Demystifying ethical AI: A guide to modern-day governance

> **"** Law is order, and good law is good order. **"**
>
> - Aristotle

As artificial intelligence (AI) continues to evolve at an unprecedented pace, global discussions increasingly stress the urgent need for regulatory frameworks to ensure AI serves society responsibly. Policymakers and industry leaders worldwide recognize that without well-structured laws, AI could disrupt economies, infringe on rights, and introduce unforeseen risks. The question is no longer whether AI should be regulated, but how to strike the right balance between innovation and control.

When considering the negative impacts of AI, issues such as biases in automated decisions, invasions of privacy, and job displacement arise. AI now plays a crucial role in decisions that affect individuals, such as hiring, loan eligibility, and even medical treatments. The lack of regulation heightens the potential risks posed by AI bias, making the situation quite concerning. It is crucial to focus on these negative implications to understand the significant ethical challenges organizations face when seeking to harness these technologies.

Digital ethics addresses how to respond to these challenges and ensure that AI systems are designed and used responsibly. This article seeks to explore the landscape of digital ethics by examining the first regulatory framework for these types of technologies, the European (EU) AI Act, and utilizing this advancement in regulation as a benchmark to analyze the ethical guidelines available to organizations wishing to utilize these technologies in the Middle East (ME).

This article kicks off a series exploring the regulatory landscape for organizations aiming to start their journey in a compliant and ethical way. By comparing differences and focusing on a Middle Eastern approach, insights can be gained into how various regions handle the ethical aspects of AI and digital technologies. Additionally, this article aims to help organizations tap into the resources available in the region that emphasize the importance of ethical considerations in AI and digital technologies.

Background on the AI Act

The EU AI Act (the AI Act), introduced by the European Commission, represents a significant step toward regulating AI technologies within the European Union (EU). Aimed at ensuring AI systems are safe, trustworthy, and respectful of fundamental rights, the Act is structured around key pillars to promote transparency and accountability throughout the AI lifecycle. It categorizes AI systems based on risk, detailing stringent requirements for high-risk AI applications. The Act mandates transparency, establishes oversight and governance mechanisms, and imposes conformity assessments to ensure proper measures are taken during development and deployment. These key pillars will be explored in more detail below.

The AI Act firstly introduces a risk-based approach to categorizing technologies based on their functionality and impact, prohibiting high-risk practices such as those that manipulate behavior, exploit vulnerabilities, or create and use real-time remote biometric identification systems (Article 5). By prohibiting these high-risk practices, the AI Act sets a clear boundary for AI applications that pose significant dangers. This distinction is crucial, as it addresses both present and potential future challenges, paving the way for responsible AI development and deployment.

The AI Act undertakes the ambitious task of regulating the entire lifecycle of AI technologies, from conception through to deployment. It assigns distinct roles and responsibilities at each stage, ensuring that every player in the AI ecosystem, from developers to end-users, is held accountable. This comprehensive oversight is intended to ensure that AI systems are not only innovative but also ethical and secure.

Consumer protection is at the heart of the AI Act. By prohibiting manipulative AI practices and the exploitation of vulnerabilities, the AI Act ensures that consumers are shielded from undue influence and harm. This focus on consumer protection is vital in an era where AI systems can deeply impact personal decisions and everyday life.

Data governance is another critical area addressed by the AI Act. The requirements for accuracy, transparency, and security in high-risk AI systems are particularly stringent when it comes to data management and bias detection. These measures are designed to root out inaccuracies and ensure that AI systems function as intended without unintended harmful consequences.

Privacy remains a top priority under the AI Act, as reflected throughout the legislation. The Act makes multiple references to the General Data Protection Regulation (GDPR), requiring entities to ensure that AI systems protect personal data as per GDPR requirements. It mandates that personal data be processed in strict compliance with existing data protection laws. This is coupled with stringent conditions for handling sensitive data, ensuring that individual privacy rights are respected throughout the AI lifecycle.

The objectives of the AI Act are clear and multifaceted. It aims to ensure safety and uphold fundamental rights by banning certain high-risk AI applications and implementing strict guidelines for high-risk systems. Transparency and accountability are further promoted through mandatory information and logging requirements for deployers of high-risk AI systems. Additionally, the AI Act fosters innovation by including regulatory sandboxes—secure environments where new AI technologies can develop under close scrutiny.

Several core principles underpin the AI Act. Proportionality and non-discrimination are key, ensuring that AI systems do not unfairly affect individuals or groups based on their behavior or characteristics.

