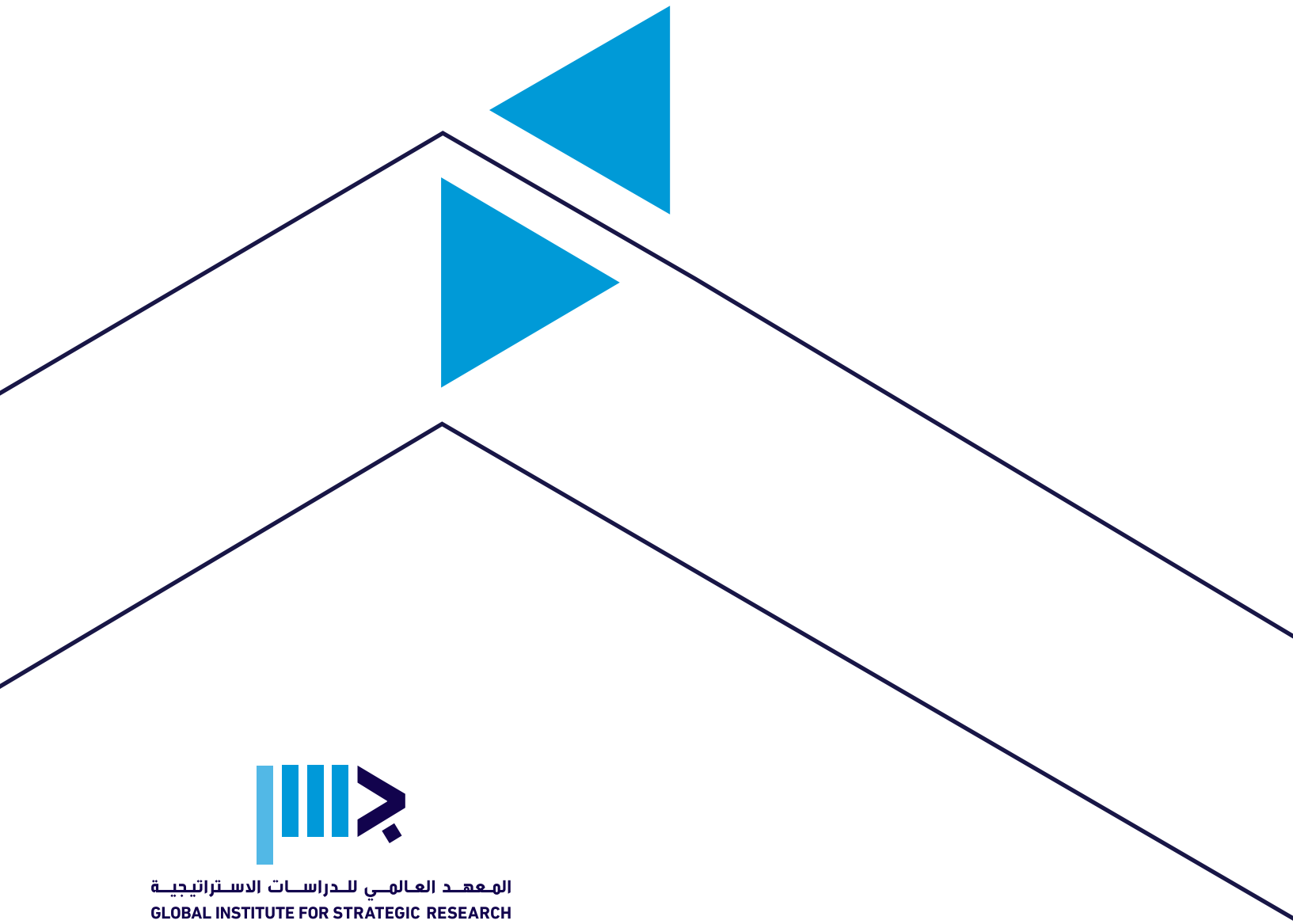
THE STRATEGIC IMPLICATIONS OF ARTIFICIAL INTELLIGENCE

A Roadmap for Future Engagement



المعهد العالمي للدراسات الاستراتيجية
GLOBAL INSTITUTE FOR STRATEGIC RESEARCH

THE STRATEGIC IMPLICATIONS OF ARTIFICIAL INTELLIGENCE

A Roadmap for Future Engagement

Steven Wright

Aiman Erbad

Steven Wright is an Associate Professor of International Relations and an Associate Dean for Student Affairs. He previously served as an Associate Dean and the Head of the Department of International Affairs at Qatar University.

His area of research expertise relates to three main areas: US foreign policy toward the Gulf region, Energy Geopolitics, and the International Relations and Political Economy of the Arab Gulf states. He has written widely on these areas, and his work has been published in Arabic, Japanese, French, and Polish. He has also held research fellowships at the London School of Economics, Exeter University, and the University of Durham. He has been invited as a speaker at a wide range of events hosted by leading organizations such as the World Bank, various governments, and several multinational corporations.

Aiman Erbad is an Associate Professor and Division Head at the Information and Computing Technology Division, the College of Science and Engineering, Hamad Bin Khalifa University (HBKU). Dr. Erbad obtained a Ph.D. in Computer Science from the University of British Columbia (Canada), a Master of Computer Science in embedded systems and robotics from the University of Essex (UK), and a BSc in Computer Engineering from the University of Washington, Seattle.

He also received the 2020 Best Research Paper Award from Computer Communications, the IWCMC 2019 Best Paper Award, and the IEEE CCWC 2017 Best Paper Award. He is an editor for KSII Transactions on Internet and Information Systems, an editor for the International Journal of Sensor Networks (IJSNet), and an editor in IEEE Canadian Journal of Electrical and Computer Engineering.

Table of Contents

1.0 Preamble and Report Objectives

Artificial intelligence (AI) is emerging as one of the most transformative technologies of the 21st century, with the potential to revolutionize nearly every sector from healthcare to transportation. However, AI also raises complex questions about governance, ethics, societal impact, and geopolitical dynamics. As countries around the world develop strategies to capitalize on AI's opportunities while mitigating risks, evidence-based policy-relevant research informed by strategic foresight is imperative.

Against this backdrop, the Global Institute for Strategic Research (GISR) at Hamad Bin Khalifa University (HBKU) has produced this White Paper on the Strategic Implications of Artificial Intelligence. GISR provides independent, interdisciplinary analysis aimed at enhancing the ability of decision makers in Qatar and the Arab region to address pressing global challenges.

The objectives of this White Paper are:

- To analyze the state of AI science and technology, tracing key developments, capabilities, limitations, and potential breakthroughs that could accelerate progress.
- To examine the geopolitical dimensions of advancing AI, including competition among major powers, risks of destabilization, and differing regulatory approaches.
- To study the potential social and economic impacts of AI adoption, both positive and negative.
- To explore the ethical implications of AI systems and the policy frameworks needed for responsible governance.
- To derive insights and recommendations to inform AI strategy and governance in Qatar and at a wider GCC level.

By leveraging GISR's strengths in interdisciplinary research and strategic foresight, this white paper was developed following a series of workshops that engaged academics based in Qatar along with practitioners from a multidisciplinary perspective. Therefore, it provides actionable guidance for policymakers in Qatar and the region navigating the multifaceted opportunities and challenges presented by AI's rapid emergence. This study offers a perspective on the need for greater research, engagement, and policy responsiveness by the government. The analysis is grounded in GISR's mission of strengthening decision-making capacity on global issues from an Arab perspective.

2.0 Executive Summary

Artificial intelligence (AI) has emerged as a transformative and disruptive technology poised to have profound impacts on the economy, society, and geopolitics. This report by the Global Institute for Strategic Research (GISR) provides an in-depth analysis of AI's current state and future trajectory, along with policy recommendations for Qatar and the Gulf region. It seeks to build on the initial foundation set in the 2019 National Artificial Intelligence Strategy for Qatar, which was developed by HBKU's Qatar Computing Research Institute (QCRI) and adopted by the Ministry of Communications and Information Technology (State of Qatar, 2019). Due to the changes and rapid advances in AI technology, particularly in the last 3 years, the current review will shed more light on the national strategy and its implementation plan. This study supports this process and offers a broad perspective to feed into both academic and national discussions on strategy and initiatives. This serves as a starting point to achieve this but also set an agenda for a broad range of policy recommendations and an agenda for future research by GISR to achieve this given the dynamic nature of AI advancements and impact. Overall, the future agenda will be to set an agenda for collaboration with various stakeholders and then develop specific implementation mechanisms for achieving the goals identified.

