DEMYSTIFYING THE COMPATIBILITY OF ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS

By Smruti Poola
From Symbiosis Law School, Hyderabad

ABSTRACT

Technological advances offer a plethora of benefits for humankind. These benefits should be embraced with gratitude and vigilance. Human thoughts, beliefs, and ideas are constantly changing to wholly fit the landscape of the modern era. The shift in people’s attitude and lifestyle also dictates a shift in the rights that humans have agreed to. The existing international legislations on human rights scarcely contemplate the effects of modern inventions, especially the instant integration and development of Artificial Intelligence (‘AI’). The correspondence of AI and human rights is undeniably the foundation for a sound future; however, certain discrepancies are found in the ethics of AI. Concerns around privacy, racist experiences, and instances of sexism are commonly reported in conjunction with AI which preview the possibility of harmful AI that jeopardizes human rights. International authorities and local governments find themselves in an imbroglio with rapid integration and rise of AI without adequate acknowledgement towards its risks and human rights impact. Though, efforts to introduce and promote ethical AI have been introduced and discussed widely by multiple authorities, there still exists a lot of ambiguities concerning AI’s impact on human rights. The author will discuss the various ethical concerns of AI and the present regulatory framework, across the world, to combat the dangers of AI and harness its potential, simultaneously.

Keywords: Artificial Intelligence, Human Rights, Regulation, Ethical Concerns, Risks

INTRODUCTION TO HUMAN RIGHTS AND AI

Human rights law embodies a spirit of philanthropic intentions that act as a parachute for guaranteeing basic rights that ensure the liberty, freedom, and dignity of all human beings. The evolutionary history of human rights as a political force and ideology has allowed the development of suitable legal mechanisms to strengthen its objective.¹ The violation of human rights attracts rigid penalty on an international level and more notably, leads to intense criticism. However, the enforcement of human rights assumes a back-seat when influenced by economic or political agendas.

Human rights are vested in the principle of universality. Towards the end of the World War II, there was an urgent need to secure basic rights of people around the world. Universal Declaration of Human Rights (1948) was the foundational force that solidified the enforcement of human rights and continues to remain so.² Human rights assured individuals protection from arbitrary

decisions of sovereign authorities. The United Nation Declaration on Human Rights (UDHR) along with a series of treaties and conventions such as the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR) enacted by the United Nations, in respect of the premier declaration are recognized as the International Bill of Human Rights. These binding guidelines secure every person’s basic rights irrespective of their geographical location. Development and human rights are intertwined on various levels. The advancement of technology and development in resources promote human rights, either in design or in outcome. However, these concepts may not always overlap especially when the advance measures or items turn a blind eye to human rights protocols. To understand the implications of compromising human rights through introduction of developments, an analytic and clear view of the consequences must be considered. Substantiating human rights to a higher value in its principles would possibly facilitate their relationship.

The transient nature of affairs in the world allows the discovery of novel technologies that induce change in mankind’s system of operation. The historian, Melvin Kranzberg believes that “technology is neither good nor bad; nor is it neutral”, whereas, political theorist Langdon Winner opines that “certain technologies are inherently political”. Artificial Intelligence (‘AI’) is one such development in technology that has forced the world to adapt and enforce mechanisms to govern such a dynamic concept. AI is recognized as a branch of computer science that allow machines to possess or replicate human intelligence. AI systems possess super-intelligence and can be classified in two ways, the first indicating a system that relies on coded instructions to draw a logical conclusion; the second type pertains to self-driving systems that are constantly improving their decision-making talents, and are structured as “neural networks” which closely resemble human intelligence. The ever-growing presence and rapid integration of AI into daily work, has dragged the question of its governance into the limelight. AI raises some pertinent questions on its regulation, transparency and ethics that remain unattended to. The growing internet reliance on AI draws attention to the potential threat on human rights. The Institute of Electrical and Electronics Engineers (IEEE) which is recognized to be the largest organization of technical professionals published a report highlighting that the importance of respecting international human rights in order to ethically align AI’s design. This was

---

8 Institute of Electrical and Electronics Engineers, ‘Ethically Aligned Design’ (The IEE Global Initiative 2019) <https://standards.ieee.org/content/dam/ieee-
supported by the United Nations’ International Telecommunications Union during its 2018 Global Good Summit\(^9\) wherein it was concluded that fundamental human rights should not be harmed or infringed upon. The Guiding Principles on Business and Human Rights, 2011 released by the United Nations provides adequate framework for regulating businesses and developments to secure international human rights. Although not exhaustive, this document provides a starting point for the intersection of preservation of human rights and technology.