Transparency is another vital principle, requiring deployers to inform individuals about the operation of AI, especially in high-risk applications like biometric categorization and emotion recognition. Lastly, data protection remains central, with the AI Act aligning closely with existing EU regulations to ensure that AI systems treat personal data with the utmost care and security.

With over 300 AI-related laws, guidelines, and regulations either implemented or under development worldwide, AI systems must meet stringent security standards. The AI Act sets a precedent for a balanced, ethical approach to AI, providing a robust framework that can serve as a benchmark for regions worldwide. By establishing stringent standards and clear guidelines, AI technologies will be able to benefit society while mitigating potential risks and addressing ethical concerns. A key aspect of the AI Act is the requirement for AI systems to handle personal data in ways that protect privacy, in line with GDPR requirements.

Overview of digital ethics guidelines in the Middle East

Although the AI Act establishes a balanced and ethical framework for AI that can serve as a global benchmark, it's important to recognize that regulators in other regions have also made significant strides. Governments worldwide have recognized the need to shift from voluntary AI ethics frameworks to embedding enforceable regulations in their respective landscape in order to ensure there are no delays in alignment. Efforts to rigorously implement ethical considerations for organizations within their jurisdictions are advancing notably in UAE, Saudi Arabia, Qatar, and Oman. While noteworthy advancements have been made in the region, two specific frameworks from this region will be highlighted to emphasize these advancements.

The United Arab Emirates' Minister of State for Artificial Intelligence, Digital Economy, and Remote Work Applications Office has most recently issued the UAE Charter for the Development and Use of Artificial Intelligence (the Charter).[1] This initiative aims to transform the UAE into a global hub for AI development, reflecting its leadership's vision. It focuses on ethical use, privacy, and compliance with existing legislation, promoting innovation while safeguarding rights.

Governments worldwide have recognized the need to shift from voluntary AI ethics frameworks to embedding enforceable regulations in their respective landscape in order to ensure there are no delays in alignment

A critical aspect of the Charter is its focus on ethical and responsible use of AI, ensuring that all applications are developed and deployed with strict adherence to privacy and data security principles that are already active in the regulatory landscape of the country. By balancing technological advancement with social values, the Charter aims to encourage innovation while protecting individual rights (similar to the AI Act) and promoting economic growth. The UAE seeks to create an environment where AI technologies can thrive safely and securely, building public trust and fostering transparency and accountability in their usage.

One of the key objectives of the Charter is to strengthen human-machine ties, ensuring that AI developments prioritize human well-being and progress. It also emphasizes the irreplaceable value of human judgment and oversight in AI, aligning AI applications with ethical values and social standards to correct any errors or biases that may arise. Governance and accountability are crucial, with the UAE adopting a proactive stance to ensure that AI technologies are used ethically and transparently. Safety of the technology's use is a top priority, with stringent standards required for all AI systems, promoting the modification or removal of systems that pose risks and ensuring that AI technologies do not compromise safety standards. The Charter advocates for the responsible development of inclusive AI technologies that support diversity, respect individual differences, and ensure equal technological benefits for all.

Keeping in mind the purpose of utilizing these AI technologies, the Charter promotes technological excellence as a key consideration, reflecting the UAE's vision for digital, technological, and scientific advancement. The Charter aims for global leadership in AI by driving innovation, enhancing competitiveness, and improving quality of life through effective solutions to complex challenges.

Saudi Arabia is also making significant strides in the ethical adoption of artificial intelligence through its new AI Adoption Framework,[2] setting a benchmark that aligns closely with Europe's comprehensive approach. Along with the release of the framework, the Saudi Data and AI Authority (SDAIA) established 23 AI-dedicated offices across various governmental sectors. These offices aim to incorporate AI into key areas of operations within the government, building on the well-known efforts of the country's Vision of Saudi 2030, which focuses on economic diversification and modernization. By integrating AI with national objectives, Saudi Arabia aims to maximize the technology's societal benefits while ensuring its responsible use.

The AI Adoption Framework acts as a strategic roadmap for organizations navigating the intricate process of AI integration. It addresses the needs of all stakeholders, from top executives to mid-level managers, and lays out a phased approach, starting with the identification of priorities, creation of AI-centric units, and evaluation of organizational preparedness. This phased strategy ensures that AI initiatives are in line with organizational goals and are flexible enough to meet the demands of a rapidly changing technological environment.

A cornerstone of the framework is its strong focus on data as the backbone of AI development. It calls for robust data infrastructure and significant investment in specialist training, recognizing the need for both technical robustness and skilled human resources. This focus leverages existing national regulations and stresses that organizations wishing to adopt AI must have mature data governance and data privacy systems in place before advancing in this field. Organizations need to ensure their data practices are well-structured and compliant to effectively and responsibly harness AI technology.

