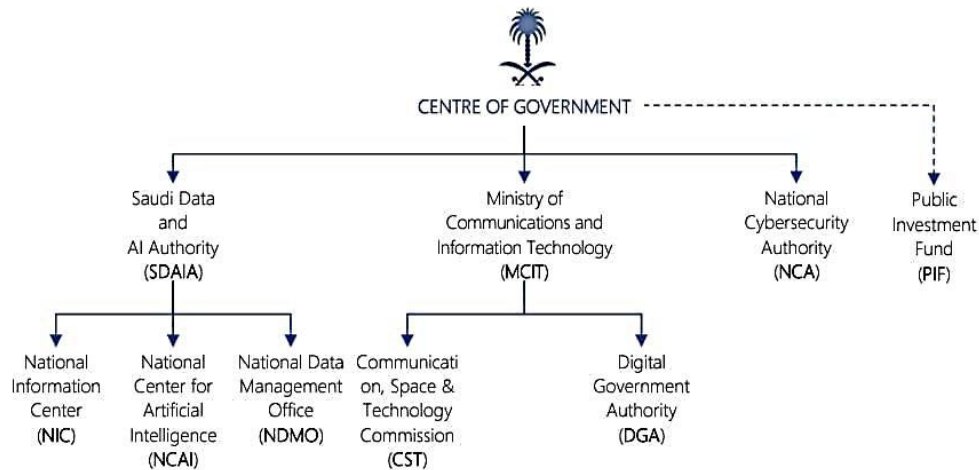

Source: Global AI Index 2024

This success reflects Saudi Arabia's successful efforts to position itself at the forefront of AI innovation at the technical, ethical, legal, regulatory, and governance levels. At the heart of this success is the National Strategy for Data and AI (NSDAI), launched by SDAIA, which sets out a transformative AI roadmap "*for becoming one of the leading economies utilizing and exporting Data & AI after 2030*" (NSDAI, 2020: 7).

This competitive positioning in data and AI is supported by a robust governance structure centered on the SDAIA and its three sub-entities:

• National Information Center (NIC).

• National Data Management Office (NDMO).

• National Center for Artificial Intelligence (NCAI) (NSDAI, 2020: 14).

In collaboration with key government organizations such as the Ministry of Communications and Information Technology (MCIT), the Communications, Space and Technology Commission (CST), the Digital Government Authority (DGA), the National Cybersecurity Authority (NCA), and the Public Investment Fund (PIF), this ecosystem (Figure 3) has propelled Saudi Arabia to become a leading nation in AI innovation at the technical, ethical, legal, regulatory, and governance levels (NSDAI, 2020: 15).
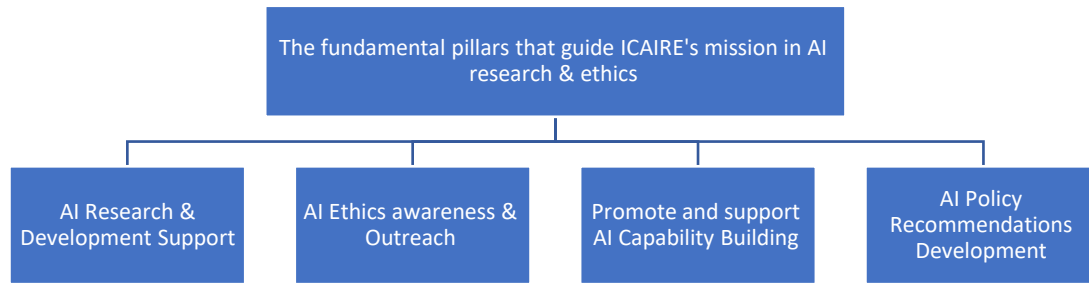
Figure 3. Saudi National AI Governance

Source: UNESCO (2024: 20)

The National Strategy for Data and AI (NSDAI, 2020), led by SDAIA, forms the foundation of Saudi Arabia's ambition to become a global leader in AI. This strategy is also complemented by specific guidelines developed by SDAIA, such as the *AI Ethics Principles* (SDAIA, 2021), the *GenAI Guidelines* (SDAIA, 2024c), and the *Deepfake Guidelines* (SDAIA, 2024d). These AI-specific guidelines are complemented by a robust legal framework including the *Personal Data Protection Act* (SDAIA, 2023), the *Freedom of Information Policy* (SDAIA, 2020), and the *Anti-Cybercrime Law* (BECM, 2007), positioning Saudi Arabia as a pioneer in responsible and secure AI adoption, including a set of regulations: *Organizations' Social Media Accounts* (NCA, 2021a), *Telework Cybersecurity Controls Accounts* (NCA, 2021b), *Critical Systems Cybersecurity Controls Accounts* (NCA, 2019), *Operational Technology Cybersecurity Controls* (NCA, 2022a), *Data Cybersecurity Controls* (NCA, 2022b), *Essential Cybersecurity Controls* (NCA, 2024a), and *Cloud Cybersecurity Controls* (NCA, 2024b).

Regarding AI regulations, which aim to ensure that the development and use of AI products and productions respect fundamental rights and Saudi values, SDAIA was tasked with regulating the data and AI sectors by establishing policies, standards, and rules; processing the relevant regulatory provisions; disseminating them to relevant government and non-government entities; and ensuring their compliance. In this regard, it also has the role of raising awareness of policies, provisions of relevant systems, regulations, and decisions concerning data and AI.