AI is a popularly discussed tool within the legal community as well, however, the effectivity of the same is still ambiguous.\(^10\) The risks associated with AI display a tendency to overpower the benefits. An individual’s digital presence, data, and internet activities share a vulnerable position in terms of their privacy and security that may be encroached upon through modes of AI.\(^11\)

AI lacks moral or ethical considerations in its implementation, the application of AI solely depends on face value, and is unaccounted to other human values. The protection of human rights might be diluted if AI isn’t regulated through an exhaustive and proper mechanism. The role of AI isn’t widely discussed or clarified in many national legislations which creates gaps between AI and other potentially effected laws.\(^12\)

Therefore, approaching AI with a perspective of human rights aids in the identification of its various ethical implications and offers certain measures to ensure the maximum reduction of such negative impacts.

**CONCERNS OF AI AND HUMAN RIGHTS**

AI complicates the whole discussion around personhood and human rights. This confusion was fueled when Saudi Arabia, became the first country to grant an AI robot named “Sophia” citizenship and thus, establishing a legal identity of her own.\(^13\) The continuance of this trend could potentially result in AI assuming a personhood of its own through its own experience and knowledge. The concerns around this would eventually equate the problems of AI with the flawed nature of human personhood, which would inevitably result in serious implications for human rights, for itself and human intelligence.\(^14\)

The gaps in AI, once legitimized through recognized legal standards/standards/web/documents/other/ead1e.pdf?utm_medium=PR&utm_source=Web&utm_campaign=EAD1e&utm_content=geias&utm_term=undefined> accessed 16 August 2020
\(^14\) Maggie Redden, ‘Sophia: The Intersection of Artificial Intelligence and Human Rights’ (2020) 10 Journal of Global Rights and Organizations Annual Review 155
personhood would attract various different subsets of law and enjoyment of law, which would invariably allot a pedestal to AI’s existence over humans. Even with the undeniable advantages, AI seems to encompass a large set of issues in terms of human rights that rests on the connection between morality and rationality.

**Equality, Dignity and Non-Discrimination**

The principle of equality serves to ensure that all classes are treated in the same manner and the principle of non-discrimination serves to ensure that there is no social or political bias against one demographic over another. These principles are embedded into various international treaties and conventions to create a protective framework against any arbitrary practice of the State. AI compromises human dignity by inadvertently supporting social bias and discrimination. The inherent discrimination perpetrates socio-economic gaps, racism and historical inequity. AI replicates or follows patterns of human decision making, therefore, determining the status quo and hence, allowing scope for social discrimination.

The debate on the social values of AI and its regulatory mechanisms surrounds the inherent issue of social bias and discriminatory practices through its operation or algorithm. AI systems, in the past, have been observed to follow a biased pattern in the detection and interpretation of skin colors. Moreover, the usage of AI in surveillance technology can be harmful to marginalized groups.

There have been certain speculations that the Chinese Government is in the process of developing an AI model that classifies and categorizes citizens on the basis of their social characteristics. The system could potentially target the poorer or marginalized sections of the country, resulting in the creation of an unfair divide. The unintentional but real discrimination in AI systems was also addressed by the World Economic Forum, in a recent publication, which stated that success of AI is determined through parameters based on profit which might threaten the operation of human rights.

The rise of use of facial recognition systems, especially in concerns of national security, could potentially add to the AI problem of equality. Faces with strong ethnic features generate vague results. In one instance, Google images classified pictures of “gorillas” as African-Americans. However,

---

the problem was only fixed by censoring search words and tags, instead of effectively developing or correcting the AI algorithm. The identification accuracy is largely favorable towards men with Caucasian features, whereas, detection of dark-skinned individuals or individuals with ethnic features, is inaccurate. The inaccuracy and gaps in the idea of facial recognition, especially pertaining to one section or group of the society leads to the creation of inherent biases that could prove to be detrimental in certain situations. The use of AI for predictive justice allows for the scope of uncertainty that could potentially lead to human rights violations.

