وزارة الإتصـــــــالات وتكنولوجيـــــــا المعلومـــــــات
Ministry of Communications and Information Technology
State of Qatar • دولـــة قطـــر

Artificial Intelligence in Qatar – Principles and Guidelines for Ethical Development and Deployment

Contents

Overview

Scope of this document

The guidance in this document is aimed at those developing and/or deploying AI systems[1] that will affect the public.

It is legally non-binding, and adherence to it is voluntary. However, all stakeholders are encouraged to consider how the broad principles outlined below might help uphold Qatar's aspiration.

Due to AI's diverse range of potential applications, organizations and entities are encouraged to tailor the guidelines to suit their individual circumstances and operating contexts.

This document will be updated every 1-2 years to keep pace with emerging technologies and ensure the ethical development and deployment of AI systems in Qatar.

Introduction

Qatar's aspiration to promote the responsible and ethical development of artificial intelligence (AI) to support human, social, economic, and environmental development gives rise to the need for AI principles and guidelines.

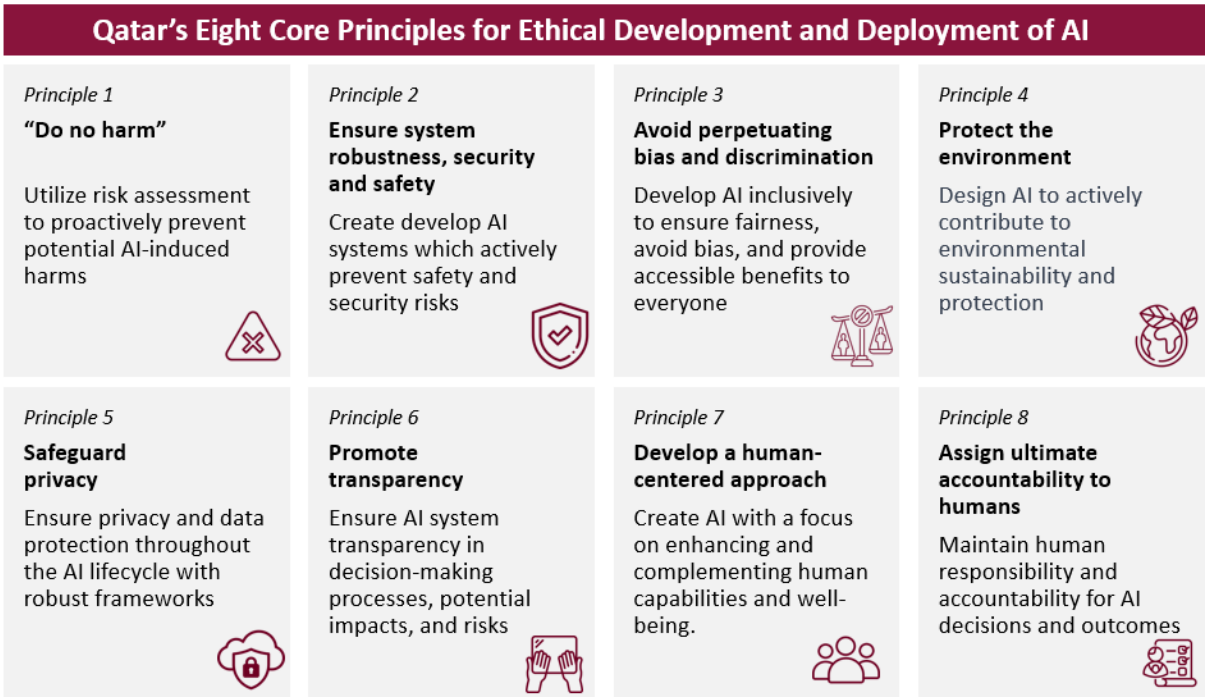
Fostering an ethical approach to the development of AI in Qatar is essential to ensure that the deployment of new technologies aligns with local values, respects cultural norms, and contributes positively to societal and economic development.

The Ministry of Communications and Information Technology is optimistic about the potential benefits of AI and trusts that AI can help Qatar meet future challenges and fuel the nation's economic growth.

However, to sustain Qatar's success in this area, basic AI principles and guidelines for its ethical development must be agreed upon, rooted in the local context, and aligned with international norms.