In terms of ethics of AI, Saudi Arabia concluded an agreement with UNESCO and the Riyadh-based International Center for Research and Ethics in Artificial Intelligence (ICAIRE), in 11 September 2024. According to this agreement, ICAIRE conducts research and training activities in AI ethics. This confirms Saudi Arabia's role in promoting international and regional cooperation in AI policy, ethics, and research (SPA, 2024), as illustrated by ICAIRE's four strategic pillars (Figure 4).
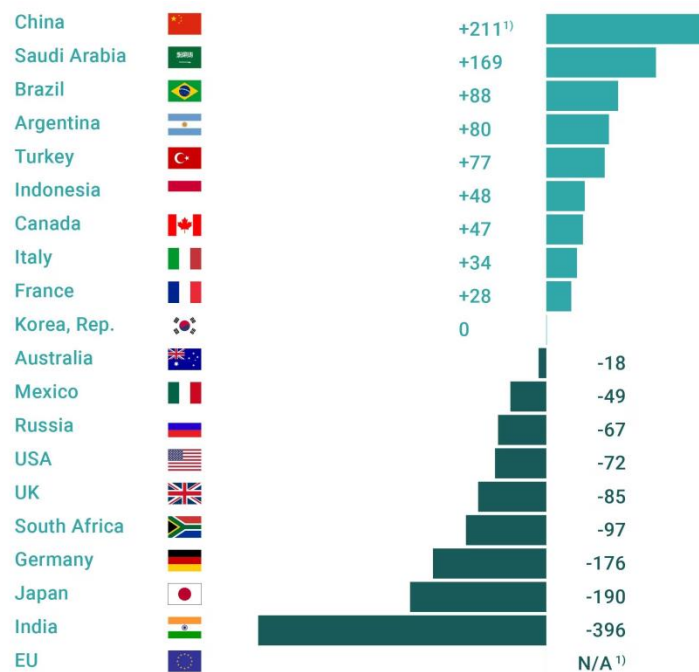
Figure 4. ICAIRE's strategic pillars

Source: ICAIRE website

Saudi policies, standards, and rules on AI aim to regulate the development, placing on the market, and use of AI systems. The regulations establish harmonized measures to ensure that AI systems respect fundamental rights, Saudi values, and security requirements. They are based on a risk-based approach, categorizing AI systems according to their potential impact, from minimal to unacceptable risk. In addition to protecting citizens and limiting harmful uses, the regulations encourage innovation, particularly for startups and small and medium-sized enterprises, and aim to strengthen trust in rapidly expanding AI technologies and raise Saudi digital competitiveness (Figure 5).

Figure 5. The Saudi Digital Competitiveness

[1] Due to missing data for China on the qualitative indicators used in the study changes in these dimensions represent changes between 2020 and 2017. However given the magnitude of changes observed in China these differences in time do not alter the rankings.

[2] The EU is not included since it is a collection of countries.

The objective of policies, standards, and rules is to foster trustworthy AI and ensure respect for the values, security, and fundamental rights of users, and to strengthen adoption, investment, and innovation in AI in Saudi Arabia. As Figure 6 illustrates, there are no AI systems that do not pose risks. In this regard, it is appropriate to distinguish four levels of risk:

● Unacceptable risk AI systems, which are prohibited.

● High-risk AI systems, subject to strict compliance requirements.

● Limited-risk systems, governed by transparency obligations.

● Minimal or no-risk systems, for which no specific obligations are imposed.

Figure 6. The four levels of risk for AI systems

Source: SDAIA (2023: 15)

The principles of AI ethics developed by SDAIA (2023: 12-27) encompass fairness, privacy and security, humanity, social and environmental benefit, reliability and safety, transparency and explainability, accountability and responsibility (Figure 7). These principles aim to ensure that AI is developed and used in ways that respect human values, promote social well-being, and avoid harm.

Figure 7. AI Ethics Principles

Albara Abulaban[1] and Abderrazak Belabes[2]

Source: SDAIA (2023: 12-27)

To achieve trustworthy AI, fundamental ethical principles must be applied and evaluated throughout the AI system's lifecycle. These requirements are interconnected, equally important, and mutually supportive. They are not exhaustive but represent systemic, individual, and societal aspects, as shown in Table 1, which compares Saudi Arabia's ethical principles with those of UNESCO (2022: 20-23), the OECD, and the European Union (Belabes, 2025b).

The human-centered approach to AI aims to place human values at the heart of the development, deployment, use, and monitoring of AI systems, ensuring respect for human dignity, which recognizes that every human being has intrinsic and inherent worth, regardless of their characteristics or status. This leads to the need to take into account the natural environment and other living ecosystems of which humans are an integral part, interacting with them and being affected by them, as well as a sustainable approach allowing the flourishing of future generations.

Table 1. AI Ethical principles of UNESCO, OECD, and European Union

UNESCO	OECD	EU
● Proportionality & do no harm ● Safety & Security ● Right to Privacy & data protection ● Multi-stakeholder, Adaptive Governance & collaboration ● Responsibility & accountability ● Transparency & explainability ● Human oversight & determination ● Sustainability ● Awareness & literacy ● Fairness & non-discrimination	● Inclusive growth, sustainable development & well-being ● Human rights & democratic values, including fairness & privacy ● Transparency & explainability ● Robustness, security & safety ● Accountability	● Human agency & oversight ● Robustness & safety ● Privacy & data governance ● Transparency ● Diversity, non-discrimination & fairness ● Societal & environmental well-being ● Accountability

Source: Authors' own

Through a proactive approach, SDAIA (2023: 29) ensures the review and updating of principles relating to AI ethics and monitors compliance with them. The authority also develops national guides, standards, and guidelines that ensure the effective management and dissemination of AI ethics and enable the achievement of the desired objectives. To this end, it has implemented five tasks as illustrated in Figure 8.

Figure 8. The five tasks to achieve the objectives of AI ethics

Source: SDAIA (2023: 29)

When implementing ethical requirements for AI systems, conflicts may arise between different principles, making certain trade-offs unavoidable. These decisions must be made in a reasoned and transparent manner, based on available technical knowledge, and by assessing the risks to fundamental human rights. If no ethically acceptable trade-off is possible, the system must not be used as is. Decisions must be documented and reassessed based on new data or changing circumstances, and those responsible must bear the consequences of their actions and possibly be subject to sanctions or remediation measures. In the event of unfair negative impact, accessible recourse mechanisms must be provided, with particular attention to vulnerable people.

