THE RISE OF ARTIFICIAL INTELLIGENCE POSES UNPRECEDENTED CHALLENGES TO HUMAN RIGHTS AND REQUIRES A GLOBAL FRAMEWORK FOR ETHICAL AI GOVERNANCE

Natasha Rama Rocha[*]

ABSTRACT

Artificial Intelligence is the most powerful and prevalent form of technology, and with such innovations come many benefits and threats. As the AI system benefits citizens around the globe, there remain many ethical questions about the intrusion of AI into every aspect of our private and professional lives. The advancement of AI has forced humans to forecast their future where jobs will be automated and lead to a huge amount of unemployment. Here AI is not just a player but the hegemon that had made us sub-servants. This AI will be controlled by the hands of people who will operate the proliferation of it and squeeze out everyday livelihood from the marginalised community. There will be no incentives to create employment and welfare for the people because the backbone of the economy is tied to an entity that does not require a salary and health insurance, this itself closes doors to employment which will lead to an increase in suicidal rate and hamper our fundament rights mentioned in the Constitution of India. This paper explores the concerns of human rights that will be affected by AI and raises awareness of the unprecedented challenges and how the application and use of AI technologies impact individuals' fundamental rights and freedoms. What are the ethical and legislative challenges that arise when it comes to AI governance and regulation? How can we ensure that AI is used for good and not Evil?

Keywords: Artificial Intelligence, Human Rights, Governance.

INTRODUCTION

We are living in a world where Artificial Intelligence is worn as a halo on the heads of every mankind and such advancements come with threats to human rights. The most demanding question here is, how does the development of deploying an Artificial Intelligence system affect the fundamental rights and freedom of individuals and society?

[*]BBA LLB, SECOND YEAR, KARNATAKA STATE LAW UNIVERSITY, ISBR LAW COLLEGE BENGALURU.

Artificial Intelligence or AI, is a technology built up using science and engineering with a mechanism that enables computers and machines to vitalize human cognitive works, intelligence and social problem-solving capacities of human beings. This innovative blend of computer and technological systems makes human functions easier and faster. AI is very versatile and quite accessible to the world from using this technology in-house to humongous IT companies and other organizations.

Alan Turing, the founding father of AI defines AI as the science and engineering of making intelligent machines especially intelligent computer programs*"*[1]

The groundwork for AI started somewhere in the early 1900s-1950's when the media created explicit attention to the idea of artificial humans and artificial brains which is now termed as Robots by society. In 1950-1956 AI was Born and came up as a head after Alan Turning published "Computer Machinery and Intelligence" which proposed a test of machine intelligence called The Imitation Game, the Turing test is even used to date. Later in 1952 Arthur Samuel a computer scientist developed a programme and John McCarthy in 1955 held a workshop on AI at Dartmouth which popularised the concept of AI throughout the world. Later on, various events took place that grew the recognition of AI in the world, for example in 1997 IBM's IA Deep Blue successfully defeated the world chess champion, Gary Kasparov. Google introduced a voice identification system in its Android phone in 2008 and in 2016 Google Deep Mind's Alpha Go system won over Go's champion Lee Sedol, etc. in today's world most of the applications work on the built-up AI systems taking from OpenAI, ChatGPT, Gemini and more; the electronics gadgets are already updated with AI software such as Copilot.[2]

ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS

Human Rights are fundamental and universal rights that are inalienable and belong to every human being, notwithstanding their race, religion, gender, status or nationality. They possess political and civil rights, such as rights to life and liberty, rights to speech and expression, rights to participation and rights to privacy. It also includes rights of social, cultural and economic aspects. Human rights being a global agenda, these rights are encoded in various international instrumentalities' which include the Universal Declaration of Human Rights

---

[1] Bhagyashikha Saptarshi, Artificial Intelligence and its Legal Ramifications, [2024], (Information Technology/ Cyber Law), (pg. no. 1)

[2] Ibid.

(UDHR), the International Covenant on Economic and Cultural Rights (ICESCR) and the International Covenant on Civil and Political Rights (ICCPR) exclusively. Late 2022 and early 2023 have sparked up the growth of AI making it accessible to every individual, no doubt the system provides numerous benefits but also possesses a few threats that are invisible as of yet.

Artificial Intelligence can sabotage human rights by creating challenges and risks that may threaten National security, surveillance, discrimination, violation and manipulation by alternating the capabilities and using them for malicious actions. AI can also expedite the unequal distribution of economies by centralizing market powers in the hands of a few hegemons and making others sub-servants by superseding jobs and producing undesirable conditions through extortive data practices, introducing alternative biased narratives and harmful environmental determinants. For example, AI can enable mass and intrusive surveillance, by collecting, analysing and sharing the vast amount of personal and sensitive data, without the consent or knowledge of the individual's concerns.[3] AI can also be used to manipulate and influence individuals and groups by generating and disseminating false or misleading information such as fake news, fake characters that can influence their perception and behaviours and Deepfake is one of the greatest challenges, it's an enormous menace to individual reputation and privacy.