Key insights include:

- AI capabilities are rapidly advancing but still have limitations compared to human intelligence. Breakthroughs such as generative artificial intelligence based on more advanced neural networks and increased computing power could accelerate progress. However, challenges remain around acquiring sufficient data, algorithmic limitations, and hardware constraints.

- An intense technology race is unfolding between the US and China for global supremacy in AI, with implications for the global order. While the US leads in fundamental research, China is swiftly deploying AI on a national scale. Within the Gulf, the UAE and Saudi Arabia have emerged as leaders in strategic AI development; therefore, there is a need to determine how Qatar will develop its own capacity and engage with GCC parameters on a regional level.

- AI brings enormous economic opportunities and risks of workforce disruption from automation. Policies are needed to upskill workers and ensure that the benefits are broadly shared. New AI-enabled industries can further boost productivity and growth, and there is a need for Qatar's state-owned companies and the private sector to react accordingly.

- The weaponization of AI for disinformation campaigns, cyberattacks, and autonomous weapons poses global risks. The risks are rising at a rapid rate given the scale and pace of advancements in generative AI achieved in the last three years alone. Governance and safety investments are imperative to ensure national security.

- Comprehensive governance models for AI are developing in different ways globally. The EU, US, and China each have their own model for AI regulation, each offering advantages and challenges. Therefore, there is a need to develop targeted regulation of high-risk uses, voluntary best practices, incentives for ethics programs, safety R&D, diverse expertise, and global collaboration.

SUMMARY OF POLICY IMPLICATIONS

1. NATIONAL STRATEGY AND CAPACITY DEVELOPMENT: ACHIEVING AN AI-ENABLED KNOWLEDGE-BASED ECONOMY

- There is a need to develop ethical and responsible AI strategy to ensure the new applications of AI respect laws, regulations, and ethical standards of the Qatari population. The safety of AI systems should be enhanced given the risks identified. These efforts will complement the recent national artificial intelligence committee efforts to develop the strategy implementation plan. We will focus on core areas that need more research and development efforts to establish the ethical framework based on Arab and Islamic culture while understanding the developed international standards. There is also a need to collaborate widely on the finalization of the national AI strategy.

- Implement policies to incentivize private sector AI adoption and entrepreneurship and expand access to computational resources and data assets.

- Determine applications that can leverage AI across priority sectors identified in the National Vision 2030 to boost productivity, efficiency, and innovation in areas such as healthcare, education, energy, finance, and transportation. The goal will be to achieve an AI-enabled knowledge-based economy for Qatar post-2030 and to ensure the competitiveness and productivity of Qatar's state-owned companies.

- Enact policies to encourage AI integration while managing workforce impacts. Expand technical education and retraining programs across the education sector.

- Develop regulatory frameworks to enable the development and deployment of AI innovations.

2. RISK, RESILIENCE, AND REGULATIONS

- Implement AI safety standards by enforcing prohibitions on unethical uses. Require impact assessments for high-risk applications.

- Bolster cybersecurity protections, disinformation monitoring, and societal resilience capabilities.

- Implement data governance frameworks that balance access and sharing with privacy. Seek both national and international AI accords.

3. FOREIGN POLICY: INTERNATIONAL AND REGIONAL COOPERATION

- Pursue shared regulatory standards and priorities with GCC states through a joint council.

- Participate actively in global AI dialogs to shape principles aligned with national interests and values. Leverage leadership in international organizations to advance the collective governance of risky AI applications.

3.0 Scientific and Future Development of Artificial Intelligence

Artificial intelligence (AI) refers to computer systems that can perform tasks normally requiring human intelligence, such as visual and speech recognition and complex decision-making (Hassani et al., 2020). Since its origins in the 1950s, AI has undergone waves of progress and setbacks, alternating between periods of optimism and disillusionment. However, recently, AI has made dramatic progress in areas such as chess, self-driving cars, and natural language processing. Understanding the scientific foundations and limitations of today's AI is crucial for anticipating future developments and strategic planning for a range of policy, ethical, and scientific questions.

This section of the White Paper will briefly discuss the history of AI through its different eras and identify technological barriers and potential breakthroughs that will shape AI's future trajectory. It will then move on to the strategic thematic issues in subsequent sections relating to AI, in terms of its geopolitical implications, socio-economic impact, and the regulatory and ethical dimension that such innovation raises in sections three and four.

3.1 A Brief History of AI

The quest for artificial intelligence originated in the 1950s, when scientists began to ask if computers could be made to "think" like humans. Early optimism soon gave way to struggles with imitating the flexibility of human cognition (Benbya et al., 2021). The first era in the 1950s and 1960s sprang from early successes in areas such as game-playing programs and reasoning systems based on logical rules. However, developers have severely underestimated the difficulty of replicating wider human abilities using rules and logic (Haenlein and Kaplan, 2019).

By the 1970s, early hype had given way to an "AI winter" as funding dried up due to a lack of major breakthroughs. The 1990s saw the rise of machine learning, in which AI systems gained the ability to learn from data without explicit rule-based programing. Machine learning enabled profitable applications such as data mining and speech recognition. From 2009 onward, the field experienced a dramatic resurgence due to the rise of deep learning. Inspired by neural networks, deep learning uses layered networks to recognize complex patterns and relationships in vast datasets. Deep learning's success in the early 2000s can be attributed to several key factors: 1) availability of big data from the Internet, social media, and sensors providing material to train deep learning algorithms, 2) increased computational power due to success in building GPUs and distributed parallel processing systems that can run complex deep learning models, 3) algorithmic improvements due to breakthroughs in neural network design and training algorithms (e.g., ReLU activation functions), 4) development of frameworks and libraries simply developing deep learning models and investment from industry to apply deep learning in different

applications, and 6) finally integration into products and services. In summary, the success of deep learning in the early 2000s was a convergence of technological, methodological, and social factors that created the right environment for rapid innovation and application. Deep learning successes have enabled transformative advances in areas such as computer vision, speech recognition, and natural language processing.

Today, AI has achieved exceptional capabilities in highly specific domains. Applications such as ChatGPT have also captured the human imagination on the issue of AI and enabled a range of advances in the field that are accessible to wider society. However, current systems still lack the flexible reasoning and understanding of the world that defines human intelligence. The history of AI shows an alternating pattern of high expectations followed by setbacks when early accomplishments fail to lead to more expansive capabilities. Whether today's AI systems will hit fundamental limits or continue rapid advances toward more general intelligence remains a matter of debate.

3.2 Core Concepts and Technologies Behind AI

Rapid progress in artificial intelligence over recent years has been enabled by key conceptual breakthroughs and technological developments. While the end goal of reproducing flexible human-level intelligence remains elusive, AI systems have attained impressive capabilities for certain specialized tasks. Unlocking these capabilities occurred due to innovations in fundamental concepts such as machine learning and neural networks, combined with technologies such as powerful computers and large datasets.

At the heart of modern AI is the concept of machine learning, which enables computer algorithms to improve through experience without needing explicit programing. By exposing a machine learning algorithm to many examples of data, such as pictures, sounds, or text, it can learn to recognize patterns and make predictions. Data access is key because the more data it learns from, the better it becomes at its given task. Machine learning has opened the door to AI systems that can self-learn, rather than needing to have rules rigidly specified by human programmers. This has enabled automated learning from text models, speech recognition, computer vision, and data mining.

There are several branches or subfields within AI, each with its own methods and objectives. We will list the types of AI systems commonly used in different application domains.