Although part of a global context aimed at boosting digital competitiveness, Saudi Arabia seeks to remain true to its cultural preferences and its higher level of protection against the social risks posed by AI, including those affecting privacy, data protection, and discrimination. In this regard, Saudi's guidelines promote an AI system that is reliable, legal (compliant with all applicable laws and regulations), ethical (ensuring compliance with fundamental ethical principles), and robust in order to mitigate the negative impacts of AI systems (economic, psychological, social, etc.) and to address potential threats (security, political, etc.) (SDAIA, 2022: 3). This reveals the importance of the proposed analytical framework of AI ethics, which is not limited to purely individual considerations, including threats to national and International security (UN, AI Advisory Body, 2024: 30), threats to the rule of law, and threats to humanity resulting from the loss of human control to malicious AI (Bengio et al., 2024).

6. Analysis of the UNESCO Recommendation on AI ethics in light of the analytical framework

Aware of the "*the profound and dynamic positive and negative impacts of artificial intelligence (AI) on societies, environment, ecosystems and human lives, including the human mind, in part because*

of the new ways in which its use influences human thinking, interaction and decision-making and affects education, human, social and natural sciences, culture, and communication and information" (UNESCO, 2022: 5), and other equally important considerations, UNESCO adopted the Recommendation on the Ethics of AI on 23 November 2021.

Member States are called upon to "*apply on a voluntary basis the provisions of this Recommendation by taking appropriate steps, including whatever legislative or other measures may be required, in conformity with the constitutional practice and governing structures of each State, to give effect within their jurisdictions to the principles and norms of the Recommendation in conformity with international law, including international human rights law*" (UNESCO, 2022: 7). This is what Saudi Arabia did, like the rest of the UNESCO member states that participated in the development of AI ethics recommendations.

Member States are also recommended to "*engage all stakeholders, including business enterprises, to ensure that they play their respective roles in the implementation of this Recommendation; and bring the Recommendation to the attention of the authorities, bodies, research and academic organizations, institutions and organizations in public, private and civil society sectors involved in AI technologies, so that the development and use of AI technologies are guided by both sound scientific research as well as ethical analysis and evaluation*" (UNESCO, 2022: 7).

It should first be noted that, given its status as a UN institution, UNESCO does not, through its Recommendation, aim at a regulatory ambition comparable to that of other national and international initiatives on the ethics of AI, in particular the regulation of the European Parliament and Council (2024). As a normative instrument intended for member States, the UNESCO Recommendation attempts to define specific guidelines, leaving the scope for member States to establish legal and regulatory frameworks that they consider necessary, depending on institutional specificities.

In addition, the UNESCO Recommendation serves as a basis for ethical impact assessments throughout the lifecycle of AI systems. Ethical impact assessment is a process aimed at examining the moral implications of decisions and actions. It helps identify and mitigate ethical risks, ensuring that practices respect fundamental rights and values, which include respect for human dignity, individual freedom, equality, non-discrimination and solidarity, respect for human autonomy, prevention of harm or otherwise causing harm to human beings, fairness, and explainability by transparently communicating the objectives of AI systems.

The lifecycle perspective is welcome. The AI development lifecycle is a structured framework for creating, deploying, and maintaining AI solutions. It includes steps such as problem identification, data preparation, model development, deployment, and ongoing maintenance. The ramifications of the AI lifecycle are as complex as the algorithms developed by specialists in the field. In this regard, the ethical impact of AI algorithms must be assessed at each stage, from data creation to model deployment, to develop a comprehensive approach that addresses and considers all challenges from AI design to production (De Silva, Alahakoon, 2022).

The purpose of the UNESCO Recommendation is to first present the ethical values affected by AI and to specify their link with AI in the proposed principles. These principles are then implemented in areas of political action which concern in particular the impact studies of AI

systems, the establishment of governance mechanisms which allow for effective management and monitoring of AI systems, while guaranteeing their compliance and controlling the associated risks, the protection of the environment and ecosystems aimed at preserving nature and limiting the negative impacts of AI systems on natural environments and biodiversity, culture, education and science. AI is used to select topics, generate hypotheses, design experiments, collect, process, and interpret data, assist in writing and editing research papers (Erduran, 2023).

It is worth noting that the UNESCO Recommendation opposes the narrative that AI systems stimulate economic growth, create new industries, motivate innovation, generate employment opportunities, develop new economic activity niches for young people, support the preservation of cultural heritage, contribute to solving complex and urgent challenges in the areas of finance, health, agriculture, education, and public service delivery, and can be a powerful ally for the alternative economic models, helping its actors achieve their social and environmental objectives more effectively (April, 2023; Ishak, Mohamed, 2023).
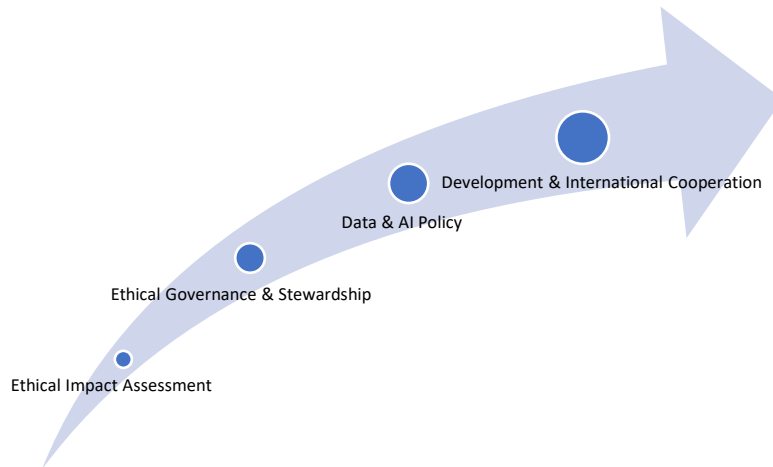
Furthermore, it should be noted that the use of AI systems impoverishes the human mind by achieving the hegemony of a quantitative and cephalocentric approach to intelligence, which emphasizes the brain as the primary seat of intellectual activity (Favier-Baron, 2023), influencing the way individuals think, interact, and make decisions. In this regard, since AI systems do not doubt, do not question, and do not enter into constructive disagreement, which can open up varied perspectives, they reduce the capacity of humans to question, to think, and ultimately to create in the noble sense of the term.

