Responsible Artificial intelligence

sanskritiias.com/current-affairs/responsible-artificial-intelligence

(MainsGS-3: Science and Technology- developments and their applications and effects in everyday life.)

Context:

- Optimizing logistics, detecting fraud, composing art, conducting research, providing translations: intelligent machine systems are transforming our lives for the better.
- Intelligence comes from learning, whether you're human or machine. AI machines uses data in which they "learn" to detect the right patterns and act according to their input.
- If we rely on AI to bring us into a new world of labour, security and efficiency, we need to ensure that the machine performs as planned, and that people can't overpower it to use it for their own ends.

What is artificial intelligence?

- Artificial intelligence (AI) is wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence.
- AI is an interdisciplinary science with multiple approaches, but advancements in machine learning and deep learning are creating a paradigm shift in virtually every sector.
- The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal.
- A subset of artificial intelligence is machine learning, which refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans.
- Deep learning techniques enable this automatic learning through the absorption of huge amounts of unstructured data such as text, images, or video.

Categorization of Artificial Intelligence

- Artificial intelligence can be divided into two different categories: weak and strong.
- **WEAK AI:** embodies a system designed to carry out one particular job. Weak AI systems include video games such as the chess example from above and personal assistants such as Amazon's Alexa and Apple's Siri. You ask the assistant a question, it answers it for you.
- **STRONG AI:** these are systems that carry on the tasks considered to be human-like.
- These tend to be more complex and complicated systems.
- They are programmed to handle situations in which they may be required to problem solve without having a person intervene.
- These kinds of systems can be found in applications like self-driving cars or in hospital operating rooms.

AI's exponential growth:

- AI is embedded in the recommendations we get on our favourite streaming or shopping site; in GPS mapping technology; in the predictive text that completes our sentences when we try to send an email or complete a web search.
- It promises to be even more transformative than the harnessing of electricity. And the more we use AI, the more data we generate, the smarter it gets.
- In just the last decade, AI has evolved with unprecedented velocity — from beating human champions at Jeopardy! in 2011, to vanquishing the world's number one player of Go, to decoding proteins last year.
- Automation, big data and algorithms will continue to sweep into new corners of our lives until we no longer remember how things were "before".
- Just as electricity allowed us to tame time, enabling us to radically alter virtually every aspect of existence, AI can leapfrog us toward eradicating hunger, poverty and disease — opening up new and hitherto unimaginable pathways for climate change mitigation, education and scientific discovery.

AI for betterment:

- AI has helped increase crop yields, raised business productivity, improved access to credit and made cancer detection faster and more precise.
- It could contribute more than $15 trillion to the world economy by 2030 adding 14% to global GDP.
- Google has identified over 2,600 use cases of "AI for good" worldwide.
- A study published in *Nature* reviewing the impact of AI on the Sustainable Development Goals (SDGs) finds that AI may act as an enabler on 134 — or 79% — of all SDG targets.
- We are on the cusp of unprecedented technological breakthroughs that promise to positively transform our world in ways deeper and more profound than anything that has come before.

Worries created by AI:

- Yet, the study in *Nature* also finds that AI can actively hinder 59 — or 35% — of SDG targets.
- For starters, AI requires massive computational capacity, which means more power-hungry data centres — and a big carbon footprint.
- AI could compound digital exclusion.
- Robotics and AI companies are building intelligent machines that perform tasks typically carried out by low-income workers: self-service kiosks to replace cashiers, fruit-picking robots to replace field workers, etc.
- Days are not far when many desk jobs will also be edged out by AI, such as accountants, financial traders and middle managers.

Creating inequality:

- Without clear policies on reskilling workers, the promise of new opportunities will in fact create serious new inequalities.
- Investment is likely to shift to countries where AI-related work is already established, widening gaps among and within countries.
- Together, Big Tech's big four — Alphabet/Google, Amazon, Apple and Facebook — are worth a staggering $5 trillion, more than the GDPs of just about every nation on earth.
- In 2020, when the world was reeling from the impact of the COVID-19 pandemic, they added more than $2 trillion to their value.
- The fact is, just as AI has the potential to improve billions of lives, it can also replicate and exacerbate existing problems, and create new ones.
- For instance, the documented examples of AI facial recognition and surveillance technology discriminating against people of colour and minorities.

Privacy worries:

- AI also presents serious data privacy concerns.
- The algorithm's never-ending quest for data has led to our digital footprints being harvested and sold without our knowledge or informed consent.
- We are constantly being profiled in service of customisation, putting us into echo chambers of like-mindedness, diminishing exposure to varied viewpoints and eroding common ground.
- Today, it is no exaggeration to say that with all the discrete bytes of information floating about us online, the algorithms know us better than we know ourselves.
- They can nudge our behaviour without our noticing.
- Our level of addiction to our devices, the inability to resist looking at our phones, and the chilling case of Cambridge Analytica, in which such algorithms and big data were used to alter voting decisions

- These serves as a potent warning of the individual and societal concerns resulting from current AI business models.
- In a world where the algorithm is king, it behoves us to remember that it is still humans, with all our biases and prejudices, who are responsible for it.

Ensuring our humane future:

- How then do we ensure that AI applications are as unbiased, equitable, transparent, civil and inclusive as possible? How do we ensure that potential harm is mitigated, particularly for the most vulnerable, including for children?
- Without ethical guard rails, AI will widen social and economic schisms, amplifying any innate biases at an irreversible scale and rate and lead to discriminatory outcomes.
- It is neither enough nor is it fair to expect AI tech companies to solve all these challenges through self-regulation.
- First, they are not alone in developing and deploying AI; governments also do so.
- Second, only a "whole of society" approach to AI governance will enable us to develop broad-based ethical principles, cultures and codes of conduct.
- It is shared responsibility to ensure the needed harm-mitigating measures, reviews and audits during design, to inculcate the transparency, accountability, inclusion and societal trust for AI to flourish and bring about the extraordinary breakthroughs it promises.

AI and India:

- Many countries, including India, are cognisant of the opportunities and the risks, and are striving to strike the right balance between AI promotion and AI governance.
- NITI Aayog's Responsible AI for All strategy, the culmination of a year-long consultative process, is a case in point.
- It recognises that our digital future cannot be optimised for good without multi-stakeholder governance structures that ensure the dividends are fair, inclusive, and just.

Conclusion:

- Agreeing on common guiding principles is an important first step, but it is not the most challenging part.
- We must be prepared for deep, difficult, multi-stakeholder ethical reflection, analyses and resolve. Only then will AI provide humanity its full promise.