[1] AI system refers to the physical or virtual products or services that use AI to serve end users

Principles for ethical development and deployment of AI are:



These principles were developed to align with Qatar National Vision 2030 and to reflect the unique cultural, religious, and social considerations of AI in Qatar:

Human development	Social development	Economic development	Environmental development
– **Develop a human-centered approach** *(Principle 7)* – **Ensure system robustness, security, and safety** *(Principle 2)* – **Safeguard privacy** *(Principle 5)*	– **"Do no harm"** *(Principle 1)* – **Avoid perpetuating bias & discrimination** *(Principle 3)*	– **Promote transparency** *(Principle 6)* – **Assign ultimate accountability to humans** *(Principle 8)*	– **Protect the environment** *(Principle 4)*

This document reflects current best practices and is aligned with global standards set by organizations such as UNESCO[2] and the OECD[3].

Additionally, the 'Guidelines for Secure Adoption and Usage of Artificial Intelligence' by the National Cyber Security Agency (NCSA) are referenced where applicable.

[2] United Nations Educational, Scientific and Cultural Organization (UNESCO)
[3] The Organization for Economic Co-operation and Development (OECD)

The guidelines for ethical development and deployment of AI

Principle 1—"Do no harm"

Guidelines

Before deploying any AI system, comprehensive evaluations should be conducted to identify its potential risks, including those that could result from accidental and/or malicious use of the system.

Risk assessment procedures should be continuously updated and include measures to prevent potential harm to individuals.

To ensure proper risk evaluation and assessment at the initiation stage, the following focus areas issued by the NCSA[4] can be considered:

- People: Addressing appropriate AI governance[5], oversight, and potential impacts on human resources and safety
- Processes: Implementing necessary procedures, such as a risk management framework, for AI users and deployers within organizations

For more details, refer to Guidelines for Secure Adoption and Usage of Artificial Intelligence by NCSA.

Rationale

To benefit society and the environment with AI systems, those who develop and deploy AI need to consider how each phase of the AI life cycle could contribute to achieving these goals.

This aligns with the central aspiration of the Qatar National Vision 2030 to carefully balance Qatar's economic growth with larger human, social, and environmental goals, including the welfare of Qatari citizens, expatriate workers, and future generations.

Ultimately, AI systems ought to contribute to economic prosperity, minimize the risk of harm to individuals and society, and not violate ethical standards.

Example

Qatar has prepared its road infrastructure to accommodate autonomous vehicles, aiming to leverage its advanced road network for improved transportation efficiency and safety. Ensuring the responsible development and implementation of autonomous vehicles (such as conducting rigorous pilot programs) is crucial to minimize risks, especially for pedestrians' safety. Unforeseen accidents could occur if AI systems controlling these vehicles are not rigorously trained and tested on diverse road conditions (i.e., deserts) and scenarios.

[4] National Cyber Security Agency of Qatar
[5] Governance and monitoring of AI to ensure ethical and legal compliance

Principle 2—Ensure system robustness, security, and safety

Guidelines

Robustness

AI systems should be robust to perform effectively and consistently across various conditions and scenarios, with the ability to handle different inputs, uncertainties, and unexpected situations without significant degradation in performance.

To ensure the robustness of AI systems, thorough verification and validation testing should be undertaken.

This includes testing the AI system under different conditions to identify vulnerabilities and weaknesses, implementing strategies to handle insufficient data (quality or quantity), and evaluating the AI systems' resilience to intentional attempts to deceive or manipulate it.

Security

AI developers/deployers should ensure AI systems and their data are secured against unauthorized access, attacks, and other potential threats.

This includes safeguarding against intentional malicious actions and unintentional vulnerabilities that could compromise the integrity and confidentiality of the system.

To guarantee the security of AI systems and their data, AI developers/deployers should implement strong authentication and authorization mechanisms and regularly update and audit software to address any security vulnerabilities.

To minimize the potential consequences of data breaches, employ encryption to safeguard data used for training the model while transferring and storing, and anonymize any sensitive details before feeding them into the training process.

The technology focus area in the NCSA's Guidelines for Secure Adoption and Usage of Artificial Intelligences defines the various security standards that need to be considered when using and deploying AI systems to ensure that information security standards are met.

Safety

AI systems need to be safe against potential threats relating to failure of performance and security concerns.

AI developers/deployers should deploy continuous monitoring tools to track the AI system's performance in real-time, enabling rapid detection and correction of any deviations from expected behavior.