7. Challenges related to implementation of the UNESCO Recommendation on the Ethics of AI in Saudi Arabia

In Saudi Arabia, the implementation of the UNESCO Recommendation on the Ethics of AI has received significant attention since its official publication. Saudi Arabia aims to lead the development of global AI ethics, policy, and research (SPA, 2025). It therefore appears that while Saudi Arabia is taking the UNESCO Recommendations as a starting point, it does not intend to stop there. Implementing a recommendation involves translating the suggested actions into concrete measures, assigning responsibilities, and establishing a timeline. It involves transforming an advisory document into effective and proactive actions. These are actions that are not only effective in achieving a goal but also taken anticipatorily—that is, in anticipation of future needs or potential problems, rather than in reaction to existing situations. Currently, the conditions for implementing the UNESCO Recommendation on AI Ethics in Saudi Arabia encompass four major areas as illustrated in Figure 9.

Figure 9. Major Areas of AI Ethics in Saudi Arabia

Source: Authors' own in light of the data compiled so far

7.1. Ethical Impact Assessment

The UNESCO (2022: 26) Recommendation calls on Member States to implement an ethical impact assessment, which is a structured process to identify and assess the impacts that an AI system may have. In this regard, Saudi Arabia has begun to reflect in terms of potential impact, to identify and assess the benefits and risks of AI systems, and to propose appropriate measures for the prevention, mitigation, and monitoring of threats to avoid their harmful consequences, including security, and safety threats that are increasingly frequent, diverse and complex (SDAIA, 2022: 3).

The competent Saudi authority relies on a graduated typology of risks of AI systems that allows the degree of supervision to be adjusted according to the potential danger of a given system (SDAIA, 2023: 15). While the Saudi state is involved in the development of ethical standards for AI, which would then serve as a reference for impact assessment tools, the field remains open to private, academic, and civil society actors to actively enrich the ethical impact assessment, by developing mechanisms to identify and prevent any negative consequences related to the impact of AI systems and to address them responsibly (NSDAI, 2020: 26).

To strengthen ethical impact assessments of AI in Saudi Arabia, it would be beneficial to review existing ethical documents from the OECD, the Council of Europe, and the EU, recognizing that there are significant differences in how ethical principles for AI are interpreted, why they are deemed important to the issue, field, or actors they are intended for, and how they should be implemented according to the weight of each. The spirit of the UNESCO Recommendation is that AI ethics refers to principles that lead to the construction of trustworthy systems. This leads to local choices based on each country's historical trajectory. As a next step, it would be beneficial to examine Saudi Arabia's added values in AI ethics through a comparative analysis.

7.2. Ethical Governance and Stewardship

Self-assessments by public authorities regarding their use of AI (UNESCO, 2022: 27), are taken into account within the Sudanese public authorities (SDAIA, 2023: 8). This need is retracted as soon as the opportunity arises (SDAIA, 2024b: 4), knowing that the development of AI on a solid

basis that benefits everyone requires clear and applicable rules (Corrêa et al, 2023: 12). While some reservations remain about taking UNESCO's recommendation into account, one thing is certain: the emphasis is now on the law, as well as on the practical operationalization of standards (Daly et al, 2022: 201).

Regarding AI governance (UNESCO, 2022: 27), the relevant Saudi authority recommend "*evaluating the current adherence to regulations and policies and the maturity of internal processes that govern the use of AI technologies and applications*' (SDAIA, 2024b: 7). This involves the integration of the multi-level review systems envisaged in the UNESCO Recommendation, including ethical standards and procedural consultations and reviews.

The current approach to AI process management in Saudi Arabia is clearly open to ensure greater diversity and stakeholder participation in the design and development of AI systems through inclusive approaches to AI governance. This involves establishing systems and practices to ensure that AI projects are carried out efficiently, respecting objectives, budgets, deadlines, quality standards, and ethical principles (SDAIA, 2024b: 3).

7.3. Data and AI Policy

The National Strategy on Data and AI was approved by the Supreme Council on 17 July 2020. SDAIA is working in collaboration with other entities to implement the strategy so that Saudi Arabia can build the foundations for a competitive advantage in key niches domains and compete internationally as a leading economy using and exporting data and AI based on creating long term value through innovation that boosts competitiveness, sustainability, and positive impact (NSDAI, 2020: 19).

In this regard, SDAIA is mandated to develop policies, governance mechanisms, standards and controls related to data and artificial intelligence and to monitor their compliance upon publication, analyze global practices and standards to develop the AI ethics framework (SDAIA, 2023: 4), in accordance with the UNESCO (2022) Recommendation. In this regard, AI policies, guidelines, regulations and ethical frameworks should always be respected.

The challenge is to coordinate international data protection regulations with national data and AI strategies (NSDAI, 2020: 5). This is particularly the case with point 74 of the UNESCO Recommendation, which recommends strengthening existing data protection regulations to protect personal data, including particularly sensitive data, which "*if disclosed, may cause exceptional damage, injury or hardship to individuals* (UNESCO, 2022: 29). In addition, economic, scientific and civil society circles should exchange and collaborate more intensively to optimize the impact of efforts. In this regard, organizations in the AI sector should, in accordance with point 77 of the UNESCO Recommendation, be encouraged to share the data they collect in order to strengthen competitive research and innovation capacities (UNESCO, 2022: 30).

7.4. Development and International Cooperation

The fourth area of action of the UNESCO Recommendation concerns development and international cooperation. Provision 78 addresses UNESCO Member States and transnational corporations and calls on them "*to prioritize AI ethics by including discussions of AI-related ethical issues into relevant international, intergovernmental and multi-stakeholder fora*" (UNESCO, 2022: 30).

The National Data and AI Strategy states that Saudi Arabia aims to become "*the global hub where the best of Data & AI is made reality*". The Kingdom also aims "*to be at the center of a large local and international network of partners*" (NSDAI, 2020: 7). This ambition has been embodied in the establishment of the International Center for Research and Ethics in AI (ICAIRE) in Riyadh, which will serve, in close collaboration with UNESCO, as a "*the lighthouse for ethical AI, empowering nations to advance ethical AI for the good of humanity*".

