The Global Ai Race: Exclusivism V. Free Data Authoritarianism

Dr. Supreet Gill*

Jayendra Singh Kanwar**

1. INTRODUCTION

The special characteristic that sets humans apart from animals and makes them more advanced living beings is intellect. It is our intellect that has driven the human 'strive for more' way of life. By nature, everyone is competitive and has historically strived for more – the biggest, tallest, longest, strongest. Lately, this ambition has extended to achieving the 'smartest'. This is why Artificial Intelligence (AI) has become a key term in policy discussions. After the global pandemic, which allowed us to pause and reflect on our lives and thoughts, one might have expected humanity to emerge more content and less power-hungry. One might have thought that we humans would have learned from the global races of the past, such as the arms race and the space race of the Cold War era, and that these would be relics of the past. However, just like a tiger never changes its stripes, it seems apt to say that humanity's desire for more is insatiable and will inevitably resurface, even after a deadly pandemic; and the latest manifestation of this is the pursuit of AI.

Like nuclear energy and semiconductors, AI is a dual-use technology with civilian as well as military applications. However, with an ever-looming danger of regional conflicts escalating to global conflict; compounded by skewed electoral democracies giving way to autocracies and non-State entities running amok; misuse of AI is a reality. Even though it gives the impression that AI is a novel phenomenon and would herald some revolution; contrary to popular misconception, AI has been around for quite some time. Despite being around for a while the way AI notoriously escapes any definition is what the scholars call the AI Paradox. In an attempt to find an apt definition of AI when one compares definitions offered by various scholars, some similarities and

* Dr. Supreet Gill is an Assistant Professor at the University Institute of Legal Studies, Panjab University Chandigarh. An alumna of Symbiosis Law School, Pune and the prestigious London School of Economics and Political Science, UK.

** Mr. Jayendra Singh Kanwar is an advocate and the Punjab and Haryana High Court, Chandigarh.

overlapping is inevitably apparent, which then have been categorised accordingly. The importance of defining a phenomenon which may revolutionize the way we humans lead our life becomes exceedingly urgent and mandatory because it is only when AI is defined that any view can be formed of its liability. It is a compelling discussion that challenges the orthodox jurists which further begs the questions - how far is humanity willing to push? In circumstances where AI interacts and intertwines with various facets of law like that of Intellectual Property Rights (IPRs), this already knotty debate is further enlivened.

On one end of the spectrum, there are authors whose work is fed to the AI, without acknowledging them, in the name of fair dealing. On the other end, there are questions around AI-generated work. It impliedly holds the potential of shattering the monopoly of humans over creativity as well as standards of creativity. Moreover, the data scarcity which the global AI race has resulted in, will bring in the standard bearers of AI development in conflict with the IPR holders asserting their restrictive rights. This data inequity calls for *data justice*. The present paper dwells on the debate surrounding definition and legal personality of AI and then, unravels the IPR considerations of AI at the input as well as the output stages. It brings in the socio-economic dimensions, which necessarily play out in the context of widening inequality on different fronts. It concludes with policy recommendations with recent geo-political developments in view.

2. EMANCIPATION OF THE AI

Artificial Intelligence is not a novel phenomenon, but can be traced back to the year 1950 wherein in a paper titled, 'Computing Machinery and Intelligence'[1], Alan Turing[2] furthered his view that if a machine could pass the Turing test it would be considered intelligent. The test involved a human judge asking questions to a human and a computer. If he could not distinguish the responses from the human respondent and from the computer, the computer would be taken to have passed

[1] Alan Turing's "Computing Machinery and Intelligence" is a foundational article concerning artificial intelligence. The paper, which was initially published in Mind in 1950, was the pioneering work that introduced the general public to his notion of the Turing Test.

[2] Alan Turing (born June 23, 1912, London, England—died June 7, 1954, Wilmslow, Cheshire) was a British mathematician and logician who made major contributions to mathematics, cryptanalysis, logic, philosophy, and mathematical biology and also to the new areas later named computer science, cognitive science, artificial intelligence, and artificial life. (Mar. 24, 2023, 10:55 AM) https://www.britannica.com/biography/Alan-Turing

the test. Half a decade later, the term 'artificial intelligence' was coined by John McCarthy[3] (popularly known as the father of AI) at the first AI conference held at Dartmouth Summer Research Project on AI in 1956.

Putting aside these initial endeavors in the implementation of artificial intelligence, it is indisputable that the three functions that an AI is designed to execute i.e., classification, recommendation and prediction have lost their novelty to the internet. Search engines and social platforms have been making use of the AI to perform these functions since their very inception. This is the so-called 'AI paradox' which implies that as people get accustomed to new AI technology, it is no longer considered AI; since the novelty soon wears off. Nonetheless, the buzz that AI has generated lately is not without a cause.