The gender gap is also amplified through the application of AI. A handful of social media networking applications have been designed in a way that is more favorable towards men, in terms of proposing higher paying jobs, job advertisements and depiction of activities such as sports and housework. A popular illustration that displays bias in AI is the chatbot “Tay”, which was released by Microsoft in 2016 for conducting research through human interaction. However, the system adapted to the darker side of humanity and in a short while tweeted offensive content that undermined feminists and supported racism.

The larger question that is considered in terms of inherent algorithmic discrimination in AI application is the ethical construes that exist to protect such discrimination from occurring. Viewing the application of AI, through an ethical perspective can benefit the disadvantaged sectors by erasing prejudices through a more developed and defined model. The European Union has recognized the need and published the “Draft Ethics Guidelines for Trustworthy AI” that offered deep insights into the balance of advantages and threats of AI and suggested an anthropocentric approach to AI.

Privacy, Personal Data and Digital Misinformation

The growth in technology highlights the importance of cyber security. However, new hardware and software developments can ensue threats to cybersecurity by making data and privacy vulnerable to third party attacks. There have been multiple instances of data leaks through corruption, hacking or attacks. According to Metcalfe’s Law, the value given to a network multiplies greatly with the addition of new users. The internet is a vast

---

hub of data which is transmitted across all platforms. The application of AI also mimics the theoretical aspects of this law. This data is susceptible to cyber-attacks that includes data theft or exploitation and surveillance, which concerns the privacy rights of an individual.\textsuperscript{27} The information logged into the internet by users can be modified, studied and exploited by large corporations to further capitalistic agendas by studying the consumer profiles. This manifests itself into the user’s online activity with specifically tailored advertisements according to the data shared by them. Moreover, data or online activity, if unlawfully tracked and traced could be harnessed to extort an individual.

Privacy is on the pathway of being traded for monetary gains. If this trend does arise, it could prove to be detrimental to the social freedoms of individuals and can potentially push towards a totalitarian regime.\textsuperscript{28} Recently, Stanford University developed a form of AI algorithm that could accurately determine the sexual orientation of users through their internet activity, despite not having their consent.\textsuperscript{29} This could be exponentially harmful towards the individuals since identification of such personal information could land them into trouble in countries where the legal mechanisms contain laws against homosexuality. In the same vein, Amazon’s facial recognition software also raises serious privacy concerns that are linked to discriminatory inclinations.\textsuperscript{30} Additionally, the spread of digital misinformation breeds another cause of concern in AI application. Misinformation can be spread to fulfill political goals and challenge the integrity of cybersecurity. Information is recognized as a valuable asset, twisting such an asset to suit diplomatic or political agendas could incite violence or mistrust and rattle democratic mechanisms. The allegations on Russia for manipulating the 2016 US Presidential elections is an example of the use of digital technology to specifically identify, amplify and target persisting social, economic and political differences.\textsuperscript{31} Simultaneously, AI can be a gold-mine for governments which could allow the development of policies, security and specifically benefit the operation of intelligence agencies.

AI can be equipped to infer certain facts from the available data that is shared that could potentially expose political opinions, emotional states and personal information. The vulnerability of privacy rights could potentially result in the infringement of other connected and related rights such as right to freedom of speech and expression. The key factor to be considered to tackle such invasion is the discrepancy between individual human rights and algorithmic risks.\textsuperscript{32} Hence, the analysis and storage of data should be justified beforehand.