Through this, Saudi Arabia supports the creation of framework conditions for AI ethics in Arab and Muslim countries through the Riyadh "*Charter on AI Ethics a Moral Compass Anchored in Islamic World Values*" (ICESCO, 2025). This cooperation is in line with UNESCO Recommendation 79, which calls on member States to ensure that the use of AI in key sectors of development cooperation is consistent with the values and principles set out in this recommendation, particularly in the fields of culture, education, science, health, information, communication, agriculture and food supply, environment, natural resource management and infrastructure, economic planning and growth (UNESCO, 2022: 30).

8. Conclusion

Since 2020, Saudi Arabia has been actively contributing to the development of AI ethics through an open and inclusive process that translates ethical principles into concrete policymaking tasks. In this regard, it is building on the UNESCO recommendation, which includes ethical standards for AI systems centered on human rights, accepted by all Member States.

Reading Saudi publications related to data and AI, it is clear that the Kingdom considers AI ethics as a societal mission that calls for a formal commitment to contribute positively to society and the environment, beyond purely economic considerations of AI systems that could stimulate economic growth, contribute to the creation of new jobs, strengthen digital competitiveness and support the attractiveness of foreign direct investments.

For the implementation of this societal mission and its successful achievement, all stakeholders are made aware of the potential and risks of AI, this includes the government, public institutions, AI companies and start-ups, banking and financial markets industry, law firms, judicial courts, charities, *awqaf*, foundations, trusts, academic institutions, research centers, specialized institutes, experts in the field, general workforce and general public.

At the political level, the national data and AI strategy considers AI ethics as a set of directly applicable standards. It thus formulates concrete tasks for national action regarding the development and use of AI in various policy areas, including education and science, communication, mobility, health, energy, and environment. The focus is not on abstract requirements, but on tangible improvements in the quality of the ethics of AI systems.

In this regard, our contribution in this research is to draw attention to the fact that the conceptualization of AI ethics must take into account the historical trajectories of countries, their religious beliefs; and be rooted in socio-cultural contexts and articulate with particular moralities, otherwise the universalist claim of AI ethics risks becoming the imposition of a dominant thought. While ethical effort is to be encouraged to ensure that AI systems are developed and used in a way

that is beneficial to society, it must not replace the legislative framework and moral values in social regulation

The interest of the proposed analytical framework based on the Pentagon SMLET is to start from the fact that AI systems are able to impose behavioral models, which underlie particular worldviews, and ultimately influence decision-making at both the individual and collective levels, based on the fact that the family is the basic unit of society, which plays a crucial role in the formation of the individual and social cohesion. Hence the need to develop an ethics of AI that goes beyond the erection of the individual as the supreme and central value of human life. Highlighting this epistemological presupposition invites the integration of religious, moral, legal, and societal values broader than individual privacy and protection of personal data.

Based on the distinction between unforeseen events, hazards, and risks, the proposed analytical framework distinguishes collective disaster from individual risk in the context of AI systems and advocates a holistic governance approach that prioritizes multi-stakeholder collaboration and proactive policies. In light of this framework, the research analyzes Saudi Arabia's institutional and regulatory framework for AI systems to promote trustworthy AI.

Furthermore, the article compares Saudi Arabia's ethical principles for AI with UNESCO's recommendations, illustrating how the Kingdom not only adopts these guidelines but also develops them by integrating its own cultural and strategic considerations, with the aim of becoming a global leader in responsible AI, considering both *Sharīʿah*, moral, legal, and ethical aspects that must accompany AI systems designed to analyze data, identify patterns, and use this knowledge to make more accurate and informed decisions to ensure the improvement of the quality of life at the individual, family, and societal levels.

In terms of recommendation, the analytical framework based on the Pentagon SMLET could be useful in the enrichment of analytical foundations of the Riyadh Charter on AI ethics (ICESCO, 2024) and the Charter of Ethics of Artificial Intelligence developed by the Arab League Educational, Cultural and Scientific Organization (ALECSO, 2025) for developing AI systems in line with Islamic values and principles.

References

Agbavon, Tiasvi Yao Raoul (2024). L'exigence d'une normativité éthique pour une confiance éclairée en l'IA, *Communication, technologies et développement*, 15. https://doi.org/10.4000/123ir

ALESCO (2025). *ALECSO Ethics Charter for Artificial Intelligence*. Publications of the Arab League Educational, Cultural and Scientific Organization, Tunis.

Al-Shatibi, I. (2009). *Al-Muwafaqat*, Dar Ibn al-Qayyim Dar Ibn Affan, Riyadh and Cairo.

Al-Sunhuri, A. (1936). *ʿIlm Usūl al-Qanūn*. Matba'at Fath Allah Ilyas Nuri, Cairo.

Al-Suyūtī, J. (1983). *Al-Ashbāh wa-al-nazāir fī qawāʿid wa-furūʿ fiqh al-shāfiʿiyah*. Dar al-Kutub al-Ilmiyah, Beirut.

April, K. (2023). AI-induced solidarity economy: The need for stewardship orientation. *Effective Executive*, 26(3), 30–45.

Basic Law of Governance of Saudi Arabia. Royal Order No. A/90, 1st March 1992, https://cutt.ly/arABj7qF

Beaudouin, V. & Velkovska, J. (2023). Enquêter sur l'éthique de l'IA. *Réseaux*, 240(3), 9–27.

BECM (2007). *Anti-Cyber Crime Law*. Bureau of Experts at the Council of Ministers. Riyadh. https://cutt.ly/ArDlJFUM

Belabes, A. (2017). The notion of risk in economics revisited in the light of Arab-Muslim classical literature. *Journal of King Abdulaziz University*: *Islamic Economics*, 30(2), 3–12.

Belabes, A. (2024). The Grothendieck's toposes as the future mathematics of AI. *Philosophy International Journal*, 7(3), 1–8.

Belabes, A. (2025a). *Economic life beyond economists*. Routledge, London and New York.

Belabes, A. (2025b). Islamic finance, Artificial Intelligence, and the debt embedded in the ex nihilo monetary creation system, *International Journal of Multidisciplinary Research and Analysis*, 8(2), 485–491.

Belabes, A. (2025c). The localization of the use of data in endowments in Saudi Arabia in light of the ethics of AI (in Arabic). *Waqf Journal*, 13, 21–82.