The world is not bipolar as it used to be during the Cold War. It is not even unipolar as it was for some time following the collapse of the USSR. It is rather multipolar. New poles are surfacing or countries with new found wealth or new-found confidence are vying to pose themselves as the axis around which their neighbors will revolve. The tension this phenomenon has generated has not left artificial intelligence unscathed. By its very nature artificial intelligence is data guzzling. Data as an asset is bound to become scarce as a consequence of its over-exploitation. Thus, the fight over who develops more advanced AI and the concern about the impending data scarcity have reinvigorated the interest of the world in AI.

Artificial neural networks which enable AI to mimic human neurological activity is of recent vintage. Large language models have been developed as a corollary to these neural networks. These are able to generate texts on prompts which makes it difficult to distinguish artificial intelligence from human intelligence. Then there is generative AI which is not restricted to generating text in response to prompts but can generate even images and videos. Hyper-realistic videos generated by *Sora* of OpenAI indicate the level of advancement achieved in a short span of time by the startup. Sam Altman, the face of OpenAI, suffered a corporate coup d'état before he was reinstated as the CEO. Satya Nadela, the CEO of Microsoft was eager to induct Sam Altman

[3] John McCarthy (born September 4, 1927, Boston, Massachusetts, U.S.—died October 24, 2011, Stanford, California) American mathematician and computer scientist who was a pioneer in the field of artificial intelligence (AI); his main research in the field involved the formalization of common sense knowledge.

after he was ousted by his startup. Microsoft has been funding OpenAI and Nadela's stance comes as no surprise. Elon Musk was signatory to the Asilomar Principles[4] which seek a cautious development of AI technology. But a corporate stalling the development of a technology at one front and making breakthroughs at other fronts at the same time may be indicative of his wider efforts at buying time for himself to develop an AI system of his own which matches or surpasses that of Altman.

2.1 CAN AI BE APPROPRIATELY DEFINED?

All the talk around generative AI or large language models is devoid of depth because it just deals with technical differences between different AI technologies. Conversely, it is imperative that any definition of AI must not confine itself to technical nuances. Several definitions have been proposed by several scholars coming from varying walks of life. Broadly, the definitions use either human intelligence benchmark or the rationality benchmark. Taking AI as a software, it is said to be intelligent if it can think and act like a human. This is the human intelligence benchmark and such a definition is human dependent definition. But a software is considered intelligent if it can achieve ideal outcomes as a rational agent. This is the rationality benchmark and it is a human independent definition. Not taking sides, it is appropriate to leave the definitional quandary with a remark that it is the context which best defines the phenomenon. What it means for a legal researcher may not be the same what it means to a software developer.

One issue which is pertinent to mention here is that of the legal personality of AI. This is because, in order to define the contours of rights and duties of AI, a pre-existing assumption of its legal personhood beckons. This arena is highly polarised because on one hand there are those who do not see any difference between artificial and human intelligence and on the other there are those who do not see some qualities peculiar to humans in AI.[5] Treading the middle path who do not want to get into this question jump onto the product liability bandwagon. Admittedly, the debate whether AI is a subject or an object can be best dealt by philosophers. Still jurists consider product

[4] Asilomar AI Principles are 23 guidelines for the research and development of artificial intelligence (AI). The Asilomar Principles outline developmental issues, ethics and guidelines for the development of AI, with the goal of guiding the development of beneficial AI.

[5] R.D. Brown, *Property Ownership and The Legal Personhood Of Artificial Intelligence,* 30:2 INFORMATION & COMMUNICATIONS TECHNOLOGY LAW 208, 220-234 (2021).

liability as the ideal solution so as to enforce any liability for AI. But then what will be the nature of this liability?

2.2 AI AND ITS VARIED INTERPRETATION

The EU takes a risk-based approach in its AI Act as it divides AI systems into unacceptable, high and low risk applications. But this approach has been criticized as an exercise in 'taxonomy.' These critics categorise the potential claims that can arise from the applications of AI into personal injury and death; dignitary or reputational injuries; damage to property and other losses. They suggest a framework wherein there would be a no-fault strict liability for personal injuries and death because these are less likely to be covered by insurance and a fault liability for dignitary or reputational injuries. Other losses would more likely be covered by insurance.

In this scheme of things, the user and the software supplier would be jointly and severally liable. But the one who substantially benefits from the use of AI would bear the burden to pay the damages if something goes wrong. This liability would be parallel to product liability. Since, consumer protection defines a defective product as one which does not meet the expected standard of safety, defective AI cannot be defined this way as people cannot expect any standard of safety whatsoever. This is in turn because AI is shrouded in mystery as to what it would bring in its wake. But then certain factors have been brought out which would be kept in view—function which the AI is reasonably expected to do, defects at the time of production and adaptability of the algorithm. Moreover, how can fault liability be enforced against AI as it does not have legal personality? The courts then are advised to judge whether a decision reached by AI would be considered negligent had it been taken by a human in comparable situation.