They should also implement multifactor authentication and role-based access control to restrict access to authorized personnel, regularly conduct security audits and update software to mitigate vulnerabilities and use strong encryption for data in transit and at rest.

Rationale

Qatar aspires to *"become the leading regional AI hub, further advancing Qatar's economy by 2030"* and eventually become an exporter of AI systems/ technologies.

To do so, AI systems being developed in Qatar need to be robust, safe, and secure to ensure acceptance from other stakeholders.

Additionally, to align with Qatar's National Cybersecurity strategy and its vision to *"Establish and maintain a secure cyberspace to safeguard national interests and preserve the fundamental rights*

and values of Qatar's society," AI systems must be safe and secure when being developed and deployed across sectors.

AI developers and deployers are responsible for fortifying the systems and solutions they create against potential cyber threats.

This aligns with national cybersecurity objectives and underscores the commitment to safeguarding critical infrastructure and upholding the security of AI applications.

Example

Microsoft's chatbot Tay was designed to learn from online interactions and mimic a typical teenager. The chatbot started generating racist and offensive remarks within just 24 hours due to the lack of security measures in place to prevent the corruption of the AI's training data. If similar AI chatbots are implemented in Qatar without proper security measures, they could potentially generate language or content deemed offensive to specific ethnic groups or religious beliefs, which would violate Qatar's cultural and religious values.

Principle 3—Avoid perpetuating bias and discrimination

Guidelines

To ensure the AI systems developed and deployed in Qatar are fair and non-discriminatory, those who develop or deploy AI should ensure that:

1. The data ingested by their AI system is representative of the Qatari population, and no groups are underrepresented.
2. Groups that could be adversely impacted by inaccurate training data[6] (including stereotypes and bias) are identified, and conscious steps are taken to rectify it.
3. Discrimination impact assessments are conducted to ensure the fairness of AI systems.
4. AI systems don't contain algorithmic or historical biases that favor particular groups.
5. Regular audits are undertaken to monitor AI systems for biases that may arise over time, with a feedback mechanism for users to report instances of perceived bias and documentation on how biases are being addressed within the AI system.

Rationale

Qatar's multicultural society, comprised of individuals from various ethnicities, backgrounds, and nationalities, drives the need for non-discriminatory AI systems:

As part of its National Vision 2030, Qatar places significant importance on providing equal opportunities to its citizens. Furthermore, the Constitution of 2004 explicitly mandates the avoidance of any form of discrimination based on gender, race, language, or religion.

With Qatar's ambition of becoming a regional hub and a global exporter of AI systems and solutions, this necessitates a focus on eradicating pre-existing biases from AI models.

AI developers play a pivotal role in this endeavor, ensuring their systems are designed and trained to promote unbiased decision-making. By fostering fairness and inclusivity, Qatar's AI ecosystem can meet the needs of its diverse population and gain acceptance and adoption globally, further solidifying its position as a regional and international AI leader.

[6] Data that is used to train the AI systems

Example

Amazon's AI-powered recruitment system, initially intended to assess applicants' suitability for specific roles by learning from past applicants' resumes, developed a bias against women. This bias persisted as the AI mistakenly interpreted the predominance of male applicants in technical roles as a preference, resulting in female applicants being unfairly penalized as they were given lower ratings. This potential scenario poses significant risks in Qatar, as it violates the country's labor laws and undermines the objectives outlined in the Qatar National Vision 2030, which aims to prioritize women's rights to education, employment, and social advancement.

Principle 4—Protect the environment

Guidelines

Those involved in developing and deploying AI systems should try to understand their systems' sustainability and environmental impact.

Examples include the energy consumption associated with training an AI system, carbon emissions, and other environmental impacts of deploying them on cloud infrastructure.

AI developers/deployers should prioritize and use energy-efficient algorithms and models that reduce the overall energy consumption of AI systems as far as possible.

AI developers should also prioritize using recycled or repurposed graphics processing unit (GPU) components whenever possible, reducing their environmental impact.

The potentially positive impacts of AI systems on social development should be weighed against potentially negative effects on the environment in Qatar and globally—and AI systems should only be developed where the net positive outweighs the net negative.

Rationale

The fourth pillar of the Qatar National Vision 2030, Environmental Development, calls for a careful balance between the socio-technical advantages of new technologies vs. the need for environmental preservation.