Belabes, A. (2025d). Conceptualizing the limits of the use of generative AI in economic analysis via a deep meaning of habs. *Journal of King Abdulaziz University*: *Islamic Economics*, 38(2), 23–38.

Bélisle-Pipon, Jean-Christophe, Monteferrante, Erica, Roy, Marie-Christine & Couture, Vincent. (2023). Artificial intelligence ethics has a black box problem. *AI & Society*, 38, 1507–1522. https://doi.org/10.1007/s00146-021-01380-0

Bell, R.H. (1998). *Simone Weil*: *The Way of Justice as Compassion*. Rowman & Littlefield Publishers, Lanham.

Bengio, Yoshua et al. (2024). International Scientific Report on the Safety of Advanced AI (Interim Report). GOV.UK. https://cutt.ly/NrSCbpWR

Berlin, I. (2002), 'Liberty', in Henry Hardy (ed.), Liberty, Oxford: OUP, 283–286.

Boccara, B. (2022). Simone Weil: Du droit à l'obligation. *Implications philosophiques*, 24 March, https://cutt.ly/qrSezsZ7

Brandtzaeg, Petter Bae, Marita Skjuve, and Asbjørn Følstad (2025). 'AI Individualism: Transforming Social Structures in the Age of Social Artificial Intelligence', in Philipp Hacker (ed.), Oxford Intersections: AI in Society (Oxford, online education, Oxford Academic. https://doi.org/10.1093/ 9780198945215.003.0099

Burgess, J. Peter, Floridi, Luciano, Pols, Aurélie & van den Hoven, Jeroen (2018). *Towards a digital ethics*: *EDPS ethics advisory group*. The Ethics Advisory Group established by the European Data Protection Supervisor, the EU's independent data protection authority, https://philpapers.org/archive /BURTAD-3.pdf

Callahan, D. (2013). 'Minimalist Ethics', The Roots of Bioethics: Health, Progress, Technology, Death (online education, Oxford Academic). https://doi.org/10.1093/acprof:oso/9780199931378. 003.0003

Carrillo, M.R. (2020). Artificial intelligence: From ethics to law. *Telecommunications Policy*, 44(6). https://doi.org/10.1016/j.telpol.2020.101937

Coeckelbergh, M. (2019). Artificial Intelligence: some ethical issues and regulatory challenges. *Technology and Regulation*, 19, 31–34. https://doi.org/10.71265/a9yxhg88

Coeckelbergh, M. (2025). Three challenges for a global AI ethics: towards a more relational normative vision. *AI Ethics*. https://doi.org/10.1007/s43681-025-00791-9

Corrêa, Nicholas Kluge et al. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, 4(10), https://doi.org/10.1016/j.patter.2023.100857

Daly A, Hagendorff T, Hui L, et al. AI, Governance and Ethics: Global Perspectives. In: Micklitz H-W, Pollicino O, Reichman A, Simoncini A, Sartor G, De Gregorio G, eds. Constitutional Challenges in the Algorithmic Society. Cambridge University Press; 2021:182-201.

Dazai, O. (1973). *No Longer Human*. New Directions, New York.

De Silva, Daswin & Alahakoon, Damminda (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, 3(6), 1–13. https://doi.org/10.1016/j.patter.2022.100489

Derrida, J. (1972). *La pharmacie de Platon. In La dissemination.* Seuil, Paris, 69-198.

Douglas, David .M., Lacey, Justine & Howard, David (2025). Ethical risk for AI. *AI Ethics*, 5, 2189–2203. https://doi.org/10.1007/s43681-024-00549-9

Erduran, S. (2023). AI is transforming how science is done. Science education must reflect this change. *Science,* 382(6677). 10.1126/science.adm9788

European Parliament and Council (2024). Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence. *Official Journal of the European Union*. 12.7.2024, 1–144. https://cutt.ly/WrS4URvQ

Favier-Baron, E. (2023). Que fait l'Intelligence Artificielle à l'intelligence? *Appareil*, 26, 1–13. https://doi.org/10.4000/appareil.6943

Filipova, I.A. (2024). Legal Regulation of Artificial Intelligence: Experience of China. *Journal of Digital Technologies and Law*, 2(1):46–73. https://doi.org/10.21202/jdtl.2024.4

Foster, C. & Miola, J. (2015). Who's in charge? the relationship between medical law, medical ethics, and medical morality?, *Medical Law Review*, 23(4): 505–530, https://doi.org/10.1093/medlaw/ fwv004

Godelier, M. (2010). *Au fondement des sociétés humaines. Ce que nous apprend l'anthropologie*. Flammarion, Paris.

Goodpaster, K.E. and Nash, L.L. (1985) . *Policies and persons*: *A casebook in business ethics*. McGraw Hill, New York.

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds & Machines*, 30, 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hoffmann-Riem, W. (2020). Artificial Intelligence as a Challenge for Law and Regulation. In T. Wischmeyer, & T. Rademacher (Eds.), *Regulating Artificial Intelligence* (pp. 1–29). Springer, Cham.

Ibn Hazm, A. (1983). *Al-Ihkām fī Usūl al-Ahkām*. Dar al-Afaq al-Jadida, Beirut.

ICESCO (2024). At the Global AI Summit: ICESCO Director General Announces the launch of Riyadh AI Ethics Charter for the Islamic World. https://cutt.ly/xrDxiNH1

Ishak, Muhammad Nur & Mohamed Adjila (2023). Harmonization of Islamic Economics With Artificial Intelligence: Towards an Ethical and Innovative Economic Paradigm. *Al-Kharaj*: *Journal of Islamic Economic and Business*, 5(4), 457–474. https://doi.org/10.24256/kharaj.v5i4.4387

Kang, Sun Gyoo (2024). Analysis of Artificial Intelligence and Data Act Based on Ethical Frameworks. *International Journal of Law, Ethics, and Technology*, 4(4): 1–28.

King Jennifer and Meinhardt, Carline (2024). Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World, White Paper, Stanford University, Human-Centered Artificial Intelligence, https://cutt.ly/SrS4kBu9

Knight, F. (1921). *Risk, Uncertainty, and Profit*. Houghton Mifflin Company, Boston and New York.