The legislative background to protection of privacy rights in technology is still in the

---

\textsuperscript{27} Nabile M. Safdar, John D. Banja, Carolyn C. Meltzer, ‘Ethical Considerations in Artificial Intelligence’ (2020) 122 European Journal of Research

\textsuperscript{28} Maria Stefania Cataleta (n 19)

\textsuperscript{29} Sam Levin, ‘New AI can Guess Whether You’re Gay or Straight From a Photograph’ (\textit{The Guardian}, September 08 2017) <https://www.theguardian.com/technology/2017/sep/0>

\textsuperscript{30} Mark Latonero (n 24)

\textsuperscript{31} Alex S. Wilner, ‘Cybersecurity and its Discontents: Artificial Intelligence, the Internet of Things and Digital Misinformation’ (2018) 73(2) Canada’s Journal of Global Policy Analysis 310

\textsuperscript{32} Maria Stefania Cataleta (n 19)
stage of infancy. The General Data Protection Regulation\(^{33}\) regulates and ensures certain limits are set on the level of data collection in Europe, thereby, imposing certain standards that are mandatory for companies to comply with while managing their data stock. However, the GDPR widens the regulatory gap between Europe and companies that have originated out of Europe. The main concerns arise on the question of priority, whether the consumers privacy rights are to be safeguarded or protection is to be given to industry agents. However, it is pertinent that a balance is struck to ensure that the benefits of AI are absorbed while granting protection to an individual’s privacy. Moreover, expectations on citizens to educate themselves of these harms and spread carefully in light of these issues would be mundane and burdensome.\(^{34}\) Hence, the requirement shifts from treatment of privacy as an ethical obligation to a fundamental right, which would benefit the legal and regulatory framework of technology developments.

**Freedom of Speech, Expression and Movement**

The right to freedom of speech and expression is an integral part of ensuring individuality and self-thought. Every person has an inherent power to be able to express their own opinions, reasons, consent and dissent. It is recognized as an integral human right under UDHR and ICCPR.\(^{35}\) The internet provides a wide outreach in matters of public discussion and hence, the freedom of speech and expression is undeniably connected with the internet and the content an individual wishes to publish.

The application of AI in regulating or monitoring such speech posted online could result in some serious damage especially in terms of content moderation. Though, content moderation could be necessary to regulate the spread of fake news, hate speech or arbitrary opinions, it could also potentially infringe upon people’s freedom to speech and expression. AI could imaginably and inadvertently filter out the content related to any prejudiced group’s opinions and rights. Moreover, in a study conducted by Cornell in collaboration with Facebook, it was found that platforms can undertake decisions in accordance to the users’ expressions that could mask one reality over another and thus, manipulate the user’s expression.\(^{36}\) It also raises the question of free speech given to the AI system itself and the application of this right on non-humans and the risks associated with it.\(^{37}\)

The right to liberty also encompasses physical freedom concerned with individual autonomy. Data determination tends to infringe on such liberties accorded to individuals. Data determination refers to the interpretation of data to form certain behavioral assumption by AI algorithms through the data collected online. Such predictive inferences are dependent upon the data published by the user online. Thus, these conclusions on the basis of a person’s information patterns could curb the freedom

\(^{33}\) General Data Protection Regulation [2016] OJ L 119/1
\(^{34}\) Isabella Banks & Leonore ten Hulsen (n 5)
\(^{35}\) Universal Declaration of Human Rights (adopted 10 December 1948 UNGA Res 217 A(III) (UDHR) art 19; and International Covenant on Civil and Political

Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR)
\(^{36}\) Mark Latonero (n 24)
to expression and speech on the internet. Therefore, individuality and speech are significantly impacted by the algorithmic profiling.\textsuperscript{38} However, the freedom of speech must be balanced in light of other cyber issues such as public shaming and misinformation. The design process to determine and outline such discrepancies is imperative in striking the desired balance through AI systems.

**Political Participation**

Freedom is a strong derivative of democracy. The developments in AI could allow bad-faith agents to undercut democratic regimes, hence, challenging the integrity of personal freedoms. A popular case of political manipulation is the alleged Russian interference with the 2016 US Presidential elections with automated bots on popular social media networking site, Facebook.\textsuperscript{39} Such bots are trained to replicate and mirror human dialogue that could hinder detection of such bots and in turn, lead to weaponization of information through popularizing of false narratives.

The dangers of such political manipulation manifests in shaping public opinion and views towards a political view or agenda. In another instance, a verified and state-operated Qatari news agency’s Twitter account and Website was hacked, resulting in the publication of false quotes said to be stated by the Emir of Qatar. The news sparked turmoil in Qatar’s political relations and benefitted its rival countries and resulted in Saudi Arabia, UAE, Egypt and other countries ending their diplomatic ties or relation with Qatar with serious implications on travel, trade and other political mistrust.\textsuperscript{40} Thereby, the cyber hack induced a situation of conflict between various political regimes.