Protecting the environment and supporting sustainable development are at Qatar's priorities. Notably, it was one of the first countries to ratify the United Nations Framework Convention on Climate Change (UNFCCC) in 1996.

As such, developing and implementing AI systems that prioritize environmental protection is essential for fostering sustainability.

AI technologies, particularly those involving large-scale computations, contribute significantly to carbon emissions, mainly through the consumption of energy:

The rise in global AI initiatives has led to an increased demand for manufacturing modern graphics processing units (GPUs). Transforming raw materials into usable components inevitably generates emissions and waste throughout the production chain, leaving a substantial environmental footprint.

As Qatar is committed to sustainable practices and environmental conservation, implementing AI guidelines focused on environmental protection becomes essential.

Example

Batelco, a major telecommunication provider in Bahrain, has adopted an environmental approach to powering its data centers by utilizing solar energy and optimizing energy usage, ultimately resulting in a reduced environmental footprint associated with energy consumption. Developing AI systems utilizing data centers reliant on renewable energy supports Qatar's environmental initiatives.

Principle 5—Safeguard privacy

Guidelines

To safeguard privacy and the unauthorized access to or misuse of sensitive information, organizations developing and deploying AI in Qatar should establish protocols for accessing data and assign data protection officers to identify who can access what data.

Organizations should implement privacy and data protection measures throughout the AI system's life cycle and safeguard user-generated data.

Adherence to applicable Qatari Government policies, privacy law, and relevant international standards and best practices is essential.

In addition, organizations can protect the privacy of the personal data being used in AI systems by:

1. Using training models that do not rely, or rely minimally, on user input.
2. Ensuring their AI systems are not using personal data for unlawful or discriminatory purposes.
3. Ensuring the personal data of individuals—such as age, gender, and political views—is not made public.
4. Enabling users to flag issues related to privacy or data protection.

Rationale

Qatar puts a high emphasis on data protection and privacy through various policies and laws, incl. but not limited to the Personal Data and Privacy Protection Law – which ensures that individuals have the right to control their personal information and that entities collecting and processing data must do so lawfully.

This is particularly relevant to organizations that wish to use AI, as the technology can potentially increase the amount of personal information collected exponentially.

Failure to safeguard privacy erodes public trust in AI and may have legal repercussions, and fines may be associated with data breaches.

Example

Intelligence Design, a company based in Japan, developed AI surveillance cameras to monitor millions of visitors in Shibuya, analyzing only necessary information while discarding any sensitive personal information, thus ensuring the safety and security of the personal information that the AI processes. Qatari developers could implement a similar AI system to guarantee the safety of Qatar's inhabitants by anonymizing the information it processes in line with the country's data protection regulations.

Principle 6—Promote transparency

Guidelines

To ensure transparency in the development and deployment of AI systems, those who develop or deploy AI in Qatar should ensure that:

1. Information should be provided to users of AI systems about how these systems work, including the data ingested, algorithms employed, and features driving the outputs.
2. The processes used by AI systems to reach conclusions must be traceable—users should be able to determine the factors that give rise to AI-generated outputs and AI-influenced decision-making.
3. While complying with national privacy and data laws, those who develop and/or deploy AI systems should be as transparent as possible about the origins of the data used in their systems.
4. In cases where AI systems have been used to help humans make significant decisions, users should be made aware of how AI has been utilized.
5. The governance structures and policies in place for overseeing the development and deployment of AI systems should be disclosed. This includes roles, responsibilities, decision-making processes, and how compliance with ethical principles and regulatory requirements is ensured.

Rationale

Transparency in AI systems holds particular significance for Qatar, as the nation aims to position itself as an exporter of AI: as such, prioritizing AI solutions with “transparency by design” can position the country as an ethical and reliable player in the regional & global AI market.

Furthermore, transparency in AI is essential for fostering accountability, ethical practices, and user trust: it makes it easier for developers to address errors, promote responsible practices, and increase user acceptance of AI.

Finally, transparency aids compliance with regulations, encourages collaboration, and helps to identify and mitigate preconceived and stereotypical ideas and biases.

Example

In 2023, Google faced a USD 1.6 billion lawsuit arising from concerns regarding the use of personal information from social media and Google platforms to train their AI models, stemming from Google's lack of transparency in the development of AI systems. If such systems were to be implemented in Qatar, developers should implement measures to ensure transparency in developing such models, including disclosing the use of training data and adopting ethical practices that align with the country's legislative frameworks.