Lauer, D. (2021). You cannot have AI ethics without ethics. *AI Ethics*, 1, 21–25. https://doi.org/10.1007/s43681-020-00013-4

Li, Z. and Wa, X. (2025). Ethical challenges and innovations in AI-driven predictive policing: the case of China, *Law, Ethics & Technology*, 2(2), 1–10.

Manheim, K. and Kaplan, L. (2019). Artificial Intelligence: Risks to Privacy and Democracy. *The Yale Journal of Law & Technology,* 21, 106–188.

Martin, E. (2021). L'éthique de l'intelligence artificielle, ou la misère de la philosophie 2.0 à l'ère de la quatrième révolution industrielle. *Cahiers Société*, (3), 189–218. https://doi.org/10.7202/1090182ar

Ménissier, T. (2023). Les quatre éthiques de l'intelligence artificielle. *Varia*, 13(2). https://doi.org/10.4000/rac.29961

Mihály, H. (2020). A criticism of AI ethics guidelines. *Információs Társadalom*, 20(4), 57–73. https://dx.doi.org/10.22503/inftars.XX.2020.4.5

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1, 501–507. https://doi.org/10.1038/s42256-019-0114-4

Munn, L. (2023). The uselessness of AI ethics. *AI Ethics*, 3, 869–877. https://doi.org/10.1007/s43681-022-00209-w

Murgia, M. (2025). Overcoming AI ethics, towards AI realism. *AI Ethics*, 5, 2257–2262. https://doi.org/10.1007/s43681-024-00552-0

NCA (2019). Critical Systems Cybersecurity Controls. https://cutt.ly/xrDliAEV

NCA (2021a). Organizations' Social Media Accounts. https://cutt.ly/VrDlyo7f

NCA (2021b). Telework Cybersecurity Controls. https://cutt.ly/7rDluhRT

NCA (2022a). Operational Technology Cybersecurity Controls. https://nca.gov.sa/otcc_en.pdf

NCA (2022b). Data Cybersecurity Controls. https://cutt.ly/VrDlazRF

NCA (2024a). Essential Cybersecurity Controls. https://cutt.ly/XrDlscZN

NCA (2024b). Cloud Cybersecurity Controls. https://cutt.ly/vrDlrFQ1

NSDAI (2020). *Realizing Our Best Tomorrow*. Strategy Narrative. National Strategy for Data and AI, Riyadh. https://cutt.ly/4rDuHOwP

Ogien, R. (2007). *L'éthique aujourd'hui*: *Maximalistes et minimalists*. Folio, Paris.

Pizzolato, F. (2019). "Les droits font-ils obstacle à la justice?". In *Mondialisation ou globalisation? Les leçons de Simone Weil*, edited by Alain Supiot, Collège de France, https://doi.org/10.4000/books.cdf.6076

Plato (1868). *The Phaedrus of Plato*. Whittaker & George Bell, London.

Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 4, https://doi.org/10.1016/j.jrt.2020.100005

Sadin, E. (2015). *La Vie algorithmique*: *Critique de la raison numérique*. L'Échappée, Paris.

Schiff, Daniel, Biddle, Justin, Borenstein, Jason & Laas, Kelly (2020). What's Next for AI Ethics, Policy, and Governance? A Global Overview. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 153–158. https://doi.org/10.1145/3375627.3375804

SDAIA (2020). Freedom Information Policy. https://cutt.ly/IrDk6xyp

SDAIA (2021). National Data Governance Policies (in Arabic). https://cutt.ly/YrDcgOzY

SDAIA (2022). *AI Ethics Principles*. Version 1.0, August. The Saudi Data and Artificial Intelligence Authority, Riyadh. https://cutt.ly/IrSCaSol

SDAIA (2023a). *AI Ethics Principles*. Version 1.0, Septembre. The Saudi Data and Artificial Intelligence Authority, Riyadh. https://cutt.ly/ArSUr3vY

SDAIA (2023b). Personal Data Protection. https://cutt.ly/CrDk7Io0

SDAIA (2024a). *AI Adoption Framework* (in Arabic). The Saudi Data and Artificial Intelligence Authority, Riyadh. https://cutt.ly/mrDiXQjL

SDAIA (2024b). *AI Adoption Framework* (in English). The Saudi Data and Artificial Intelligence Authority, Riyadh. https://cutt.ly/brDiXpUi

SDAIA (2024c). Generative Artificial Intelligence Guidelines. Public. Version 1. January. https://cutt.ly/FrDk8yM5

SDAIA (2024d). Deepfakes Guidelines. Version 1. https://cutt.ly/6rDk4i42

SDAIA (2024e). The Draft Data Sovereignty Public Policy. https://cutt.ly/0rDx3VLm

SDAIA (2024f). Regulation on Personal Data Transfer Outside the Kingdom. https://cutt.ly/qrFZnIiG

Smuha, N.A. (2021). Beyond the individual: governing AI's societal harm. *Internet Policy Review*, 10(3), 1-31. https://doi.org/10.14763/2021.3.1574

SPA (2024). ICAIRE Classified as UNESCO Category 2 Center, The Saudi Press Agency, 11 September, https://www.spa.gov.sa/en/N2170054

SPA (2025). SDAIA Supports Saudi Arabia's Leadership in Shaping Global AI Ethics, Policy, and Research. Saudi Press Agency, Riyadh, 27 July. https://cutt.ly/arDyeFeH

Stahl, Bernd Carsten & Brooks, Laurence & Hatzakis, Tally & Santiago, Nicole & Wright, David (2023). Exploring Ethics and Human Rights in Artificial Intelligence – A Delphi study. *Technological Forecasting and Social Change*, 91. https://doi.org/10.1016/j.techfore.2023.122502

Stiegler, B. (2007). Questions de pharmacologie générale. Il n'y a pas de simple pharmakon. *Psychotropes*, 13(3-4), 27-54.