AI can be used to encourage violence and harm by reproducing and amplifying existing assumptions in data and assumptions in algorithms, which can result in unfair and harmful outcomes for certain individuals and groups. It can also be used to encourage violence and harm by facilitating the production and use of lethal autonomous weapons, which can target and kill humans without human intervention or accountability.

ETHICAL PRINCIPLES AND EMERGING GOVERNANCE TECHNOLOGY

The 2017 release of the Development plan for the new generation of Artificial Intelligence, stresses on dual technical and social attributes of AI which must carefully be managed to ensure that AI is trustworthy and reliable. Ministry of Science and Technology of the People's Republic of China in 2019, established a National Governance Committee for the governance of the new generation of AI and Released the *Governance principle for the new*

---

[3] Bostrom and others, The Ethics of Artificial Intelligence, [2011], K. Frankish & W.M. Ramsey, The Cambridge Handbook of Artificial Intelligence (pg. no. 316-334).

generation artificial intelligence responsible artificial intelligence,[4] and more, these identified specific keywords:

1. Security and privacy: AI systems should be secure and should respect privacy.

2. Safety and reliability: AI systems should perform reliably and safely.

3. Transparency: The AI system should be understandable.

4. Accountability: The AI system should have accountability.

5. Fairness: The AI system treats all people fairly.

The security and privacy principal inclusion are of utmost importance, data security is the most common and basic requirement of ethical AI governance. Many countries and governments are establishing legislations and acts to protect the data and provide security and privacy for not just individuals but entities, governments, etc. Recently in 2023, the Government of India introduced the new Digital Personal Data Protection Act (DPDP Act) 2023, the United Nations (UN) also declared a humanistic attitude towards AI by promoting AI ethics at the United Nations Education, Science and Cultural Organization (UNESCO), the European Unions in 2018, also enforced the General Data Protection Regulation (GDPR) and China kin 2017, ratified the Cybersecurity Law of the People's Republic of China.[5] This initiation of such extensive regulations aims to protect the personal privacy of the users and face new challenges that may arise and challenge the AI-driven development that is commonly introduced today.

AI components when worked in uncertain and dynamic environments which are embedded in today's autonomous systems, inevitably increase the challenges of reliability, transparency and trustworthiness. Because as being state of the art, the AI system employs complex DNS and end-to-end training schemes, it acts as a black box which not only hinders the developers from completely understanding its structure and behaviours but also introduces implicit and potentially harmful inputs to the malicious models. Therefore, AI governance must involve multiple approaches that will enable AI engineers to practice a well-systematically developed analysis that will increase the public's confidence in the AI

---

[4] Zeng Y, Lu E and Huangfu C., Linking Artificial Intelligence Principles, [2018], arXiv.1812.04814.
[5] Wenjun Wu, Tiejun Huang, Ke Gong, Ethical Principles and Governance Technology Development of AI in China, [2020], Engineering, (Volume 6, Issue 4).

system. AI can also affect safety and act in robustness for taking into consideration Adversarial instances. Recently adversarial examples have become a very popular topic in machine learning for DNNs due to their vulnerability, in which the DNNs can be misled by adding inputs with insignificant imperceptible perturbation which results in improper and incorrect results. For example, the social watchdogs that are the hackers can meticulously add spiteful changes of small magnitude to any image of street crossing with pedestrians on the highway, resulting in confrontational situations that can tempter DNNs into dismissing the pedestrians on the scene, because of this the significant limitation of practice deep learning tools such as automated driving and facial recognition systems, adversarial examples, etc can result in fatal accidents, financial losses, emotional damage, and global peace imbalance.

The Transparency and critical accountability of AI are very important and critical for the public in underrating and making decisions for formal applications like, bank loan application management, medical diagnosis reports or law enforcement, etc it helps the general public to better understand the medical, legal or banking terms with basic easy understanding and with improved and innovative AI being transparent, people are more comfortable to use these AI tools to assess the legitimacy and accountability of sovereign autonomous systems and make decision with referring AI prospects. These AI interpretations aid in deciphering the intricate inner workings of deep learning of the algorithms and proving human-understanding explanations of reasoning and inferences based on the cognitive ability of understanding. This transparency of artificial intelligence has opened a new path of misuse, the most leady and prominent subject matter Deepfake.