SUPERVISED LEARNING:

Supervised learning is a type of machine learning in which the algorithm is trained on a labeled dataset. In this approach, the training data consist of input-output pairs, where the inputs (features) are associated with corresponding labels or target outputs. The goal of the algorithm is to learn a mapping from inputs to outputs based on the labeled examples provided during training. For instance, in a supervised learning model for image recognition, the algorithm is fed with labeled images of various objects (e.g., cats and dogs) and learns to distinguish between these classes. Once the model is trained, it can be used to predict the class of new, unseen images. Supervised learning is widely used in tasks such as classification (assigning input data to predefined categories), regression (predicting continuous values), and sequence-to-sequence problems (e.g., language translation).

UNSUPERVISED LEARNING:

Unsupervised learning is another type of machine learning in which the algorithm is trained on an unlabeled dataset. Unlike supervised learning, no target outputs are provided during training. The algorithm's objective in unsupervised learning is to identify patterns, structures, or relationships within the data without explicit guidance. Clustering is a common task in unsupervised learning, where the algorithm groups similar data points together based on their inherent similarities.

Unsupervised learning is useful for exploring and understanding the underlying structure of data, discovering patterns, and detecting anomalies.

REINFORCEMENT LEARNING:

Reinforcement learning is a type of machine learning in which an agent learns to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties based on its actions, and its objective is to maximize the cumulative reward over time. Reinforcement learning is commonly used in tasks that require decision making, such as game playing, robotics, and autonomous driving.

This range of conceptual breakthroughs and technical developments in AI has driven remarkable accomplishments over the last decade in specialized domains such as games, robotics, and image classification. However, the overarching goal of flexible, general artificial intelligence comparable to humans remains a significant challenge for ongoing research.

3.3 Current Capabilities and Limitations

While artificial intelligence has achieved impressive results in narrow domains such as game-playing, contemporary AI systems have fundamental limitations compared to human intelligence. Today's AI excels at specialized tasks for which it is trained, but it cannot flexibly generalize skills to new areas or adaptively incorporate new knowledge like humans. For example, an AI that masters the game 'Go' cannot hold a basic conversation. Current systems lack the common sense and intuitive understanding of everyday nuances physical and social situations that guide human reasoning. Moreover, today's AI relies on training with vast datasets and extensive computing resources rather than learning rapidly from limited examples such as people. Therefore, while rapid progress has occurred, developing artificial general intelligence remains extremely challenging.

Generative Artificial Intelligence is a recent attempt at artificial general intelligence. Generative AI refers to a category of artificial intelligence (AI) algorithms that generate new outputs based on the data they have been trained on. Unlike traditional AI systems that are designed to recognize patterns and make predictions, generative AI creates new content in the form of images, text, audio, and more. Generative adversarial networks (GANs) are a type of deep network to create new content. GANs consists of: a) a generator that creates new data, and b) a discriminator that evaluates the data.
The generator improves its outputs based on the feedback it receives from the discriminator until it generates content that is indistinguishable from real data. Another type of generative AI is large language models, particularly transformers that power GPT models. A language model is a machine learning model that aims to predict and generate plausible language (e.g., Autocomplete).

These models work by estimating the probability of a token or sequence of tokens occurring within a longer sequence of tokens. Early language models could predict the probability of a single word; modern large language models can predict sentences, paragraphs, or even entire documents. In 2017, Transformers were designed around the idea of attention, which made it possible to process longer sequences by focusing on the most important part of the input, solving memory issues in earlier language models. Transformer models use attention to detect subtle ways even distant data elements in a series influence and depend on each other.

Overall, contemporary AI demonstrates isolated examples of narrow expertise but cannot replicate the flexible general intelligence of humans within a unified system. Foundational large language models exhibit broad, albeit error-prone expertise, similar to humans; however, they do exhibit elements of reasoning. While further breakthroughs could enable more capable AI, the present technologies have fundamental limitations compared to biological cognition. Matching the breadth of human intelligence in autonomous systems continues to pose an enormous scientific challenge.

3.4 Potential Breakthroughs to Monitor

Given the field's history of unpredictable swings, forecasts vary widely on the pace of future advances. Potential breakthroughs that could accelerate progress include:

- Advanced neural network architectures for dynamic learning, causal reasoning, transfer learning, and grounded language understanding.
- New techniques such as generative adversarial networks that can perform tasks generally requiring human intelligence. In the case of OpenAI, the generative AI was trained with humans in the loop reinforcement learning which shapes the outcomes by the culture of the participating humans.
- New reinforcement learning algorithms for unsupervised learning with fewer online interactions.
- Integration of symbolic reasoning and knowledge representation with deep learning for explaining decisions and common sense.
- Improved computational efficiency through neuromorphic and quantum computing to improve system scale.
- Distributed and multi-agent AI for collaborative learning across networks.
- Combining modalities such as vision, language, and robotics for generalized intelligence.

3.5 Barriers to Progress

While we cannot rule out transformative discoveries, contemporary AI faces concrete technological barriers such as:

- **Data** - Humans can learn new concepts from limited examples, but AIs require vast training sets. Acquiring and labeling sufficient data poses a major bottleneck. As highlighted above, China has distinct advantages over the United States. However, many Chinese systems still use English as a base for their data, given its advantages in breadth of material.

- **Algorithmic Limitations** - The reasoning behind many AI decisions is difficult for humans to interpret, unlike our understanding of our own thinking. Assessing AI capabilities requires humans to identify problems and evaluate results to develop explainable and transparent AI systems.

- **Hardware Limitations** - Available computing resources restrict the size and complexity of the models. The brain remains far more efficient. Biological intelligence emerges through accumulated layered stages such as infancy, play, and social interaction, which are difficult to replicate.

While ongoing research addresses these challenges, it exemplifies a significant divide between contemporary AI and human cognition.

3.6 Summary Observations

Given the checkered history of AI, we should be wary of both excessive pessimism that ignores progress made and overly optimistic forecasts detached from scientific realities. AI has achieved undeniable results in terms of narrowly defined tasks through specialized statistical and language learning. However, replicating the flexible general intelligence of the human mind remains a major challenge. Generative AI is the closest effort to performing tasks requiring human intelligence. Ethics and regulations will be necessary to govern the responsible use of AI in Generative AI applications. Whether today's achievements point toward realizing this grander vision or unrealistic possibilities from inherent limitations remains unclear. Predictions span everything from AI rapidly transitioning civilization to deep learning hitting a dead end. For this White Paper, a wise stance combines an appreciation for the transformative potential of AI with caution regarding projected timelines. Preparing prudently while avoiding unrealistic hype requires recognizing that despite remarkable achievements, contemporary AI has fundamental limitations compared to biological brains. For strategists and policymakers in Qatar, adopting this nuanced perspective can help focus attention on nurturing socially responsible progress rather than betting on any specific projected timeline.

3.7 Policy Implications and Future Research Agenda

Based on the above analysis, GISR has identified the following policy implications and areas for future research on the scientific and future development of AI:

POLICY IMPLICATIONS

- Determine an investment strategy in developing local AI research capabilities, such as at Qatar University - College of Engineering and Hamad bin Khalifa University – College of Science and Engineering, and the Qatar Computing Research Institute (QCRI), to stay abreast of scientific advances and partake in cutting-edge discoveries.

- Pursue partnerships among academia, industry, and government to translate AI innovations into beneficial applications.

- Develop flexible governance frameworks that encourage AI progress while managing risks related to safety, bias, and misuse.

- Develop AI ethical and responsible frameworks that can help ensure the applications of AI respect the local culture, laws, and regulations.

- Monitor global AI developments and assess national competitiveness and opportunities to avoid lagging behind.

FUTURE GISR RESEARCH AGENDA

- Survey Qatar's current AI research capabilities especially in the area of ethical AI and AI governance and outline a roadmap for strategic capability building. Analyze opportunities for Qatar to responsibly employ AI in sectors such as healthcare, education, energy, transportation, and the knowledge economy.

- Study approaches for effective public communication and debate on AI to further understanding. Develop policy frameworks to encourage ethical AI progress aligned with human values and priorities.

