

Inversion of power between human beings and AI: are we heading towards losing our fundamental rights?

Muhammad Sohail Asghar

Assistant Professor of Law, University of Okara, Punjab, Pakistan

Kashif Mahmood Saqib

Assistant Professor of Law, University of Okara, Punjab, Pakistan

Hamid Mukhtar

Assistant Professor of Law, University of Okara, Punjab, Pakistan

Hafsa Naz

Visiting Lecturer (Corresponding Author), University of Okara, Punjab, Pakistan

Abstract

Recent times have witnessed a drastic expansion in the utilization of technology. The extensive use of sophisticated algorithms for decision making in high stake domains has ignited the realization that machines, like humans, must conform with the law. The existing liaison between the human beings and Artificial Intelligence (AI) is at the verge of radical transformation. Until recently, the human capability to dominate technology was unquestioned but the contemporary trend is indicative of a fading power differential and the likelihood of an inverse power relationship in the near future. This paper discusses the influence that the inversion of power between human beings and their technologies has regarding protection of fundamental human rights.

Keywords: Fundamental Rights, Human Rights, Artificial Intelligence, Machine Learning, Robotics

Article History:

Received: 21st Apr. 2022

Accepted: 17th May. 2022

Published: 24th Jun. 2022

1. Introduction

A set of instruction designed in a sequential order to perform a specific task is referred as algorithm (TechTarget, 2020). Algorithms are everywhere, whether that be a photo recognition software or a simple automated soap dispenser (Dickson, 2018). Use of these algorithms to create more complex machine learning systems falls under the ambit of Artificial Intelligence (AI) (Kavlakoglu, 2022). The use of algorithms is inevitable for the further technological advancements (Padamkar, 2020) and to create such machines that possess human like knowledge and consequently are capable to execute such tasks that are generally performed by humans (Schroer, 2020). The ultimate drive fueling the creation of Artificial Intelligence was our determination to understand the thought process of humans and create such machines that could mimic this pattern (Shubhendu & Vijay , 2013). To this end, large amount of sample data is processed through intelligent algorithms. This process enables a machine learning system to learn from the patterns within the data and finally apply human-like skills such as perception and reasoning etc. Artificial Intelligence is a very wide term, normally it includes advanced algorithmic mechanism such as machine learning and deep learning, that finally fulfil the ultimate objective of Artificial Intelligence (Kavlakoglu, 2022). Machine learning algorithms are quite different as compared to the traditional algorithms, these are not finite in nature, meaning thereby these algorithms are programmed to train a system how to complete a particular task by analyzing the data which will be provided over a period of time. Since the data could change over the time, the performance of the algorithms changes as well, making them adaptive in nature (Wakefield, 2020). The functional capability of machine learning resembles like human brain as it continues to learn by experience (Oberoi, 2017). Machine learning algorithms are the foundation of deep learning. Deep learning is a more advanced form of Artificial Intelligence, it mimics the thought process of human beings. In order to create a deep learning system that mimics human brain, deep learning algorithms are created in a layered form that are referred as Artificial Neural Networks (ANNs). As ANNs are structured to replicate the functioning of human brain, their capabilities are much advanced as compared to the machine learning algorithms. Unlike learning

from the data to form improved decision over a span of time, ANNs are capable to initially learn and make decision on their own without persistent analysis and collection of the new data (Grossfeld, 2020).