PROBLEMS OR EXPLOITATION OF ARTIFICIAL INTELLIGENCE

As Artificial Intelligence slowly advances every day, the voices warn of the dangers that artificial intelligence can cause by its rise, the alarms become more intense and many people have come into public to warn the global economy of its repercussions. Geoffrey Hinton, *"Father of AI"* who worked on machine learning and neural network algorithms said, *"These things could get more intelligent than us and could decide to take over, and we need to worry now about how we prevent that happening"*. In 2023 he even left his position at Google, to talk about the dangers of AI. In the same year Elon Musk, the founder of SpaceX and Tesla, together with 1000 other tech leaders opened a letter to put a pause on large AI experiments, citing that the technology can "pose a profound risk to the society and

humanity".[6]

We're slowly moving towards a boring dystopia and biases, Unease exists on a variety of fronts, including the increasing automation of certain occupations, gender and racial-biased algorithms or autonomous weapons that operate without human supervision and oversights, we are still in the early stages of what AI can do. Here are some repercussions associated with the exploitation of AI and what this technology can do unless supervised and governed under a government with the proper framework.

Artificial Intelligence in the wrong hands is definitely an undoubtingly powerful weapon. How likely will we have a job in the future? Are we safe on social media? Will I get hacked? How confidential are my bank details? These questions are not just doubts but a few are happening in recent times. Some of the major examples are listed below.

Unemployment due to AI Automation: job losses due to AI automation are a pressing concern. The Hi-Tech global industries such as healthcare sectors, marketing and manufacturing industries have started adopting artificial intelligence in their daily activities. AI will create almost 97 million jobs mostly in the upcoming years but, there will be no specialized persons to take up the roles, unskilled employees will not be able to operate such sophisticated technology due to lack of specialized education and not many employees will be required for the operating AI unlike other labour job posts. AI robots have entered into medical sectors and manufacturing units, majority of the tasks are not handled by a single individual who only operates the AI machines and robots.

In the medical sector, the synergy of AI robots has revolutionised this sector. AI can be used to help doctors and medical officers to assess the patient accurately using their data and other information, and provide diagnosis and treatment plans for the same. In the recent COVID-19 emergency, many hospitals worldwide switched to AI robotic procedures, that is to perform functions like disinfecting and screening patients and employees at the entry, these have shown success in AI robotics use in healthcare sectors. However, AI has its challenges with human doctors who can make quicker decisions in complicated situations, whereas AI robots will only perform procedures that are updated in their software.

In manufacturing sectors, companies like Amazon use robots to move items back and forth,

---

[6] Mike Thomas, 12 Risks and Dangers of AI, [2024], Builtin Report.

pick and pack orders, operate 3D printers, etc. Cobots or collaborative robots are commonly used in warehouses and manufacturing plants to lift heavy car parts or handle assembly, these bots are capable of learning tasks, avoiding physical obstacles and also work side-by-side with humans. The robots save time by running them unsupervised overnight, whereas else humans need time to process and perform the act.[7] Many other organizations and companies use AI to ease their workload and gain rapid progress, this very act has left many hard labourers and daily wage earners unemployed.

Social Manipulation due to AI: social media has become a threat today, posting a picture or video on an online social platform can cause any person to fall under the trap of Deepfake. In light of AI algorithms, AI-generated images, videos and even voice alterations can be generated. For example, the Rashmika Mandana Deepfake Case.[8] Here a 24-year-old male, Eemani Naveen was arrested from Andhra Pradesh, who created a deep fake video of Rashmika Mandana to increase this fan page account followers.

On the other hand, taking into consideration TikTok a platform that runs on AI algorithms, has also led to raising concerns over its ability to protect its users from misleading information. For example, Ferdinand Marcos, Jr., wielded a TikTok troll army to capture younger Filipinos votes during the Philippines 2022 election.[9] Rachel a famous TikToker also faced harassment after someone shared deep faked nudes of her on social media after which she suffered serious mental agony and distress.

Autonomous Weapons Created by AI: Technological advancement has controlled and uplifted the welfare of society. When it comes to AI some are just very eager to use this in different forms and explore it for other uses. The most exquisite car, the Tesla functions with AI-powered lane keeping and autosteer features, there are possibilities that hackers can manipulate this feature and use the automobile for evil such as accidents or murders. For example, in March 2020. Amat Cama and Richard Zhu of a team called 'Fluoroacetate' won a Tesla Model 3 and $35,000 for hacking into its system. They targeted the infotainment system on the Tesla Model 3 and also demonstrated how they could trick a Tesla Model S to enter into the wrong lane by using an 'Adversarial Attack' a way of manipulating a

---

[7] Jeremy Bowman, How Artificial Intelligence is Used in Manufacturing, [2023], Stock Market Statistics.