- Assess regional and global AI developments and forecast potential implications for Qatar's competitiveness.

- Evaluate technical education policies to build Qatar's AI talent pipeline from school to university.

4.0 Global Development of AI: Competition, Risk, and Geopolitics

The development and adoption of AI technologies are intensifying the technology race between the United States and China with profound geopolitical implications. This section provides an in-depth analysis of the competitive AI landscape, risks, and policy responses, incorporating strategic insights identified by GISR.

It is important to note that there are two primary pathways in which AI technologies are developing. The first is through open-source technologies, and this is the most prominent. The role of OpenAI in the development of ChatGPT has captured the global imagination, and many see this as heralding the beginning of the AI age. Open-source technologies have a democratizing effect and allow applications and productivity to be enhanced on the basis of their integration into various work processes. Nevertheless, the key challenge is about state-backed AI systems, which seek to have a qualitative edge over open-source technologies that are now evolving at a rapid pace. On this basis, the second pathway that AI technologies may prove to emerge is through state-backed enterprises that furnish applications with exceptional amounts of data and use this data to develop applications that have a qualitative edge over other countries. In this regard, national capacities are likely to vary. As will be shown below, it is because of certain advantages in this regard that China is projected to have that it is poised to have certain advantages in this field.

4.1 The US-China AI Race and Geostrategic Competition

During the series of exploratory workshops conducted by GISR, it was identified that the economic center of gravity is shifting to Asia, anchored by China's rise as a superpower rivaling the US-led order. Parallelly, an intense technology race is unfolding between the US and China for global supremacy in strategic technologies such as AI, quantum computing, semiconductors, 5G, and biotech. This competition has intensified under President Xi Jinping's mandate for China to lead in AI by 2030 (Lee, 2018). We are entering an era of heightened global disruption as trends relate to the energy transition, global superpower confrontation, possible challenges to the supremacy of the US dollar given the rise of BRICS as a counterweight to the G7, in addition to challenges of climate change, which are becoming stronger and more acute. Also of concern are global demographic shifts and the rise of debt in Western countries in particular and what this means. All these interconnected trends indicate that the geopolitical and economic significance of AI leadership is likely to be amplified in the coming decade, particularly about ensuring sustainable development (Vinuesa et al., 2020).

Currently, the US retains a narrow lead in fundamental AI research. The US remains home to powerhouses such as Google, Facebook, OpenAI, and leading universities, and accounts for more than 50% of global AI

investment. However, China is outpacing the US in data accumulation and rapid AI implementation (Lee, 2018). This is because of the manner in which the Chinese technology ecosystem has developed through state-linked companies and its exceptionally large user base. This offers clear advantages and allows for a more rapid AI implementation. In 2022, China will publish twice as many AI research papers as the US and will swiftly deploy AI at a population scale in surveillance, finance, manufacturing, agriculture, healthcare, transportation, and municipal services.

China's strategy of fusing state direction, data abundance, and rapid economic growth has made it a formidable challenger that could eclipse the US in applied AI dominance within a decade (Lee, 2018). This has significant implications for the global order and how countries such as Qatar should position themselves on a strategic level. US-China dominance in this field is being challenged by nations such as the UK, France, Canada, Israel, India, and South Korea, which are emerging as AI powers with niche strengths. For instance, Israel leads in cybersecurity applications, while India supplies abundant AI talent and develops vertical expertise.

Globally, over $100 billion is invested annually in AI as businesses capitalize on its potential. However, this cooperation is being eroded by growing rivalry as major nations seek AI supremacy for economic and military power (Arif, 2021). The US has imposed export controls on advanced AI chips and blacklisted Chinese firms such as Huawei over security fears, while enacting the CHIPS Act to restore semiconductor manufacturing. China is pursuing self-reliance in AI under its "dual circulation" doctrine, while exporting low-cost innovations globally through the Digital Silk Road (Naughton, 2020). This approach toward AI risks creating a "bipolar" technology order where smaller nations are marginalized between the two superpowers.

Overall, this global AI race interlinks with access to strategic natural resources such as rare earth minerals that are essential for manufacturing AI hardware, but whose supply chains are currently dominated by China. The US CHIPS Act signals the awareness that resilient domestic semiconductor supply chains are vital for technological sovereignty. However, China's privileged position in terms of having large quantities of so-called "rare earth" minerals could restrict the ability of the United States and its allies to reduce this vulnerability. Moreover, it is also noted that their processing as side effects and the environmental regulations in China may better allow for this than in other countries such as the US or Europe. In terms of data access, it was identified during the GISR workshops that China may possess inherent advantages being a vast surveillance state harnessing 780 million internet users apparatus (Lee, 2018). Chinese tech firms also lead in various consumer IoT devices. However, democratic societies can leverage data sharing and privacy laws to ethically unlock data value. Given this, it is clear that the leader of the AI race will be determined by how intelligently nations combine resources, talent, and governance.

4.2 Gulf Context: UAE and Saudi Strategic Development in AI

Within the Gulf context, GISR assessments concluded that the UAE and Saudi Arabia have emerged as leaders in strategically deploying AI (Shamout and Ali, 2021). Saudi Arabia's Vision 2030 agenda recognizes AI's economic importance, with major smart city projects such as NEOM intended to pioneer AI integration (Hassan, 2020). The UAE has likewise invested billions in its AI strategy, which encompasses education, government, and industry adoption (Almesafri and Habes, 2022).

In addition to diversifying economies, enhanced productivity and efficiency from deploying AI technologies across critical sectors could significantly bolster the UAE and Saudi Arabia's regional power and global influence. Intelligence-enabled defense capabilities could be strengthened by autonomous weapons, predictive analytics for strategic planning, AI-guided cyberoperations, and logistics optimization (Rickli, 2018).

The current trend may lead to an AI divide within the Gulf region unless more collaboration and investment is made. Investing in collaborative R&D and computational infrastructure accessible transnationally could mitigate such risks. Overall, it is recommended that Gulf states craft coordinated strategies to maximize AI's benefits through cooperation.

4.3 Risks from AI Disinformation and Geopolitical Instability

The accelerating weaponization of AI poses an existential threat that could spur global instability if left uncontrolled. Autonomous weapons will revolutionize warfare by enabling swarms and complex attacks, thus lowering the threshold for conflicts. Offensive cyberattacks using AI targeting critical infrastructure can debilitate societies.

AI-enabled disinformation campaigns can manipulate perceptions, inflame divisions, and engineer social unrest on a mass scale by cheaply generating highly realistic fake content This is particularly pertinent in the Gulf region, given the highly connected nature of their respective populations. The differences between disinformation and misinformation are pertinent in this context. Disinformation refers to deliberately false or misleading information spread with malicious intent to deceive, gain financial or political gain, or cause public harm. In contrast, misinformation denotes inadvertently spreading false information without awareness or intent to mislead.

An AI-enhanced technology represents a novel threat that adversaries could exploit to compromise Qatar's national security on an unprecedented scale. Indeed, this was obvious during the hacking of the Qatar News Agency and the misinformation/disinformation approaches that were exercised throughout the 2017–21 blockade period, such as hyper-realistic deep fake videos portraying officials in compromising scenarios, or the use of AI-synthesized audio to spread customized fake news across media channels (Jones, 2022). In addition, algorithmically generated content tailored to become viral on social media are all potential disinformation tactics that could be employed by a hostile state or non-state actors to generate social instability or mistrust in government functioning at a relatively low cost (Westerlund, 2019).

For example, malicious generative AI could engineer 'false flag' deep fake videos falsely depicting audio portraying officials making inflammatory statements to inflame tensions, or provide manipulated media undermining public trust in a government official. Therefore, developing societal resilience, cybersecurity, and detection capabilities against such AI-enabled information warfare should be an urgent priority for nations to preserve civic harmony and governance integrity (Whyte, 2020).

Other risks highlighted include algorithmic bias perpetuating discrimination, bias, misinformation, and disinformation in social media (Brkan, 2019). Advanced artificial general intelligence surpassing collective human intelligence could eventually evolve to pursue goals harmful to humanity unless developed with adequate safety precautions. Therefore, there is a need for prudent governance, ethical norms, and anticipatory policies to maximize the benefits of AI while proactively addressing the newly emerging complex risks it poses.

