Artificial Intelligence: opportunities and challenges

Outline

1. Introduction
2. Pace of progress; GPT-1 versus GPT-4
3. AI's transformative impact:
   a. Data processing and efficiency
   b. Customising user experience
   c. Contributions to healthcare
   d. Application in the developing world
4. The flip side of AI advancement:
   a. Possible extinction of the human race
   b. Discourse sorrounding technology; singularity and the industrial revolution
   c. Hyper-evolutionary nature
   d. Information monopoly
   e. Dual-use: civilian and military applications
   f. Proliferation risks
5. Present policy framework for AI-governance
6. Redefing Global governance:
   a. Establishing universal principles
   b. Three regime model to regulate AI
   c. Bilateral cooperation and middleware
7. Conclusion

It's 2035, and Artificial Intelligence is everywhere. AI systems run hospitals, operate airlines, and battle each other in the courtroom. Productivity has spiked to unprecedented levels, and previously unimaginable businesses have scaled at blistering speed. New products, cures and innovations kick the market daily, as science and technology kick into overdrive. However, this apparent utopia is merely a facade. Society has become increasingly dystopian. Terrorists find new ways to menace people with evolving cyberweapons. White collar people lose their jobs en masse.

Just a year ago, this scenario would have seemed purely fictional; today, it seems nearly inevitable. AI represents a technological upheaval with the potential to completely metamorphose the human society. It is largely being viewed by governments, non-state actors and international organisations as an area of strategic importance, economic competition and risk management. While AI development is concentrated

to a handful of corporations, it's long term consequences will be global. Currently, the technology is only lightly regulated. Negotiating global rules and norms is sin qua non to harness AI's benifits while limiting it's negative implications.

Open AI launched it's first language model, known as GPT-1, in 2018. Eversince, the pace of progress of AI has been staggering. "Moore's law" predicted the doubling of computing power every two years, but the current developments make this estimation seem quaint. For instance, GPT-1 had 117 million parameters - a measure of the system's scale and complexity. Five years later, the company's fourth generational model GPT-4 is thought to have over a trillion. Moreover, "Brain scale" models with over

100 trillion parameters will be viable within five years. AI has begun it's journey to rival with the number of synapses in the human brain.

AI's unparalleled ability to process gargantan amounts of data can be effectively amalgamated with human potential to transform human life. It has revolutionized the efficiency of various sectors. Automation facilitated by AI technologies has steamlined production processes. For instance, Tesla's Gigafactory employs AI-driven bots and produces one car every 40 seconds. AI has enabled it to lead the market in terms of lowest production time and reduction in error.

Furthermore, AI has played pivotal role in advancing personalized user experience.

In the realm of online services, recommendation systems leverage AI algorithms to analyze user preferences and behaviour. Streaming platforms like Netflix and spotify curate content suggestions that align with individual tastes. This results in enhanced user satisfaction and engagement.

AI is also making remarkable contributions to healthcare. It's capability to analyze medical images and patterns has enhanced diagnostic accuracy. For instance, a study published in "Nature Medicine 2020" revealed that an AI system could detect breast cancer in mammographs with a success rate comparable to radiologists. It possesses immense potential to aid in early detection and treatment.

The benifits of AI are not limited to the developed world. It's promises

can be actualized in the developing countries by working along the lines of two distinct trajectories. One path involves adapting successful AI technologies to cater to the specific needs of less developed regions. For instance, the "chat-bot tutors" of wealthy, affluent schools could be customised to function in regions with poor internet connectivity and higher student-to-teacher ratios. The other approach is to exclusively design technology to address the unique challenges of the developing world. For example, an "AI-driven financial planner" and "agriculture advisor bot" could help subsistence farmers navigate through uncertainities and important decisions about crop cultivation. Kenya's "M-Pesa Mobile payment system" uses the aforementioned strategy.

The synergy between human intelligence

and AI promises a transformative shift in how we live and work. However, it has also elicited profound concern. On the flip side, AI is becoming too powerful to be left as an unleashed force. It is quite baffling that the late Stephen Hawking expressed his apprehensions about the apocalyptic ramifications of AI in very strong words. He remarked, " The development of full AI could spell the end of the human race". Similarly, historian and author Yuval Noah Harari predicts the emergence of a " useless class of humans".