In another episode, Facebook was in the spotlight when allegations surfaced that the popular social media giant was responsible for the circulation of exaggerated hate speech that led to the incitement of violence and blood-shed in Myanmar concerning the Rohingya tribe.\textsuperscript{41} The failure of its AI algorithm to detect such inciteful content was a result of gaps in Facebook’s AI algorithm, moreover, developments to the algorithm to detect such content in indigenous tones might not be an easily attainable reality.

**Criminal Justice**

The use of risk assessment tools in criminal justice could allow negative and undeniable impacts on human rights. Every person including convicts or defendants enjoy certain rights that cannot be abridged through any method. AI challenges the foundation of these rights through systems that are used to assesses factors and determine liability in the various stages of a trial. Such risk assessment tools generally rely on manually inputted information and thus, render automated judgements according to the recorded data. The data recorded relies on static and dynamic information that includes psychological history, age and gender.

\textsuperscript{38} Evgeni Aizenberg and Jeroen van den Hoven (n 15)  
\textsuperscript{39} Mark Latonero (n 24)  
Such automated descriptions of guilt might seriously impact the concept of fair trial and the value accorded to the due procedure of law. The AI systems predict the innocence of the accused based on certain factors that results in prejudicial assumptions. These risk assessment tools also have a tendency to target minorities and highlight racial disparities, and thus, reinforces discriminatory stereotypes. However, on the other hand, the impact on social media for crime detection is widely applauded.\(^{42}\)

Furthermore, some AI systems use methods of surveillance in particular areas which are dominated by the disadvantaged classes and thus, create a clear mark of distinction between classes of different societal statuses and rely on such to determine the possibility of future commission of crimes. This type of predictive justice leaves space for the degradation of human rights by forcing biases onto people through computational methods. The privacy and data protection of marginalized sections of the populations is unfairly targeted by AI algorithms.

Risk assessment tools are being equipped by various jurisdiction. The “Correctional Offender Management Profiling for Alternative Sanction” or COMPAS is an AI algorithm developed by the United States. COMPAS analyses and studies information pertaining to social and economic positions. The algorithm has been subject to critical reviews since it has been observed to discriminate on the basis of race wherein it renders dark-skinned people twice as likely to commit crimes compared to light-skinned people.\(^{43}\)

The United Kingdom also utilizes such risk assessment tool called “Harm Assessment Risk Tool” or HARM, which has also been subject to allegations of infringing upon people’s privacy and supporting discriminatory biases. Similarly, Canada uses “Level of Service Inventory-Revised” or LSI-R tool which was developed through 20 years of research which takes into account psychological profiles and other factors such as age, gender and vocational skills. It is one of the most popularly used risk assessment tool which is used by countries such as Scotland, England, Wales and the Channel Islands.\(^{44}\)

**REGULATION OF AI IN VARIOUS JURISDICTIONS**

**International Regime**

The regulatory principles to guide AI application are accorded an increasingly important role, given the wide scope of application and reach of the technology and their future development.\(^{45}\)

At the international level, the governance of AI is attempted through various documents. The OCED Principles on AI, released in 2019, acts as a guiding document for the regulation of AI. However, the provisions contained therein are non-binding in nature. The G20 economies have agreed to follow

\(^{42}\) Maria Stefania Cataleta (n 19)  
\(^{44}\) Laura Stanila (n 21)  
guiding principles that resemble the structure of the OCED Principles on AI.\textsuperscript{46}

Additionally, the United Nations is actively working and discharging functions through its various constituencies to shed light upon the effective regulation of AI.

The working body of United Nations Interregional Crime and Justice Research Institute (UNICRI) in 2015 formed a center on Artificial Intelligence and Robotics with the main objective to establish a uniform committee to focus on application of AI. The UNICRI coordinates and collaborates with various other UN organs such as INTERPOL, International Telecommunications Union (ITU) and various other bodies. Moreover, the ITU has become a significant body that studies and conducts research on the emerging concerns and developments in AI. The World Economic Forum through its Center for the Fourth Industrial Revolution is actively moving towards design a governance framework for AI application for states and corporations.