Strategic Gears (2025). *KSA CEO Outlook 2025*: *Navigating Artificial Intelligence And Economic Uncertainty*. July, Riyadh. https://cutt.ly/2rDAknMW

Thatcher, Margaret. 1987. 'Interview for "Woman's Own" ("No Such Thing as Society").' in Margaret Thatcher Foundation: Speeches, Interviews and Other Statements. London, https://cutt.ly/ArS4kaWE

The European Center for Digital Competitiveness (2021). Digital Riser Report 2021. Berlin. https://cutt.ly/irSYDBhz

Uhumuavbi, I. (2025). An Adaptive Conceptualisation of Artificial Intelligence and the Law, Regulation and Ethics. *Laws*, 14(2), 19. https://doi.org/10.3390/laws14020019

UN & AI Advisory Body (2024). *Governing AI for Humanity*. Final Report. September. Published by the High-level Advisory Body on AI, United Nations, Washington DC. https://cutt.ly/2rDDvupP

UN (2022). Principles underpinning privacy and the protection of personal data. United Nations, Special Rapporteur on the right to privacy, A/77/196. https://cutt.ly/lrSezMBx

UNESCO (2022*). Recommendation on the Ethics of Artificial Intelligence*. Adopted on 23 November 2021. Publication of the United Nations Educational, Scientific and Cultural Organization, Paris. https://cutt.ly/ErS4gHzw

Vesnic-Alujevic, Lucia, Nascimento, Susana & Pólvora, Alexandre (2020). Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks. *Journal of Responsible Technology*, 6, https://doi.org/10.1016/j.telpol.2020.101961

Walz, A. and Firth-Butterfield, K. (2019). Implementing ethics into artificial intelligence: A contribution, from a legal perspective, to the development of an AI governance regime. *Duke Law & Technology Review*, 18, 176–231.

Weil, S. (1957). *Écrits de Londres et dernières lettres*. Gallimard, Paris.

Zeller, B. and Atherton, M. (2025). Risk-based ethics—myth or reality? *Law, Ethics & Technology*, 2(1), 1–13.

Websites in Arabic:

الهيئة السعودية للبيانات والذكاء الاصطناعي (١٤٤١). الترتيبات التنظيمية للهيئة السعودية للبيانات والذكاء الاصطناعي، الصادر بقرار مجلس الوزراء رقم (٢٩٢) وتاريخ ١٤٤١/٤/٢٧هـ. https://cutt.ly/VrSTVYzv

الاستراتيجية الوطنية للبيانات والذكاء الاصطناعي. https://cutt.ly/prDu1MtQ

وثيقة الاستراتيجية الوطنية للبيانات والذكاء الاصطناعي.https://cutt.ly/SrDu5peP

إطار تبني الذكاء الاصطناعي. https://cutt.ly/MrDiIYzj

الهيئة السعودية للبيانات والذكاء الاصطناعي. https://cutt.ly/4rSYZHnC

المركز الدولي لأبحاث وأخلاقيات الذكاء الاصطناعي. https://cutt.ly/6rSYK46k

رؤية المملكة ٢٠٣٠. https://cutt.ly/3rSYJChO

النظام الأساسي للحكم. https://cutt.ly/VrSYHcPs

إطار تحليلي لأخلاقيات الذكاء الاصطناعي في المملكة العربية السعودية

ما وراء الحقوق الفردية

البراء أبو لبن[1] وعبدالرزاق بلعباس[2]

[1] أستاذ قانون الأعمال الدولي المساعد، معهد الاقتصاد الإسلامي، جامعة الملك عبدالعزيز، جدة؛ aabulaban@kau.edu.sa

[2] أستاذ الاقتصاد الإسلامي، معهد الاقتصاد الإسلامي، جامعة الملك عبدالعزيز، جدة؛ abelabes@kau.edu.sa

ملخص. يهدف البحث إلى توفير إطار تحليلي لأخلاقيات الذكاء الاصطناعي في المملكة العربية السعودية، وُسِمَ بِ: "الخماسي الشريعة، الأخلاق، القانون، الأخلاقيات، الجانب التقني"، لتوجيه تطوير الذكاء الاصطناعي ونشره، وتعظيم فوائده وتقليل أضراره. واستنادًا إلى التمييز بين الأحداث غير المتوقعة والمخاطر، من المهم بمكان التفريق بين الكوارث الجماعية والمخاطر الفردية الناجمة عن انتشار استخدام أنظمة الذكاء الاصطناعي. ويدعو البحث إلى اتباع نهج حوكمة شامل يُعطي الأولوية للتعاون بين الجهات ذات المصلحة المتعددة والسياسات الاستباقية التي تعمل على توقع المشكلات واتخاذ إجراءات لمنع حدوثها أو تخفيف تأثيرها قبل وقوعه بدلاً من الاكتفاء بردود الفعل. وفي ضوء هذا الإطار، يُحلّل البحث التقدم الذي أحرزته المملكة العربية السعودية والإطار التنظيمي لأنظمة الذكاء الاصطناعي لتعزيز موثوقية هذه الأنظمة. كما يقارن البحث المبادئ الخلقية للذكاء الاصطناعي في المملكة العربية السعودية بتوصيات اليونسكو، موضحًا كيف لا تكتفي المملكة باعتماد هذه المبادئ التوجيهية فحسب، بل تُطورها أيضًا من خلال دمج خصائصها الثقافية واعتباراتها الاستراتيجية، لتصبح رائدة عالمية في تطوير وسياسة الذكاء الاصطناعي المسؤول. من حيث التوصيات، يُمكن أن يكون النموذج التحليلي مفيدًا في إثراء الأسس التحليلية لميثاق الرياض لأخلاقيات الذكاء الاصطناعي الذي أطلقته منظمة العالم الإسلامي للتربية والعلم والثقافة بالتعاون مع الهيئة السعودية للبيانات والذكاء الاصطناعي، وكذلك ميثاق أخلاقيات الذكاء الاصطناعي الذي وضعته المنظمة العربية للتربية والثقافة والعلوم، بهدف تطوير أنظمة للذكاء الاصطناعي تتماشى مع القيم والمبادئ الإسلامية.

الكلمات المفتاحية: الذكاء الاصطناعي، الأخلاقيات، القانون، التنظيم، الحوكمة