Although contemporary Artificial Intelligence (AI) has not fully evolved yet, still it has surpassed human experts in a number of activities and is likely to outpace human functioning in many fields in the near future. The existing power relationship subsisting between the human beings and their technologies is on the verge of a major change. The impact of AI is expanding and its scope is increasing. It is time for us to devise such strategies which are capable to take into account the prospect of power inversion between human beings and Artificial Intelligence in order to safeguard human society against technological, rather than military, economic or political power (Schmelzer, 2019). This paper investigates the impact that the reversion of power between humans and technologies has regarding protection of fundamental rights. It is pertinent to mention here that the use of Artificial Intelligence has already started to create challenges for fundamental human rights protections, failure in the recognition of these snags will leave the algorithmic determinism unquestioned. There are three fundamental hurdles within current human rights mechanism that bar our way to develop a human rights regime against Artificial Intelligent technologies (Liu & Zawieska, 2017). These fundamental shortcomings need to be redressed before widespread adoption of AI technologies in our society so the upcoming technological threats to human rights may be avoided. The confronts that human rights protections are actually facing due to the advancement in AI are profoundly distinctive and serious in nature as compared to those human rights violations that could be seen through orthodox legal lens. This situation calls for an alternate approach to be used, based upon the power inversion between humans and Artificial Intelligence, to ensure and enhance human rights protections regarding and through technology by moving forward.

2. Objections Attached with The Proposal to Re-Orientate Human Rights Regime

The proposal to restructure human rights paradigm against human rights violation caused by technological advancements will be confronted by an initial objection that the state alongwith its institutions carry all the powers and the capabilities bestowed by the modern AI technologies have merely enhanced collective efficiency of humans. The real connotation of the state, therefore, is that the re-orientation of fundamental human rights is not the need of the hour as the existing human rights mechanism is sufficient to be deployed against any human organizations existing at the core of any human right violation. This objection could be satisfied in two ways: firstly, the motivation towards the development of more human centric system of human rights, as compared to technology, does not imply a relegation of current human rights guarantees against the state. Nevertheless, the contemporary human rights regime is no more than an array of “Swiss cheese” barrier in response to the occurrences of human rights violations (Reason, 2000). Secondly, our articulation of a re-orientated human right spectrum against technology is to make sure that fundamental human rights are well protected against such powerful threats which humanity has never encountered before.

The abovementioned hindrance is not the only one in the way of introducing a more developed human rights regime. There are many other obstacles as well which need to be redressed before making a clear path for a more effective human rights protections against the technology. For example, compartmentalization of the concerns with reference to the kind of the impugned right (Kennedy, 2005). This kind of classification leads towards shattered or fragmented understanding regarding true nature of the problem concerned as a whole. The fact that a large structural change is happening is not so obvious as an incomplete sketch is being painted. This kind of shortened understanding is the core reason behind the ineffective human rights protections against the technology. Another problem with the traditional human rights methodologies is that, these are effective enough against certain types of harms caused by the state or any of it's organs within its territorial jurisdiction. However, the efficiency of this mechanism is compromised if anything falls outside this formula (Veitch, 2007) making an extremely powerful entity i.e., Artificial Intelligence excluded from the ambit of responsibility and review. The last problem with contemporary human rights regime is its monopolizing nature which tends to roll out all the other relevant perspectives on the issue (Kennedy, 2005). The problem with this hegemonizing pressure is that we cannot successfully defend any human rights claim unless it is embeded within its existing logic and wording. This problem seriously restricts any possibility of deviation from the contemporary human rights constellation, despite the fact that the fundamental characteristics of this model are intrinsically problematic from its very core.

3. Contemporary Human Rights Regime and AI

There are three inherent problems attached with the contemporary human rights structure restricting its efficiency against technology. Firstly, though substantive rights have evolved for and against state power with the advancement of experimental theory of human rights (Dershowitz, 2009) but the concerns regarding technological wrongs were not developed in the same pace, making contemporary system of human rights ineffective against AI borne challenges pertaining to fundamental rights. Secondly, our traditional human rights domain is so mixed up within the existing state law that it has shortened the facility to deal with the power dynamics which has capacity to potentially affect human beings and their very individuality (Nollkaemper, Van Der, & Harmen, 2009). Finally, our legal system is inherently flawed when it comes to origin of harm, it raises the third problem i.e., if the origin of harm is distributed and its effects are indirect or peripheral the existing framework is unable to recognize it (Isaacs, Vernon, & Richard, 2011). As the legal consequences of AI are not direct or immediate and could only be seen through a widened standpoint, these could not be recognized as a serious threat to human beings under the current legal mechanism.