A Declaration on Ethics and Data Protection in Artificial Intelligence was discussed and formulated at the 2018 International Conference of Data Protection & Privacy Commissioners (ICDPPC). The contents of the Declaration elucidate upon six guiding principle and establishes commonality on governance of AI in members states. The ICDPPC lists a permanent committee to research and study the ethics of artificial intelligence.\textsuperscript{47}

The Institute of Electrical and Electronic Engineers (IEEE) has developed the Global Initiative on Ethics of Autonomous and Intelligent Systems which aims to promote ethically aligned AI design and provides guiding principles to achieve such purpose. The Initiative proposes establishing technical standards for AI in line with ethical perspectives.

India

The growth of AI in India has been remarkable and has been the primary reason for large investments to the Indian economy. However, the laws in India fall short in mandating a regime for the governance of AI and its application. The Government has, instead, chosen to focus on establishing an efficient strategy in terms of AI management.\textsuperscript{48} The Ministry of Industry and Commerce set up the Task Force on AI for India’s Economic Transformation which published a report on the challenges and opportunities in AI for various sectors in 2018. The Task Force consisted of various industry experts and received advisory participation from other bodies of the government.

The Report suggested that “AI should be seen as a scalable problem solver in India rather

\textsuperscript{46} Angela Daly and others, ‘Artificial Intelligence Governance and Ethics: Global Perspectives’ (2019) 15 The Chinese University of Hong Kong Faculty of Law

\textsuperscript{47} European Data Protection Supervisor and others, ‘Declaration on Ethics and Data Protection in Artificial Intelligence’ (40\textsuperscript{th} International Conference of Data Protection and Privacy Commissioners, Brussels, October 23 2018) <https://globalprivacyassembly.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_Ai-Declaration_ADOPTED.pdf> accessed 19 August 2021

than only as a booster of economic growth”. The Task Force recommended the Government to constitute a National AI commission with participation from different and relevant ministries to synchronize AI activities in India. Furthermore, the Report focused on delivering suggestions such as the establishment of data banks, the operations of AI are clarified, encouragement of AI education, and participation along with discussion of international practices for socio-economic development.\(^{49}\) The Report focused its study on 10 relevant sectors and the implications of AI thereon. However, the Report lacks in addressing major technical concerns of AI and is void of measuring the social and ethical impacts of AI application. The issues of data collection and surveillance remain unanswered by the Report and subsequent observations published by the Task Force.

Additionally, NITI Aayog is given the responsibility to lead the national initiative on AI regulation. The authority that assists in the development of public policy is working towards developing a national strategy for AI that would harbor overall growth and socio-economic expansion, and reduce conflicts. The suggestive guidelines given to the Government elaborate upon the necessity to consider the privacy and security concerns in AI and its ethics.

The Ministry of Electronics and Information Technology (MeitY) constituted four organizations that are headed by industry experts with the objective to facilitate research in AI.\(^{50}\) The focus of these committees is based on understanding AI according to legal, regulatory, citizen’s rights and cybersecurity perspectives. Moreover, the Government has also focused on analyzing the national security threats and opportunities of AI through the constitution of a Task Force to study such implications. The Ministry of Defense received the Report by this Task Force in 2018 which suggested the prioritization of regulation of AI weaponization that could be a considerable tool for the Defense Ministry.\(^{51}\)

These documents published are considered to be reasonably sufficient to address the issues and concerns of AI in India. However, culture of self-regulation and structured decision-making weighing the guiding principles involved in ensuring ethical and social security, is deemed as imperative in ensuring regulatory compliances without social harms.\(^{52}\)

**China**

China has been remarked to be robust with its strategic regulation of AI development and

\(^{49}\) Angela Daly and others (n 46)


ethics.\textsuperscript{53} China has demonstrated its commitment towards establishing a strong legal background to AI application in the nation since 2013. The Chinese Government published a series of national level administrative guidelines over the years to streamline and direct the implementation and integration of AI into the society.