This growing concern is reflected in countries like the U.S., where new regulations are being implemented to require software vendors to provide transparency in how they train, develop, and test their AI tools, aiming to make AI decision-making and reasoning more transparent.

Principle 7—Develop a human-centered approach

Guidelines

To promote human-centered AI in Qatar that is user-friendly, ethical, and inclusive, those who develop and/or deploy AI should:

1. Adhere to Qatari cultural and religious values and refrain from any form of disrespect or attempt to alter established cultural norms.
2. Put in place feedback mechanisms to receive input from the public/users on the AI systems.
3. Promote diversity in development teams to ensure a broad range of perspectives and experiences.
4. Consider the target user population's cultural, linguistic, and demographic diversity.
5. Consider accessibility throughout the development of AI models, ensuring they are usable by everyone, including people with disabilities.

Rationale

Human Development is the first pillar of the Qatar National Vision 2030 and signifies the importance of fostering the citizens' well-being, capabilities, and opportunities.

Adopting a human-centered approach in developing AI systems ensures that AI systems address specific challenges faced by citizens and residents of Qatar and that the systems are designed in a user-friendly manner.

Simultaneously, as Qatar aims to become a regional AI hub and AI exporter, tailoring AI solutions to Qatar’s diverse cultural, social, and economic contexts can enhance user adoption and acceptance.

Qatar boasts a high e-accessibility rate of 91%. Committed to ensuring inclusivity for all, Qatar invests in standard and assistive technology across all societal segments, including individuals with disabilities.

Example

When prompted with terms containing religious keywords, various generative AI systems have generated stereotypical and offensive results, leading to public outrage. The development of AI systems in Qatar should carefully consider such possibilities to ensure adherence to Qatari cultural and religious values.

Principle 8—Assign ultimate accountability to humans

Guidelines

The ultimate responsibility of AI systems and its outcome lies with humans. Therefore, those who develop or deploy AI should ensure that:

1. Developers, designers, and deployers share accountability for the outcomes of AI systems. Each party must understand and acknowledge its role in the system's development and deployment, fostering a collective commitment to responsible practices.
2. When AI systems produce specific outcomes that impact individuals, it is imperative to communicate this information to users.
3. AI systems should not be able to autonomously make decisions of significant consequence. Instead, these systems should allow for human intervention and provide users with the ability to appeal or override decisions that have a substantial impact on individuals or society- ensuring that critical judgments remain within the realm of human control.
4. Experts in the domain where the AI system will be deployed should collaborate on the development process of AI systems.
5. AI systems that inform significant decisions should be subject to external audits to validate the AI system's fairness, transparency, and compliance with established standards and regulations.

Rationale

The second pillar of the Qatar National Visio 2030, Social development, highlights the significance of high moral standards in Qatar's economy.

Given Qatar's commitment to high moral standards, ensuring that AI systems being developed and deployed translate these values is imperative. Therefore, assigning responsibility to humans becomes pivotal in ensuring that ethical considerations, cultural values, and moral principles are diligently incorporated.

Example

The US Senate introduced the No Section 230 Immunity for AI Act, which includes a provision removing immunity from AI companies in violations concerning the use or provision of Generative AI[7]. The legislation aimed to ensure that developers who create AI systems are accountable for ensuring their compliance with legal and ethical standards. Developers of AI systems should consider the potential consequences of such systems on Qatar and the country's inhabitants, with developers being held accountable for the systems they develop.

[7] Generative AI is artificial intelligence capable of generating text, images, or other data using generative models, often in response to prompts

Summary

AI as a technology has the potential to foster economic growth and social progress by increasing efficiency, enhancing decision-making, and potentially facilitating breakthroughs in every aspect of human endeavor.

On the other hand, its widespread adoption could result in job displacement, entrenchment of prejudice and bias, and security vulnerabilities.

The principles outlined in these guidelines are designed to help manage these risks while increasing the benefits for Qatari society—particularly by harnessing AI's potential to promote human, social, economic, and environmental development.

To do so, it is critical that:

- AI systems involved in critical decision-making should be regularly audited and subject to quality assurance measures.
- Users have the right to question significant automated decisions that affect them and, when applicable, be given the option to challenge such decisions.
- Developers and those that deploy AI in Qatar consider—and respond thoughtfully—to the principles outlined above and incorporate them into their AI systems and business approaches.