Saudi Arabia is also making significant strides in the ethical adoption of artificial intelligence through its new AI Adoption Framework, setting a benchmark that aligns closely with Europe's comprehensive approach

Expanding on the AI Ethics Principles[3] introduced by the SDAIA in 2023, the framework reinforces commitments to ethics, transparency, and human-centric AI values. The guidelines cover privacy, security, accountability, and social responsibility, ensuring that AI benefits both individuals and communities. Additionally, a maturity model within the framework classifies AI adoption into four phases—Emerging, Developed, Proficient, and Advanced—providing organizations with a clear path for development.

Additionally, SDAIA has released Generative AI Guidelines for both government employees[4] and the public.[5] These guidelines offer advice on adopting and using Generative AI systems, including examples of common scenarios that organizations might encounter. They also outline the challenges and considerations of using Generative AI, propose principles for responsible use, and suggest best practices.

With the introduction of this framework, Saudi Arabia positions itself as a leader in balancing technological innovation with ethical governance. For businesses and organizations, this provides a supportive environment that fosters innovation, mandates compliance, and delivers tangible benefits across multiple sectors. This initiative not only drives technological progress but also ensures that advancements are made without compromising core values, setting a new standard for AI integration globally.

Implementation

Having explored the regulatory landscapes of some Middle Eastern countries in depth, it is also important to consider a holistic view of what organizations must consider before implementation, as several critical steps are essential for successful integration.

First, building a robust data infrastructure is paramount, ensuring that data systems are strong and reliable. Organizations must also align with existing regulations, adhering to established data governance and privacy standards, and ensuring a high level of maturity in their practices to comply with the ethical standards established by the frameworks discussed earlier. This should be complemented by substantial investments in specialist training to develop the skilled human resources needed for managing AI technologies and ensuring AI fluency. Ethical considerations should be a cornerstone of AI strategies, with guidelines on privacy, security, and accountability seamlessly integrated into organizational practices. Developing dedicated AI-centric units within the organization can help focus efforts on AI adoption, while a thorough evaluation of organizational readiness will ensure that the necessary technical and human capacities are in place.

Impact on innovation

As with all creative endeavors, the relationship between regulation and AI innovation is intricate and multifaceted. On one hand, government efforts can provide a structured environment that promotes responsible innovation, helping to build public trust in AI technologies. Consumers are reassured that these systems are being developed and used responsibly, fostering a more supportive environment for AI adoption and driving further advancements in the field. Moreover, regulation can address ethical and social challenges associated with AI, such as minimizing bias and discrimination, promoting fairness, and ensuring data privacy. In simpler terms, isn't it better to be safe than sorry?

However, regulation also presents significant challenges. Overly restrictive regulations can stifle creativity and slow down the development of new technologies by creating barriers to entry for startups

and smaller companies. This limitation can reduce the diversity of AI solutions and hinder the pace of innovation. Additionally, the rapidly evolving nature of AI makes it challenging for regulators to keep up and even understand the very nature of the systems they are trying to regulate. This can result in regulations becoming outdated and unable to address new challenges and opportunities as AI continues to evolve. Therefore, while regulation is crucial for fostering ethical and responsible AI innovation, it must be adaptable and balanced to avoid hindering technological progress. Emphasizing governing principles such as accountability, transparency, fairness, and safety can provide the pillars upon which agile regulatory frameworks can be built, effectively accommodating the technological advancements they aim to regulate.

The discussion above only begins to unravel the complex regulatory landscape surrounding AI. To fully comprehend the path forward, we must explore the efforts of individual countries within the region to regulate this intricate domain. This deeper understanding is essential for organizations and governments to collaborate effectively in building a shared future. As these developments unfold, comprehensive insights and strategic guidance will be provided to help practitioners navigate the evolving AI regulations effectively. ●

By **Eliza Lozan**, Partner, **Sarah Yahia**, Assistant Manager, and **Aws AlMasri**, Consultant, Digital Privacy & Trust, Deloitte Middle East

Endnotes
1. United Arab Emirates Legislations.
2. sdaia.gov.sa/en/SDAIA/about/Files/ AIAdoptionFramework.pdf.
3. sdaia.gov.sa/en/SDAIA/about/Documents/ai-principles. pdf.
4. GenAIGuidelinesForGovernmentENCompressed.pdf.
5. GenerativeAIPublicEN.pdf.