In 2017, the Chinese Government through its main administrative body, the State Council released the “Artificial Intelligence Development Plan” (AIDP) proposes a uniform approach towards AI regulation to maximize its benefits and accurately present its risks and solutions. The AIDP’s primary objective rests in ensuring that China is in the forefront of AI innovation by 2030 to actively boost China’s industrial and economic development by deploying AI mechanisms to advance the sectorial growth. The AIDP outlines a clear framework to achieve such objectives in three parts. These parts determine a set of goals and steps that are to be implemented by 2020, 2025 and 2030. Hence, fulfilling the objectives laid down by AIDP by 2030.\textsuperscript{54} The AIDP aimed to establish a spirit of competition, foundational guiding principles to ensure ethical norms and generate 150 billion yuan by 2020. By 2025, the AIDP visualizes a major breakthrough in AI in order to be a world-leader in its applications and push the AI industry’s worth to 400 billion yuan, and introduce further appropriate ethical reliefs to the AI regulation. Finally, by 2030, China aims to secure itself as the best AI innovation actor and industry worth to reach 1 trillion yuan with secure regulatory policies to direct AI application and justly deal with emerging concerns.

China’s Ministry of Science and Technology along with a newly established AI Strategy Advisory Committee have been tasked with the responsibility of implementing the directions set forth in the AIDP. In order to kick-start the implementation of AIDP, the Chinese Ministry of Industry and Information Technology (MIIT) released a detailed plan titled “Three-Year Action Plan for Promoting Development of a New Generation Artificial Intelligence Industry” which elucidates on the active measures that are to be adopted by the Chinese Government to achieve AIDP’s objectives.\textsuperscript{55}

However, the AIDP is merely a strategy that provides structured guidelines for achievement of China’s AI and the main responsibility for the implementation and success of the plan is placed on the local government and private sector agents.\textsuperscript{56}

**United States of America (USA)**

The law in USA concerning Artificial Intelligence is still in its infancy stage. The main focus of regulatory authorities and the legislators has been the use of AI in autonomous vehicles and the use of AI in deployment of defense mechanisms.

The Government has released certain reports that aim to clarify the position of AI in USA and its potential impact on the society and economy. In 2019, an executive order issued titled “Trump Administration Executive Order on Maintaining American Leadership in Artificial Intelligence” was issued that is considered the foundational document on regulation of AI.

The Order directs the National Science and Technology Council (NSTC) through the

\textsuperscript{53} Mark Fenwick and others (n 45)  
\textsuperscript{54} Huw Roberts and others (n 17) 60  
\textsuperscript{55} Tariq Ahmad and others (n 51) 36  
\textsuperscript{56} Maria Stefania Cataleta (n 19)
Select Committee on Artificial Intelligence to assume an active role in the implementation of guidelines mentioned in the Order. The guiding principles specify the imperative need to consider “civil liberties, privacy and American values” in the application of AI technology. The objective of the Order is to formulate standards and comprehensive strategies that ensure AI is harvested and used in an efficient manner that would align with USA strategic economic and social goals.

However, the governance of AI in USA rests on application of various regulatory legislations that concern data privacy and protection, intellectual property and commerce.

The Federal Trade Commission (FTC) released a set of guidelines that aim to govern and coordinate the commercial applications of AI. The guidelines acknowledged the potential threats of AI while giving due credit to its benefits.

Apart from these guidelines and other related legislations, the law related to AI in USA is ill-defined.

**European Union**

The European Union (EU) has assumed a proactive role in determining the impact of AI. In 2017, the European Parliament approved of the Resolution on “Civil Law Rules on Robotics” that proposed measures in field of AI to the European Commission. The Resolution provides recommendations for the creation of a rigid structure to regulate robots, including their civil law liabilities.

The General Data Protection Regulation (GDPR) is acclaimed as the premier legislation that regulates and protects digital rights and is applicable to all EU Member States without the passing of a domestic legislation. GDPR acts as strong suit of armor against any instances of data breaches or privacy threats and is regarded as “the toughest privacy and security law in the world”. The GDPR contains certain provisions that partially restrict the validity of AI-based decisions upon individuals. However, the underlying intent of the GDPR is ambiguous due to its technical complexities. Subsequently, in 2018, the European Commission published the “Communication on AI” which provides a direction to the socio-legal and ethical values maintained by the GDPR.