(con)What we need at the moment, a feasible and balanced human rights regime to keep a check on the increasing technological power to ensure the continuity and protection of human rights. AI entities deeply and critically examine such values and characteristics that make us human, steps to strengthen these core characteristics must be taken to preserve and protect human beings in the face of technological challenges. A more convergent human rights mechanism is needed to specifically address the technology borne risks, for example AI and robotics in this case, which could be asserted in case of responsibility gap (Matthias, 2004). Building such an assistive mechanism will not only balance the responsibilities rather it would also calibrate the capacities: fixing responsibility only on humans in case a robotic risk (Liu & Kress, 2006), or putting human beings in a position where human has to take full responsibility for the malfunction of a larger system (Elish & Clare, 2016).

4. Relationality, Responsibility and Control

Our limited possibility to control or regulate AI is the fundamental reason that why people are afraid of such technologies when confronted. An individual user is afraid as he has a very limited knowledge that how the system is developed and how it actually operates. Limited access of external intervention and increased system autonomy is another point of concern. Setting up a few system preferences only gives us an illusion of the control rather than the actual control over the system. The threatening concern from a developer's perspective is that, when a system becomes extremely intelligent and capable of learning and decision making, no one can fully control its conduct and interlinked consequences. It is such a socio-cultural environment where neither users nor developers are in position to fully control the risks attached with the emergence of modern AI technologies. Nevertheless, we more than ever rely and claim on control and security (Beck, 2006). We, on the one hand, are inclined towards the creation of more independent machines as we believe that these are more efficient than human in certain circumstances, in warfare for instance (Arkin & Ronald, 2009), but on the other hand our lack of control is exactly the reason of concern. Control and liability issues are interconnected, a person could be held liable for something only if he had actual control over it. Predictability carries high importance in law while deciding accountability, foreseeing the possible outcomes and risks attached with the AI based technology is extraordinary difficult consequently making "a responsibility gap" (Marino, Tamburrini, & Guglielmo, 2006). The fixation of responsibility is not a mere application of rules (Becker, Howard, & Michal, 2009) rather a matter of discussion among all the participating actors who had sufficient authority and autonomy to be the contributors and act willfully (Walton & Marsha, 1985). Stated differently, responsibility is not such a characteristic which could be imposed, but an ability of acceptance (Dworkin, 2011). Liability is a process, not an attribute, that is the very reason that we "learn" to be responsible in the same way as we learn to respond or interact socially. That is why, the concept of responsibility is far more complex than the mere control while dealing with advanced artificial intelligent technologies. We have to reorient our control-based concept of responsibility as it goes far beyond than that, it must be conceptualized as a rationale process (Liu & Zawieska, 2017). This primal transformation from static to rationale and dynamic notion of control & responsibility may serve as founding principle which is necessary to reorient ossified human rights mechanism into a modern human right regime. Although this rationale-based concept is a challenging task to interpret into computer scientists and

engineering terms, yet it provides a foundation to explain the novel spheres of responsibility and combat the challenges posed by modern AI technologies to ensure the continuing protection of human rights in future.

5. Conclusion:

This article has enumerated the insufficiency of the existing human rights mechanism to ensure the continuity and protection of fundamental human rights to combat the challenges posed by the advent of the modern AI entities. Considering the prevalent scenario, it is the dire need of the hour to develop a reoriented human rights regime, based upon rationale dimensions rather than the on-going control-based concept of the responsibility. If we aim to avoid catastrophic destruction of contemporary human rights constellation, human rights must be developed in a manner to safeguard human against the threats of modern AI technologies by strengthening and preserving the core concept of the human beings alongwith ultimate human values. How to achieve such a goal, remains an open question and the aim of this paper is to try to commence such discussions.