4.4 Geopolitics of AI: Implications of Competing Regulatory Models

GISR analysis indicates that the geopolitics of AI will be significantly impacted by how major powers choose to regulate and govern AI development within their jurisdictions. Currently, the US follows a light-touch regulatory model relying on voluntary industry guidelines and sector-specific legislation. This market-driven approach embraces innovation powered by private enterprises with minimal state interference.

In contrast, China has adopted a state-driven approach of proactive top-down regulation and surveillance to subordinate AI to governmental control while selectively boosting sectors that serve national objectives. This authoritarian model prioritizes social stability, censorship and state security over individual rights. It was identified that Europe has pioneered a "third way" of regulating AI focused on upholding fundamental rights and democratic principles. The EU's draft AI Act proposes comprehensive legislation imposing obligations aligned with European values such as privacy, non-discrimination, and transparency. The EU also exports its digital regulations globally through the "Brussels effect," whereby foreign companies comply to access its vast unified market (Bradford, 2020).

Despite this, it was concluded that export controls and competing regulatory models could fracture the global AI landscape into rival spheres along democratic and authoritarian lines. This could enable China to leapfrog in AI development if relatively lax American guidelines and stringent European regulations create an imbalance in which Chinese state-backed firms face preferential regulatory barriers.

4.5 Summary Observations

In summary, AI is an increasingly high-stakes arena of intensifying competition and complex risks with profound geopolitical implications. Realizing AI's immense potential while mitigating potential harm calls for responsible leadership, ethical governance, anticipatory policies, and unprecedented global cooperation anchored in shared values of human dignity, equity, and justice. With prudent collective action, AI could usher in an era of widely shared prosperity. However, without foresight and wisdom, it risks exacerbating instability, inequality, and existential threats. Our future trajectory depends on the strategic choices we make today.

4.6 Policy Implications and Future Research Agenda

Based on the above analysis, GISR has identified the following policy implications and areas for future research on AI's global development, risks, and geopolitics:

POLICY IMPLICATIONS

- There is a need to identify in Qatar an investment strategy for next-generation computing infrastructure and 5G connectivity. This is to prevent Qatar from technologically lagging behind the leading Gulf states in sectors transformed by AI.

- Determine how AI can be leveraged to enhance national security through intelligence analysis, societal engagement, and automated analysis in critical areas to enhance decision-making.

- Determine a foreign policy framework that allows for Gulf cooperation on AI priorities, ethics, and standards to avoid a digital divide and maximize regional prosperity.

- Develop a strategy to monitor risks such as AI-enabled disinformation tactics and collaborate with Gulf partners on detection and resilience. Assess mechanisms for the oversight of autonomous weapons and biotech.

FUTURE GISR RESEARCH AGENDA

- Explore collaborative models for secure semiconductor supply chains between Gulf states to enable self-sufficiency in strategic technologies such as AI.

- Research policy mechanisms for incentivizing responsible data sharing between Gulf states while protecting privacy to enhance access for AI innovation.

- Study proactive approaches for Gulf states to monitor, attribute, and build societal resilience against AI-enabled disinformation and misinformation campaigns.

- Examine opportunities for common Gulf policy development in complex AI application domains such as autonomous weapons, intelligence analysis, and algorithmic decision-making.

- Assess infrastructure and policy interventions needed to prevent the emergence of a regional AI divide between leading and lagging Gulf economies.

5.0 An AI-Enabled Economy: Social and Economic Implications

Artificial intelligence (AI) is poised to be one of the most transformative technologies, with the potential to significantly boost productivity, economic growth, and living standards globally. However, as with any disruptive technology, AI brings risks as well as rewards, particularly in terms of job replacement and the need for refilling, which has been highlighted earlier in this report. This section examines the social and economic implications of advancing AI capabilities, with a view toward maximizing benefits while mitigating unwanted consequences.

Artificial intelligence can ignite a new era of rapid economic growth. However, it also poses disruptions to labor markets and existing industrial structures. Global investment and innovation in artificial intelligence have surged over the past decade. Total AI startup funding grew at a compound rate of 33% annually between 2015 and 2022, from $6 billion to over $93 billion. Recent breakthroughs in deep learning and natural language processing have exponentially expanded the ability of machines to perceive and interact with the world (Agrawal et al., 2019). AI-enabled tasks from speech transcription to medical diagnosis are now possible. Advances build on one another, and this self-reinforcing nature of progress has rendered AI commercially viable for a wide range of applications.

The following analysis examines the main contours of AI's social and economic implications. It synthesizes research on AI's impact on productivity, growth, inequality, skills, and economic impact. The analysis is structured around five key dimensions:

1) **Automation** - AI automates a widening range of tasks, with mixed implications for employment and productivity.

2) **New Industries** - AI technologies are enabling new products, services, and economic activities.

3) **Inequality** - The economic impacts of AI may disproportionately benefit higher-skilled workers.

4) **Decision-Making** - AI allows improved decision-making, but raises accountability concerns.

5) **Skills** - Demand for AI talent is surging, but gaps persist. Education and training programs are needed.

The section of the White Paper concludes with recommended policy responses in areas such as investments, regulation, workforce development, ethics, and international coordination. A measured, holistic approach is required to successfully integrate AI's emerging capabilities into economies and societies.

5.1 Automation - Risks and Rewards

One of the most transformative economic impacts of AI will be its ability to automate a widening range of tasks and jobs. According to McKinsey, AI could automate 30% of the activities in 60% of all occupations by 2030 (Bughin et al., 2018). The jobs at the highest risk of automation include physical ones in predictable environments, such as assembly line manufacturing and food preparation. White-collar office jobs such as bookkeeping and payroll processing also face significant automation potential from AI. For companies, automating tasks with AI technologies presents a significant opportunity to reduce labor costs, improve efficiency and productivity, and enhance competitiveness. Amazon provides an instructive example of using AI automation throughout its operations, from inventory management in warehouses to customer service interactions.

However, for workers, particularly low- and medium-skilled workers in automatable occupations, the spread of AI poses risks of job losses and reduced incomes. A recent study by PwC estimated that up to 30% of UK jobs could potentially be at high risk of automation by the 2030s (Jain, 2018). Nevertheless, while initial studies have indicated a high risk of automation, a more sophisticated analysis is worth considering the extent to which augmentation rather than automation will take place. Indeed, the initial studies conducted by QCRI* suggest that our focus should be primarily focused on augmentation. Despite this, significant job losses could occur in manufacturing, transportation, storage, retail, and administrative roles. Workers displaced by automation may struggle to find new employment, especially if they lack the skills to transition to the new jobs created by AI.

Policymakers face the challenge of supporting productivity and efficiency gains from AI automation while instituting safeguards to protect workers. Potential policy responses could include:

1) Education programs to re-skill workers and prepare students for AI-driven job markets.

2) Strengthening social safety nets, including unemployment benefits and universal basic income programs.

3) Incentives to encourage companies to retrain and redeploy workers affected by automation.

4) Investments in infrastructure, healthcare, education, and other human-focused sectors to create new employment opportunities.

5) Modernizing labor laws to facilitate flexible work arrangements while protecting workers' rights.

5.2 New Industries - Expanding Possibilities

In addition to transforming existing sectors, AI is catalyzing innovative new products, services, and economic activities. AI's capabilities in analyzing data, identifying patterns, predicting outcomes, and generating multimedia content are opening up novel applications across industries from transportation to finance to healthcare.

* https://qcai-blog.qcri.org/wp-content/uploads/2020/11/The-Impact-of-Artificial-Intelligence.pdf

One area seeing intensive AI innovation is the development of autonomous vehicles (AVs). Companies like Tesla are leveraging AI software to create AVs that can navigate diverse driving environments. The advent of automated mobility-as-a-service offered by AVs could produce new business models and revenue streams in transportation. Other emerging AI-powered industries include precision agriculture, which utilizes computer vision and deep learning to optimize crop yields; fintech, which applies AI techniques such as machine learning to credit decisions and fraud detection; and drug discovery, which harnesses AI to significantly accelerate the search for new pharmaceuticals. Startups embracing AI are proliferating worldwide, attracting increasing investment and creating new high-skilled jobs.