[8] Economics Times Rashmika Mandana Deepfake Case: Police Arrests Mastermind from Andhra Pradesh's Guntur, [2024], The Economic Times online News Paper.

[9] Mike Thomas, 12 Risks and Dangers of AI, [2024], Builtin Report.

machine learning model.[10]

The formation of a Lethal Autonomous weapons system can initiate rival countries, political rivals and others to start the tech cold war, they can build a huge military that can be programmed to fight and locate the opposite country movements. It will lead to a fight against equals, become a threat to civilians on the ground and even cause a cyber attack that is beyond the control of humans, causing a dystopia.

Satellites Outburst: Satellites play a pivotal role in our everyday lives, they contribute to our lively activities and well-being, and through their use, it makes it possible for us to communicate and meet the important needs on this Earth. Satellites use AI to capture photographs and transmit the images to the ground station and also to avoid satellite collision.

Imagine a hypothetical situation where hackers used Dark AI or made new AI algorithms that can hack into the satellites for example take the SpaceX satellite, the cars associated with this satellite that the Tesla cars can easily be manipulated causing havoc and even the death of people on road by either self-destruction or causing numerous accidents. There are high possible changes to using AI to hack into satellites and cause the biggest cyber-attack by spoofing and shutting down access to others which will cause a catastrophic effect making the world an enormous cyber-attack.

There were times when hackers attacked satellites for example, a serious attack was viewed in 2008, when hackers took full control of NASA Terra EOS Earth Observation satellites and two U.S. government satellites were hacked by suspected Chinese military groups in 2011.[11] If such attacks continue, these acts can jeopardise our career and personal lives, such as stopping the network signals, and electricity supply, and also quash down our basic rights to livelihood and causing the collision of satellites which can cause geostorms as well. These following incidents can be prevented only by global AI governance.

ETHICAL AI GOVERNANCE

Today's human rights violation is the cause of tomorrow's conflicts. The significance of AI

---

[10] IANS, Tesla Hacking Competition Offers $1 Million and Free Car if Someone can Hijack Model 3, [2020], Mint News.

[11] Brett Tingley, these 3 Teams just Hacked a US Air Force Satellite in Space… and Won Big Cash Prize, [2023], Space News.

ethical principles has been recognised across the world by academics, industries and even the government which have worked hard to develop AI governance technology. The framework by the government must realise the full potential development of AI and mitigate the risk of discrimination and misinformation as well as impact the environment and human rights. The proper governance of AI will promote equitable access to participate in department and knowledge of AI, share benefits and ethical guidance through the life cycle of AI system. The following actions can be followed for proper governance and to push forward the present initiatives on AI governance:

- To reach a consensus on AI ethical principles, governments, institutes and corporations should conduct cross-disciplinary, cross-sector and multinational collaborations.

- We must actively promote ethical education for every participant in AI research and development, application and management, to raise awareness of ethics and promote general principles of ethical conduct with AI.

- Open AI develops platforms with built-in ethics-relevant features that should be developed to assist all the stakeholders of different AI systems in evaluating their functional and regulatory compliance.

- In order to keep up with the rapid advancement of AI, they must intensify collaborative research and development of AI governance techniques.

- To enable experts from different disciplines to work collaboratively to solve the ethical challenges of AI, clearly define AI moral scenarios with a large social impact that should be identified.

CONCLUSION

Artificial Intelligence is a powerful and ubiquitous technology that can change every facet of human life, including human rights, democracy and governance and the adoption of AI signals a potential fundamental shift in our society. AI can have both positive and negative effects on individual's and societies' fundamental rights, freedoms and their functioning and qualities. This change can be a great opportunity to create a human community with a common future and promote a sustainable development of the natural environment and society. AI governance and regulation face many obstacles, challenges and opportunities, which require multiple disciplinary prospects as well as a multiple-stakeholder approach. However, their implications could be both acute and unprecedented, if there is no regulatory

system in place.

We also need to develop a robust and practical AI governance framework to regulate the profound development of artificial intelligence all according to human integration of values and ethics in order to ensure that these challenges are favourable and useful before they become fully embedded into the fabric of our daily lives. We can make AI accountable and effective, as well as build the public's confidence in AI products and services. AI governance and regulations must also address diverse needs and these needs include, oversight and support, flexibility and experimentation, integration, coordination, compatibility, diversity and contextuality, consistency, innovation and protection control, promotion and improvement, empowerment and facilitation.

AI is a versatile and widespread technology that can be used for good or evil and we have the freedom to make it one.