References

1. Arkin, Ronald, R., (2009) *Governing Lethal Behavior in Autonomous Robots*, CRC Press.
2. Albright, J. (2016) Algorithms might be everywhere, but like us, they're deeply flawed, <https://theconversation.com/algorithms-might-be-everywhere-but-like-us-theyre-deeply-flawed-66838>
3. Beck, U. (2006) *Living in the world risk society*, Hobhouse Memorial Public Lecture given on Wednesday 15 February 2006 at the London School of Economics, *Economy and Society*, 2006, 35(3), pp. 329–345
4. Becker, Howard, S., Michal, M., *Symbolic Interaction and Cultural Studies*, University of Chicago Press, 2009
5. Elish., Clare, M. (2016) *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction*, WeRobot 2016 Working Paper, University of Miami
6. Dershowitz, A. (2009) *Rights from Wrongs: A Secular Theory of the Origins of Rights*, Basic Books.
7. Dickson, B. (2018) What is algorithmic bias? <https://bdechtalks.com/2018/03/26/racist-sexist-ai-deep-learning-algorithms/>.
8. Dworkin, R. (2011), *Justice for Hedgehogs*, Harvard University Press
9. Grossfeld, B. (2020) *Deep learning vs. machine learning: a simple way to understand the difference* <https://www.zendesk.com/blog/machine-learning-and-deep-learning/>
10. Isaacs, Vernon, T., Richard, *Accountability for Collective Wrongdoing*, Cambridge University Press, 2011
11. Kennedy, D. (2005) *The Dark Sides of Virtue: Reassessing International Humanitarianism*, Princeton University Press
12. Kavlakoglu, E. (2020) *AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the difference?* <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.
13. Liu, R. & Kress, H. (2006) *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University Press, 2006, pp. 325–344
14. Liu, H. & Zawieska, K. (2017) *A New Human Rights Regime To Address Robotics And Artificial Intelligence* https://www.researchgate.net/publication/322043569_A_new_human_rights_regime_to_address_robotics_and_artificial_intelligence
15. Matthias, A. (2004) *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, *Ethics and Information Technology*, 2004, 6(3), pp. 175–183.
16. Marino, Tamburrini, D., Guglielmo, *Learning robots and human responsibility*, *International Review of Information Ethics*, 2006, 6(12), pp. 46–51
17. Nollkaemper, Van Der, A. W., Harmen (eds.), *System Criminality in International Law*, Cambridge University Press, 2009
18. Oberoi, A. (2017) *9 Machine Learning Examples from Day-to-Day Life* <https://insights.daffodilsw.com/blog/9-machine-learning-examples-from-day-to-day-life>.
19. Pedamkar, P. (2020) *Introduction to Algorithm in Programming*, EDUCBA,

- <https://www.educba.com/algorithm-in-programming/>
20. Reason, J. (2000) Human error: models and management, *British Medical Journal*, 320(7237), pp. 768–770
 21. Rooij, S. D. (2018) What are the major characteristic of an algorithm? <https://www.quora.com/What-is-an-algorithm-and-what-are-its-steps>
 22. Shubhendu S., Vijay, J. (2013) Applicability of Artificial Intelligence in Different Fields of Life, *International Journal of Scientific Engineering and Research* Vol. 1 Issue 1 Pp. 28-35 (explaining that John McCarthy, a former professor Massachusetts Institute of Technology, coined the phrase “artificial intelligence”).
 23. Schmelzer, R. (2019) Should We Be Afraid of AI? <https://www.forbes.com/sites/cognitiveworld/2019/10/31/should-we-be-afraid-of-ai/?sh=144896e14331>
 24. Schroer, A. (2020) What is Artificial Intelligence, <https://builtin.com/artificial-intelligence>
 25. TechTarget (2020) Algorithm, <https://whatis.techtarget.com/> (choose “AppDev” under “Browse Definitions By Topic;” then choose “Programming;” then click “algorithm.”)
 26. Veitch, S. (2007), *Law and Irresponsibility: On the Legitimation of Human Suffering*, Routledge Cavendish, Oxford.
 27. Wakefield, K. (2020) A guide to the types of machine learning algorithms and their applications. https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html#:~:text=At%20its%20most
 28. Walton, Marsha D. (1985), Negotiation of responsibility: Judgments of blameworthiness in a natural setting, *Developmental Psychology*, 21(4), pp. 725–736