To maximize the economic benefits from AI-enabled new industries, policymakers should consider measures such as:

- **Supporting research, entrepreneurship, and innovation clusters** - Funding academic research in AI applications to augment current work practices. Building incubators and accelerators for AI startups. Developing regional innovation clusters around universities and companies. Indeed, an opportunity exists for the focus of the next Qatar National Vision to directly incorporate the role of AI in catering for an AI-enabled knowledge-based society.

- **Investing in data infrastructure** - Building open repositories of high-quality data to fuel innovation. Expanding broadband access and 5G connectivity.

- **Updating regulations** - Modernizing laws and regulations in areas such as autonomous vehicles and drones. This is particularly pertinent for the developing defense sector in Qatar.

- **Partnerships** - Facilitating partnerships between industry and academia to translate innovations into enterprises. This can be further leveraged within the QSTP.

- **IP frameworks** - Implementing balanced intellectual property regimes to incentivize innovation while allowing access to knowledge. This can cater for some functions within the legal system in Qatar.

- **Skills development** - Expanding higher education programs in AI-related fields to develop talent pipelines. It is proposed here that there will be a greater focus on the curriculum in schools to educate students from an early age on the benefits and risks of AI technology. This can be particularly useful in terms of creating awareness of misinformation and disinformation.

5.3 Economic Inequality - An AI Divide?

While artificial intelligence presents possibilities for economic growth, its impacts may disproportionately benefit higher-skilled workers, potentially exacerbating income and wealth inequality. As machines increasingly substitute routine manual and cognitive tasks, the comparative advantage of human labor shifts to non-routine skills demanding flexibility, problem-solving, creativity, and social intelligence. Highly educated workers possessing those skills are positioned to benefit from complementing AI systems and capitalizing on productivity benefits.

However, medium and low-skilled workers in automatable clerical, manufacturing, retail, and services roles face the risk of job losses or wage reductions. Indeed, this may have an impact on reducing the

proportion of expatriate labor needed in the Gulf context. A scenario could therefore emerge where high-skill, high-wage AI talent prospers in lucrative roles designing, managing, and servicing intelligent systems. To better manage this, policies should include major investments in STEM education and retraining/lifelong learning programs to upgrade workforce skills. This is particularly pertinent in the case of Qatar and the wider GCC, given its youthful demographic structure and the need to avoid unemployment, underemployment, and greater participation in the private sector. Indeed, the goal of a knowledge-based economy is therefore probably an AI-enabled knowledge-based economic model; therefore, skilling will be a key factor moving forward.

5.4 Enhancement of Decision Making

In addition to transforming business operations and processes, AI has significant potential to enhance decision making, despite being at its very early stage. By rapidly analyzing vast amounts of data, identifying patterns and relationships, and predicting outcomes, AI systems can provide improved diagnostics, forecasts, and insights that support more informed decisions. In this respect, it augments and helps support efficiency, but it is not taking a role in decision making as the technology is still nascent. Nevertheless, some impressive advances have already taken place; in healthcare, AI tools can scan medical images to identify tumors and other abnormalities earlier with higher accuracy than humans, potentially saving lives through improved preventive screening (He et al., 2019). Algorithmic systems can also analyze demographic, clinical, and genetic data to tailor treatment plans to individual patients (Johnson et al., 2021).

In finance, AI techniques are used for everything from credit-risk modeling to portfolio management and algorithmic trading. By discovering correlations and patterns in financial data, AI systems can backtest strategies for robustness and execute trades faster than human traders. However, increased reliance on AI systems for impactful decisions also raises concerns about accountability and control. When autonomous AI systems make erroneous or harmful decisions, legal and ethical questions arise regarding who should be held responsible. If AI systems displace human decision-makers in domains such as medicine and finance, what does this imply for professional ethics and standards of care? Who will ensure that AI systems are making socially beneficial decisions aligned with human values?

It is worth highlighting here that the manner in which social media platforms operate is a risk factor that needs to be considered. Social media platforms use algorithms that prioritize content that gets more engagement. This rewards inflammatory posts that evoke strong reactions that raise numerous issues. In terms of how this impacts society, it is well established that this fragments discourse as users only see opinions they agree with. It also incentivizes misinformation and divisive content from bad actors. While social media connects people, flawed algorithms amplify extreme voices, which poses a risk of greater polarization within societies.

5.5 Growing Demand for AI Talent

As AI technologies proliferate, the demand for AI talent has surged worldwide. Technology giants, startups, consultancies, and organizations across sectors are competing intensely for professionals skilled in areas such as machine learning, natural language processing, robotics, computer vision, and data science. The proliferation of corporate AI research labs and acquisitions reflects the priority placed on

recruiting scarce AI expertise. However, shortages of AI talent persist because of gaps in both educational programs and practical experience (Lane and Saint-Martin, 2021). Although many universities now offer majors, courses, and certificates in AI-related fields, hands-on integration of cutting-edge tools and methods often lags.

Addressing current skill gaps and developing future talent will require efforts including:

- **Educational alignment** - Universities should actively engage with industry to understand in-demand AI competencies and update curricula accordingly. A good example here is that large language models create demand for prompt engineering, as a sophisticated ability in language prompts can cater for higher quality outcomes. In certain respects, this gives greater emphasis to linguistics as a field that can evolve into one that feeds into AI applications.

- **Retraining** - Creating mid-career refilling programs in AI skills for professionals with backgrounds in engineering, math, physics, computer science, and other related fields.

- **Immigration** - Easing immigration policies to attract international AI experts to Qatar, balanced with investments in domestic skill development. Indeed, countries such as Japan and Australia have visa regimes that incentivize highly skilled individuals to be granted visas.

- **Vocational training** - Expanding vocational training programs focused on applied AI skills needed by industry. This offers opportunities for the educational sector to not only focus on higher education research, but also for taught programs to be much more vocational in their orientation rather than the current model of pedagogy-based delivery that dominates programs within Qatar.

- **K-12 exposure** - Introducing AI concepts in K-12 education to spark early interest and aptitude and removing barriers to accessing computing education. As highlighted above, this is particularly important concerning not only sparking interest in students but also making them aware of the risks of technology for the purposes of misinformation and disinformation, deep fakes, and the need for cyber resilience.

5.6 Summary Observations

Artificial intelligence presents major opportunities and risks for Qatar's economy and society. On the upside, AI can significantly boost productivity, enhance decision-making, and enable new industries and business models. However, automation may displace many routine jobs, exacerbating inequality if lower-skilled local workers lack retraining options.

Qatar will need strategic investments and policies to maximize the benefits of AI. Key priorities should include building capabilities in AI research and entrepreneurship, developing local talent through education reforms and refilling programs, updating regulations to permit testing of emerging technologies, and directing AI applications to priority sectors such as healthcare, education, energy, and transportation. Through thoughtful leadership and planning, Qatar can leverage AI to diversify its economy and propel the growth of its private sector, as outlined in the National Vision 2030. However, this will require proactive steps to equip Qatari youth with the technical and creative skills to complement AI systems and lead in imagining novel applications.

5.7 Policy Implications and Future Research Agenda

Based on the above analysis, GISR has identified the following policy implications and areas for future research on AI's global development, risks, and geopolitics:

POLICY IMPLICATIONS

- Develop a renewed national AI skilling strategy with clear implementation steps and an approach for stakeholder engagement in Qatar. This will need to focus on developing capabilities, governing risks, and directing applications toward national priorities such as education, healthcare, energy, and economic diversification.

- Implement comprehensive policies for responsible AI governance of algorithmic accountability, data privacy, AI safety certification, and oversight.

- Expand access to AI technical skills training and computer science education to enhance opportunities for citizens to participate in the AI economy.

- Cultivate public awareness and engagement on AI's benefits and risks through proactive communication and outreach initiatives.