2.3 LEGAL PERSONALITY OF AI OR AI PERSON HOOD

Those opposed to conferment of legal personality on AI believe that legal personality is no guarantee of civil rights as has been the sad case with people of colour and women. Brandeis Marshall recommends penalties like tech probationary jail, short-term ban and long-term incarceration for algorithms, processes, systems and tools. She has formulated an AI Dependency Spectrum. Dependency exceeding 90% qualifies as AI dependent. Ranging between 50%-90%, it

would AI-assisted classification and between 10%-49%, it would be AI-enhanced classification. Below 10%, it would be AI-lite classification. She seeks to disincentivize heavy AI dependency.[6]

Jasper Doomen explores the basis of legal personhood. He quotes John Searle as holding the view that program is to computer what mind is to brain. But then program is a syntactical process but mind is something more than mere syntactical process. Doomen argues that even though humans are not programmed the way AI is programmed, still one's brought up and genotype are like his programming. He negates the argument that lack of free will amounts to lack of autonomy on the part of AI. He distinguishes fundamental autonomy which is based on dignity which is in turn based on morality from legal autonomy which can exist independent of fundamental autonomy. Moreover, he distinguishes fundamental personhood which is not restricted to natural persons from legal personhood which has no standard criteria.[7]

His conclusion gets entrenched when arguments from all the camps in the fight over legal personality have been skimmed. He concludes that there is not 'principled objection' to conferment of legal personality on AI. It is just practical objection which is made in view of the expedients and compulsions of the day. As the title of the chapter borrowed from his paper, 'emancipation of AI' conveys, granting legal personhood to AI would be along the trend of ever-expanding list of entities being recognised as legal persons—slaves, women, animals and robots. Although legal personality will come with rights, he admits that it is unclear how these rights will be exercised by AI. Sure, in a face-off between Google's *Gemini* and the ruling dispensation in India over the response generated by it over Mr. Modi and fascism, right to freedom of expression would have come to its rescue. It is a different matter that the government released an advisory without legal backing requiring AI developer to seek government permission with regard to "under testing applications". It then exempted startups from the purview of the advisory and eventually withdrew the advisory.[8]

---

[6] B. Marshall, *Labelling Your AI Dependency* MEDIUM (Mar. 23, 2024 6:45 PM) https://medium. com/@brandeismarshall/labeling-your-ai- dependency-9828194877a3

[7] Jasper Doomen, *The Artificial Intelligence Entity As A Legal Person*, 32:3 INFORMATION & COMMUNICATIONS TECHNOLOGY LAW 277, 277-287 (2023).

[8] 'IT Ministry Replaces AI Advisory, Drops Requirement of Government's Permission' *THE HINDU* (New Delhi, Mar. 16, 2024).

3. AI RACE FOR SUPREMACY

The commitment to the development of responsible artificial intelligence (AI) has been explicitly affirmed by prominent economies across the globe. Leading the way, United States, President Joe Biden has signed an executive order that establishes norms for the safety and security of artificial intelligence. The Bletchley Declaration was released during the AI Safety Summit in the UK, with the participation of 28 countries and the EU which pledged to collaborate in an inclusive manner to guarantee AI that is human-centric, trustworthy, and responsible. The International Guiding Principles on AI and a voluntary code of conduct for AI Developers were released by the G7 leaders as part of the Hiroshima AI project. With the rapid advancement of AI, countries are engaged in a fierce competition to establish global dominance in the field. The aforementioned advancements possess significant geopolitical ramifications, since several nations articulate their distinct perspectives on this technological advancement. This worldwide endeavor prompts two significant inquiries: What are the key factors that will determine the establishment of AI supremacy? What role will emerging economies such India assume in the context of this global competition?

3.1 INDIC-KNOWLEDGE HIJACKING: DATA COLONIALISM?

Knowledge and its acquisition, as well as the connection between knowledge and truth, are the key topics of epistemology. The idea that Indic knowledge systems are either grossly under-or incorrectly represented in Anglo-sphere epistemology is a perennial worry. Generative AIs have the capability to replicate creative works by specific authors, including J.K. Rowling's Harry Potter series, down to the last timbre and pitch. This holds true of works of art as well. In the background, two things are occurring: creators will cease to produce original works due to a lack of incentives (at least according to current intellectual property theory) and any copyright inherent in these works has been lost. Unexpectedly audacious for a country that operates cautiously by consensus, Japan’s recent decision to disregard copyrights in datasets used for AI training (from the blog technomancers.ai, “Japan Goes All In: Copyright Doesn't Apply to AI Training”). The recently enacted legislation in Japan grants artificial intelligence the authority to utilize any data, irrespective of its commercial or non-commercial nature, source (other than reproduction), or origin (including content obtained from illicit websites or otherwise). The Japanese government declared its intention to abstain from copyright enforcement regarding data utilized in AI training

in May 2023. The policy permits the utilization of copyrighted material for the purpose of training AI models, irrespective of its origin (illegitimate websites, commercial or non-profit). This means that businesses are free to incorporate any amount of copyrighted material into the training data of their models.