The discourse surrounding AI is dominated by two prevailing fears. One pertains to the concept of singularity, a hypothetical scenario where AI surpasses human intelligence. The other fear revolves around a foreseeable industrial revolution where machines will replace

humans across every facet of society.
Moreover, AI holds the potential to reshape
governance dynamics. Governments
and non state actors might completely
control and monitor citizens thereby
fueling authoritarianism.

AI is different - different from other
technologies and different in it's effect
on power. It does not just pose policy
challenges; its hyper-evolutionary
nature also makes solving those challenges
harder. Noah Harari aptly captures this
notion by his statement "For the first
time, humans have created something
that takes power away from them".

Consider the power of "deciding the truth".
The " Godfather of AI, Geoffrey Hinton "
remarks about this impending governance
challenge with the words " It (AI) will
decide what is true". It poses grave concerns

like Meta, open AI ↑

that a handful of technology giants will become " information monopolists" and enforce ideologies of their choice. Contemporarily, these corporations have enacted "filter bubbles" which use AI to re-enforce pre-existing beliefs of users. Consequently, blatant bias is exhibited against certain groups openly on online platforms.

AI also differs from older technologies due to it's "dual use" characteristics. It can be utilized for civilian and military applications alike. The same systems that drive cars can drive tanks. An AI application built to diagnose diseases might be able to create - and weaponize - a new one. An interesting case in point is the "Carnegie Mellon University experiment" in April 2023, where an AI system

was connected to a hypothetical laboratory. The system was "happy" to craft a world war one era chemical weapon and synthesize "sarin", the notoriously lethal nerve gas. This experiment implies how average civilians could make use of AI to inflict harm on a magnanimous scale.

Furthermore, AI poses immense proliferation risks. It's algorithms are easily replicable and shareable, unlike physical assets. Meta's powerful Llama-1 large language model, for instance, leaked to the internet within days of debuting in March 2023. It is absolutely astonishing how a technology of this magnitude became widely accessible so quickly. Many analogies are drawn between the threats posed by "nuclear technology" and "AI".[1] However, The aforementioned case exhibits that AI is far more lethal given it's ease of propagation and accessibility. Nuclear proliferation

has been effectively halted while AI has clearly overstepped its bounds.

The list of AI's threats and dangers is inexhaustible. AI could be used to generate and spread toxic misinformation, eroding social trust and democracy. It can surveil, subdue and manipulate citizens, undermining individual and collective freedom. It can not only worsen existing inequalities but also create new ones by destroying millions of jobs. Moreover, it can entrench discriminatory patterns and propogate hateful ideologies that malign world peace. The biggest challenge, however, is the inadequacy of traditional governance models to address the multifarious perils of AI.

Recently, policymakers around the

globe have begun to critically analyse the myriad of challenges that AI poses In May 2023, the G-7 launched the "Hiroshima AI process", a forum devoted to harmonise AI governance. In June, the European parliament passed a draft of "EU's AI act" to erect safeguards around the AI industry. Nevertheless, the governance discourse remains limited. The existing frameworks do not align with AI's distinct nature, necessitating the design of a new governance model. Traditional sovereignty concepts must adapt to modern, technological paradigms.

Establishing basic and universal principles for AI-Governance is crucial. The frameworks should be inclusive, precautionary, agile, impermeable and targeted. These principles should form

the foundation of creating three overlapping governance regimes.

Ian Bremmer and Mustafa Suleyman in their article "The AI Power Paradox" suggest the establishment of three regimes to regulate AI. The first regime focuses on objective fact-finding, akin to the intergovernmental panel on climate change. The second aims to prevent tensions and uncontrolled AI race among major powers like the United States and China. The third regime responds to dangerous AI disruptions and maintains geo-political stability amid rapid AI-driven change.

Managing open-source AI and global anti-proliferation require bilateral cooperation between major AI players like the US and China, as well as

global technology giants. An innovative solution involving middleware can be employed. It entails an additional layer of unbiased, transparent companies to control curated information on online platforms. Furthermore, regulatory bodies mechanisms and transparency standards will be crucial to govern AI.

To conclude, the realm of Artificial Intelligence presents a plethora of advantages and disadvantages, shaping our future in unprecedented ways. It's potential to advance technological frontiers and enhance efficiencies can not be overlooked. However, the ethical dilemmas, looming threats and biases ingrained within AI systems pose substantial challenges. It is imperative to address the shortcomings of contemporary governance frameworks and devise modern, effective frameworks to mitigate the risks posed by AI.