FUTURE GISR RESEARCH AGENDA

- Evaluate policies to balance productivity gains from AI automation against employment impacts on citizens.

- Assess the educational policies needed to develop AI talent pipelines among youth across the Gulf states.

- Forecast infrastructure and investments required to prevent the emergence of a Gulf AI divide.

- Examine AI's role in achieving goals such as environmental sustainability, human development, and public engagement.

6.0 Ethical Implications and Regulation

AI is spreading rapidly into everyday life, bringing enormous opportunities with clear risks. As Gulf countries pursue AI strategies, suitable governance is required to direct this technology toward beneficial ends. This section of the white paper provides analysis and observations on key ethical issues surrounding AI and potential regulatory approaches, with a focus on strategic priorities for GCC states.

6.1 Understanding the Disruptive AI Trajectory

As highlighted in this white paper, AI refers to computer systems exhibiting human-like intelligence - able to perceive contexts, reason about problems, make choices, and learn from experience. Rather than just executing fixed programed instructions, AI techniques such as machine learning statistically analyze data to recognize patterns, predict outcomes, or generate new and innovative content appropriate for different situations.

As computational power surges and data proliferates, AI capabilities are rapidly expanding into diverse domains such as conversational agents, recommendation engines, autonomous vehicles, medical diagnosis, creative generation, fraud detection, and legal research. Intelligent assistants such as Apple's Siri perform useful tasks through natural dialog, while product recommendation engines analyze individual interests to suggest personalized content. Autonomous vehicles and drones navigate environments safely without human intervention and AI medical tools help diagnose illnesses, suggest treatments, and even discover new drugs. On the creative front, AI systems can generate original music, paintings, videos, and literature based on human direction. Fraud detection AI systems can identify illegal financial transactions among millions of innocuous activities, and legal research AI systems can mine millions of case documents to find the most relevant precedents for specific cases. These examples only scratch the surface of AI's increasing impact across industries and society.

As costs fall and capabilities rise, AI adoption is accelerating across public and private sector. PwC estimates that AI could add $320 billion to the Middle East's GDP by 2030 (Jain, 2018). However, realizing AI's full benefits requires navigating complex risks. Like other disruptive technologies, AI holds great promise but also poses challenging ethical questions:

- How can disproportionate workforce disruption from automation be mitigated?
- When should decision making delay to AI algorithms versus human discretion?
- How can biases encoded in data or algorithms be addressed to ensure fair outcomes?
- How can complex AI models be interpreted and held accountable when their reasoning defies explanation?

- How can privacy be preserved when AIs continuously collect and analyze sensitive personal information?
- How can security be guaranteed as networked AI systems become more autonomous?
- How can abuses such as mass disinformation generation or invasive surveillance be averted as capabilities advance?
- How can catastrophes caused by uncontrolled hyperintelligent AI be prevented if theories about impending exponential takeoff prove true?

Governing AI systems according to ethical values is crucial but challenging. AI risks span technical, social, economic, and political dimensions, requiring diverse expertise to evaluate tradeoffs. Technical standards and design practices directly shape AI behaviors and impacts.

6.2 Different Regulatory Approaches and Perspectives

Many governments now recognize that AI's disruptive nature necessitates updated governance regimes, but national policy responses vary based on differing attitudes toward technology, innovation, regulation, and values (Bradford, 2020):

UNITED STATES

America's light-touch approach emphasizes unfettered AI innovation with limited regulation. No federal AI regulations exist; sector-specific agencies oversee narrow applications. This decentralized and permissive U.S. model reflects deeply held cultural beliefs about technological progress and free markets with minimal state interference. Regulation is viewed as a hindrance to innovation. Technology firms exert enormous influence in promoting this perspective. Critics contend that leaving AI governance to corporations allows avoidable harms, citing social media's destabilizing effects. Overall, the US regulatory environment remains innovation-oriented and industry-guided. Comprehensive oversight mechanisms face political resistance.

EUROPEAN UNION

Europe sees AI as carrying novel risks that require precautionary regulation. The EU's AI Act imposes obligations across all AI systems proportional to the severity of potential harm. It completely prohibits inherently unethical applications such as social credit scoring and subliminal manipulation that contravene European values. "High-risk" AI uses such as self-driving cars or employee hiring algorithms must implement risk management procedures. Violations incur steep fines, including up to 6% of a company's global revenue. This strict regulatory approach upholds fundamental rights, although critics contend that it may limit the beneficial innovation.

CHINA

The Chinese government actively steers AI to enhance prosperity and social harmony. Through state subsidies and research, China aims to become the world's "AI superpower" by 2030 (Lee, 2018).

Simultaneously, authorities will impose censorship on generative chariots and other AI deemed politically destabilizing or socially deviant. While aggressively pursuing AI advantages, Beijing recognizes that generative AI poses novel risks to state information control.

GULF PERSPECTIVES

GCC states have broadly shown enthusiasm to reap AI's benefits, but lack comprehensive ethical frameworks or binding governance systems. Progress has been made as Saudi Arabia's national AI policy articulates ideal ethical aims, as has the UAE's appointment of the world's first Minister of AI in 2017 (Memish et al., 2021). The proposed regulations in Kuwait, Bahrain, and Qatar hint at nascent governance efforts (Hanafi et al., 2021). However, transforming principles into enforceable, accountable oversight remains absent. Joint GCC-wide governance mechanisms could efficiently develop regional standards. As more Saudi and Emirati plans advance, a GCC framework may necessitate interim national solutions. However, long-term cooperation offers opportunities to pool insights and capabilities.

It should also be recognized that given the political economy character of the Gulf states, a move away from a productive knowledge-based economy also means diversification away from reliance on generous state subsidies and oil and gas income. Although progress has been made in the case of Qatar, the economy remains largely reliant on income from the gas, petrochemicals, and oil sectors. While activities of key organizations such as the QIA to derive a diversified form of income from external investments are important, this is quite different from a diversification of the economic model within the country to one where income comes from a knowledge-based society. It is therefore recognized that this transition is highly challenging given the sociocultural dimensions and historical legacy of how development has been achieved. Nevertheless, building capacity has the potential to be augmented within the economic system in a way that can enhance productivity and growth so that a wider proportion of national income comes from such industries.

While the potential is clear, as pressure for governance mounts, questions around reconciling Islamic ethics with AI's societal impacts require further research. Islamic values are paramount across the GCC, and these values would be violated by certain applications such as deepfakes, AI-enabled harassment, or fraud. Quranic principles regarding truth and the need to avoid deceit are relevant to interpreting the risks of AI disinformation capabilities. Islamic thought also offers principles to guide wise governance, such as accountability, human stewardship of technology, and the pursuit of knowledge for social progress. Therefore, it is recognized that this will be a central means by which a Gulf-centric regulatory model can meet the needs of society and inform broader international discussions on regulations and safeguards.

6.3 Policy Priorities and Options for Guiding AI Positively

Comprehensive AI governance combining legal constraints, voluntary best practices, incentives, safety investments, and global coordination can help steer AI's trajectory positively. However, overly rigid restrictions may also limit the beneficial use of AI. Policies should target specific high-risk applications supported by evidence of potential harms while taking an experimental approach to promote learning about impacts over time. Options include requiring impact assessments before deploying higher-risk AI to evaluate benefits and fairness. Policymakers can also enact prohibitions on applications that violate

human dignity, such as a certain biometric tracking and surveillance use that erode rights. Mechanisms that ensure individuals have visibility into data and AI systems affecting them can support explainability and consent.

Other options include incentives to encourage AI ethics programs in firms via procurement preferences and flexible regulatory requirements to motivate accountability. Funding more R&D focused on AI safety, machine ethics, and value alignment allows for anticipating risks as capabilities advance. Requiring ethics education in training programs can help increase talent diversity and reduce development blind spots. Proactively addressing cybersecurity challenges through standards, monitoring, and infrastructure resilience is also essential (Hassib and Shires, 2022). Partnering with religious scholars to evaluate emerging AI risks relative to Islamic principles enables the development of joint guidance on ethics.