In 2020, the European Parliament published “The White Paper on AI”, which acted as a foundation for the European Union’s latest “Proposal for the Regulation

---

57 Tariq Ahmad and others (n 51) 27
60 Tariq Ahmad and others (n 51) 6.

---

64 Ibid.
laying down Harmonized Rules on Artificial Intelligence”, released in April 2021. The White Paper addresses the dual aspects of AI in terms of promotion of investments and risks of ethical interference. The White Paper aimed to address these points by establishing policy guidelines to maximize the benefits of AI while eliminating its risks.65 The contents of White Paper telegraphed into the new Proposal that requires high-risk AI systems to maintain a standard of compliance towards the rules on data governance, documentation and record-keeping, transparency, human oversight and security.66 The Proposal is deduced to be the GDPR equivalent for AI application.67 However, the controversy surrounding the Proposal indicates serious concerns by companies and individuals pertaining to the development of AI, personal security and privacy and human rights.68 The enactment of the GDPR resulted in the setting of a global legislative standard. Similarly, the Proposal may be welcomed globally as the legal and ethically framework as well.

65 Stefan Larsson, On the Governance of Artificial Intelligence through Ethics Guidelines, (2020)7(3) Asian Journal of Law & Society 446
68 Joe Hilleary, ‘Breaking Down the New EU Regulations for Artificial Intelligence’ (Eckerson Group, May 17 2021)

CONCLUSION

AI was never an unimaginable concept to mankind. The mention of AI is found in ancient mythology and folklore that highlights pertinent social and culture issues. In Greek mythology, there are accounts of Hephaestus, a Greek god, who created mythical golden robots that were able to think for themselves but were enslaved by the gods. 69 The Golem of Prague, is another famous example of AI in folklore.70 AI application draws its own set of risks and benefits, both of which pose certain effects on the human race. On one hand, AI is applauded for introducing revolutionary measures of modern technology; on the other end, AI is condemned for its potential ramifications upon human rights. AI offers great tools that largely benefit various industries such as the healthcare sector.71 Successful strides in the development of AI have also led humans to technology that can influence neurons, such as memory teaks and implanting images, which offers health benefits in terms of curing neurological diseases but could also be used as a lethal weapon for intrusion.72

70 Ibid.
72 Melissa Heikkila, ‘Machines Can Read Your Brain. There’s Little That Can Stop Them’ (The Politico,
AI aids human productivity and adds value to human creations on its own merits. However, the dark side of AI rests on the extent of power given to this technology without adequate checks.

The overlapping of ethical and human rights concerns of AI sheds light on the possibilities of human unemployment, job market manipulation, inequality by racial discrimination and unfair wealth distribution, and security breaches. These identified threats of AI present a deeper question of the efficiency of legal framework to mitigate such risks.

Policy regulations for AI in a rights-inclusive manner would structure the development and deployment of AI. However, theoretical implications of law may have the unintended effect of slowing down the growth of AI. The debate on the necessity of a rigid regulatory framework and the extent of its application on AI raises the question of the adequacy of an ethical or technical approach to govern AI.

In the words of Stephan Hawkings, “success in creating AI would be the biggest event in human history. It might also be the last, unless we learn how to avoid the risks.” The measures needed to successfully implement AI cannot be ignored. AI’s tendency to pose serious unjust implications that concern socio-ethical dynamics of the society drags the concept of a human rights perspective on AI into the forefront.

Most countries, though have recognized the ethical gaps of AI, are cautious about determining a legal framework to not hinder their own technological goals. However, human rights solutions to AI are believed to be the key to fulfil innovation and successful interaction between real and artificial systems of intelligence in order to accurately satisfy stakeholder needs consistent to the moral and social values embodied by human rights, and hence, offers long-term benefits. The discriminatory nature of AI induces the risk of inequal division of classes. Marx pointed out that emancipation conceived fully in terms of rights is unappealing. A society that relies on rights-based ideals might not always achieve everything. However, the quality of AI to increase inequality might threaten the very existence of the society. Therefore, adoption of active measures to combat the ethical and social risks of AI become necessary.

The present legal atmosphere is struggling to catch up to technological driven innovations, therefore, applying the current human rights

Notes:
6. Mathias Risse (n 16) 15
framework on AI becomes more imperative than establishing a new ethical standard for AI.  