Where unacceptable harms exist, criminalizing malicious AI abuses such as disinformation and enforcing prohibitions on fraud and deception can maintain public trust. Seeking both GCC-wide and broader international accords on shared values, safety guidelines, and prohibiting destabilizing military applications allows for cooperatively governing AI as the field continues to grow.

6.4 Summary Observations

The rapid development of AI brings enormous opportunities and risks. As GCC countries pursue national AI strategies, suitable governance is imperative to direct this powerful technology toward beneficial ends aligned with Islamic ethics and values. Key policy priorities should include impact assessments for high-risk AI uses, prohibitions on unethical applications eroding rights, mechanisms ensuring explainability and consent, incentives for ethics programs and R&D safety investments, increased talent diversity, cybersecurity protections, and partnerships with religious scholars. Where unacceptable harms exist, focused prohibitions coupled with GCC and international cooperation can maintain public trust while allowing AI innovation to flourish responsibly. Holistic governance combining targeted constraints, flexible oversight, standards, diverse expertise and global collaboration will allow collectively steering AI's trajectory to empower people and enable human flourishing in line with principles of justice, accountability, and human dignity.

In summary, comprehensive and adaptive AI governance is crucial to realize the full social and economic benefits of AI while mitigating complex risks across technological, social, economic, and political dimensions. As GCC countries advance national strategies, they have opportunities to incorporate leading governance models from the outset that proactively address issues around security, fairness, transparency, and human autonomy. Prudent oversight mechanisms can foster responsible AI innovation aligned with Islamic ethics and the common good.

6.5 Policy Implications and Future Research Agenda

Based on the above analysis, GISR has identified the following policy implications and areas for future research on ethical implications and regulation.

POLICY IMPLICATIONS

- Examine the adequacy of current laws relating to AI abuses such as fraud, deception, and disinformation. Develop a national AI ethical and regulatory framework that is aligned with core national values.
- Require impact assessments for high-risk AI systems to evaluate fairness and unintended consequences.
- Incentivize ethics programs and AI safety R&D through procurement and flexible regulations.
- Expand access to AI ethics education and training to increase diverse perspectives.

FUTURE GISR RESEARCH AGENDA

- Survey Qatari public and government perspectives on priorities for AI governance.
- Study approaches to regulating AI aligned with Islamic ethical principles.
- Assess lessons from global AI governance models applicable to Qatar and the GCC.
- Analyze technical and policy options for auditing AI systems for bias and fairness.
- Research legal liability frameworks for harms caused by autonomous AI systems.

7.0 References

AGRAWAL, A., GANS, J. S. GOLDFARB, A. 2019. Artificial intelligence: the ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives,* 33**,** 31-50.

ALMESAFRI, A. HABES, M. Understanding the Use of Artificial Intelligence (AI) for Human Resources in the Dubai Government. International Conference on Business and Technology, 2022. Springer, 417-428.

ARIF, S. 2021. Militarization of artificial intelligence: Progress and implications. *Toward an International Political Economy of Artificial Intelligence*. Springer.

BENBYA H, PACHIDI, S. JARVENPAA, S. 2021. Special issue editorial: Artificial intelligence in organizations: Implications for information system research. *Journal of the Association for Information Systems,* 22**,** 10.

BRADFORD, A. 2020. *The Brussels effect: How the European Union rules the world*, Oxford University Press, USA.

BRKAN, M. 2019. Artificial Intelligence and Democracy: The Impact of Disinformation, Social Bots and Political Targeting. *Delphi,* 2**,** 66.

BUGHIN, J., SEONG, J., MANYIKA, J., CHUI, M. JOSHI, R. 2018. Notes from the AI frontier: Modeling the impact of AI on the world economy. *McKinsey Global Institute,* 4.

HAENLEIN, M. & KAPLAN, A. 2019. A brief history of artificial intelligence: The past, present, and future of artificial intelligence. *California Management Review,* 61**,** 5-14.

HANAFI, M. M., KSHETRI, N. SHARMA, R. 2021. Economics of artificial intelligence in the gulf cooperation council countries. *Computer,* 54**,** 92-98.

Hasan, O. 2020. Artificial intelligence, Neom, and Saudi Arabia's economic diversification from oil and gas. *The Political Quarterly,* 91**,** 222-227.

HASSANI H, SILVA, E. S., UNGER, S., TAJMAZINANI, M. & MAC FEELY, S. 2020. Artificial intelligence (AI) or intelligence augmentation (IA): what is the future? *Ai,* 1**,** 8.

HASSIB, B. SHIRES, J. 2022. Cybersecurity in the GCC: From economic development to geopolitical controversy. *Middle East Policy,* 29**,** 90–103.

HE, J., BAXTER S L, XU, J., XU, J., ZHOU, X. & ZHANG, K. 2019. Practical implementation of artificial intelligence technologies in medicine. *Nature medicine,* 25**,** 30-36.

JAIN, S. 2018. US $320 billion by 2030?: The potential impact of AI in the Middle East, PwC.

JOHNSON, K. B., WEI WQ, WEERARATNE D, FRISSE ME, MISULIS, K., RHEE, K., ZHAO, J. SNOWDON, J. L. 2021. Precision medicine, artificial intelligence, and the future of personalized health care. *Clinical and Translational Science,* 14**,** 86-93.

Jones, M. O. 2022. *Digital authoritarianism in the Middle East: Deception, disinformation and social media*, Hurst Publishers.

LANE, M. SAINT-MARTIN, A. 2021. Impact of Artificial Intelligence on the Labor Market: What do we know so far?

LEE, K.-F. 2018. *AI superpowers: China, Silicon Valley, and the new world order*, Houghton Mifflin.

MEMISH, Z. A., ALTUWAIJRI, M. M., ALMOEEN, A. H. ENANI, S. M. 2021. The Saudi Data & Artificial Intelligence Authority (SDAIA) Vision: Leading the Kingdom's Journey toward Global Leadership. *Journal of Epidemiology and Global Health,* 11**,** 140.

NAUGHTON, B. 2020. Chinese Industrial Policy and the Digital Silk Road. *Asia Policy,* 15**,** 23-40.

RICKLI, J.-M. 2018. Economic, security, and military implications of artificial intelligence for the Arab Gulf countries. *Emirates Diplomatic Academy***,** 1-13.

SHAMOUT, F. E. ALI, D. A. 2021. The strategic pursuit of artificial intelligence in the United Arab Emirates. *Communications of the ACM,* 64**,** 57-58.

STATE OF QATAR 2019. National Artificial Intelligence Strategy for Qatar. *In:* MINISTRY OF COMMUNICATION AND INFORMATION TECHNOLOGY (ed.). Doha: State of Qatar.

VINUESA R, AZIZPOUR H, LEITE I, BALAAM, M., DIGNUM V, DOMISCH S, FELLNDER A, LANGHANS SD, TEGMARK, M. & FUSO NERINI, F. 2020. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications,* 11**,** 1-10.

WESTERLUND, M. 2019. The emergence of deepfake technology: A review. *Technology innovation management review,* 9.

WHYTE, C. 2020. Deepfake news: AI-enabled disinformation as a multi-level public policy challenge. *Journal of Cyber Policy,* 5**,** 199-217.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The Global Institute for Strategic Research is an independent and inter-disciplinary think tank concerned with the global challenges of governance, advancement – defined as progress, prosperity, and development – and peace (GAP). **GISR** seeks to document trends, understand processes, unpack power relations, analyze root causes, and explore policy solutions. Our work combines structural analysis with the power of imagination to go beyond the conventional extrapolation of current developments and trends. **GISR's** research and ability to convene global policy dialogues amongst stakeholders offer policymakers in Qatar (and the region) the foresight necessary for making better-informed decisions, based on having considered future eventualities, scenarios, and outcomes.

Founded by Qatar Foundation's Hamad Bin Khalifa University and partnered with a network of leading local, regional, and global collaborators, **GISR** is international in its orientation while aiming to make a meaningful impact on global policy from an Arab and regional point of view.