In the midst of these fast paced global developments India cannot afford to remain on the periphery. It is possible to contend that this poses an existential threat to India and requires an immediate response akin to engaging in armed conflict. Although IIT Madras' *AIforBharat* initiative is a beginning, considerably more must be done. The government must promptly implement a set of policies and regulations that are extremely targeted. These measures should not only deter unauthorized access to our intellectual property and data, but also foster the development of numerous models that effectively leverage Indian ingenuity and knowledge.

3.2 THE GEN-AI COPYRIGHT INFRINGEMENT LANDSCAPE

While artificial intelligence (AI) in general has sparked a lot of conversation, the intersection of AI with IPRs is now a very contentious issue. At its best, AI has the potential to greatly facilitate efforts to maintain and safeguard intellectual property rights. However, new US court judgments suggest that AI could jeopardize human creativity, which is protected by intellectual property rules and could be used for profit. Artists, writers, and content creators are understandably worried about the impact of generative AI (Gen AI) tools like ChatGPT on their careers. Copyright law on a global scale recognizes and protects the authorship and legal standing of works conceived and created by humans. Nevertheless, the advent of AI-generated and AI-assisted works has brought about an entirely novel set of intricacies and ambiguities within this field.

WIPO in its session on "WIPO Conversation on Intellectual Property and Artificial Intelligence", categorized works created using AI as "AI-generated works" (where the output is generated without human intervention) and "AI-assisted works" (where the output is generated with material human intervention and/ or direction).[9]

A growing number of copyright infringement lawsuits are being filed in the US. The U.S. Copyright Office issued a partial copyright to the comic book "Zarya of the Dawn" by Kristina Kashtanova last year. This is an illuminating use case, as Kashtanova generated the images for her

[9] Latha R Nair and Sudensha Banrjee, *Mitigating Liability While Copyright Law Catches Up with Artificial Intelligence* K&S Partners, (Jan. 30, 2024).

comic using Midjourney. Kashtanova "is the author of the work's text in addition to the selection, coordination, and arrangement of the work's written and visual elements," according to a determination by the Copyright Office. Nevertheless, she does not possess the physical copies of the images. Despite not being granted copyright for the images, Kashtanova regarded the ruling as a positive development for American artists utilizing AI tools. Kashtanova writes, "The arrangement is protected by copyright when your images are included in a book like Zarya." "The story is also protected by copyright, provided that it was not produced solely by artificial intelligence." Although the situation is still nascent, this ruling does offer some indication of how legislators in the United States are contemplating the intersection of copyright law and the implementation of AI tools.[10] This case can be seen as a departure from the position previously taken by US Courts in 2018, in the case if *Naturo* v. *Slater*[11] (popularly known as the monkey-selfies case) wherein they held that "completely autonomous non-human entities cannot claim copyright as they no dot have any statutory standing before the court under the copyright law.[12]

3.3 AI'S COPYRIGHT OWNERSHIP: ORIGINALITY V. AUTHENTICITY

If one follows, the Turing test, it would not be wrong to point out that any computer who succeeds at the test would be considered "intelligent", which means a presumption arises in favour of it capable of possessing intellect. As pointed out, multiple times by the courts and legal scholars, one of the main objectives of copyright law is to reward the creativity of the "Author"? Then why this discrimination between human generated work and work created by a machine or a computer? This discrimination not only seems arbitrary but is also being compared to slavery.[13] If law is some countries is recognising AI with legal personhood for fixing liabilities, then AI's rights cannot be withheld on such capricious grounds.

The central inquiry in the copyright-related controversy surrounding AI revolves around the necessity and degree to which copyrightable works necessitate or ought to necessitate human

---

[10] John Donegan, *Artificial Intelligence: The US Should Look At Japan's Unique Approach To Generative AI Copyright Law User* (Feb. 24, 2024, 8:45 AM) <https://insights.manageengine.com/artificial-intelligence/the-us-should-look-at-japans-unique-approach-to-generative-ai-copyright-law/>accessed 18 March 2024

[11] *Naturo v Slater* 9th Circuit (2018); 2016 U.S. Dist LEXIS 11041.

[12] Prateek Deol, *Artificially intelligent world and its interface with Law* in V.K. Ahuja and Archana Vashishtha (eds) INTELLECTUAL PROPERTY RIGHTS, CONTEMPORARY DEVELOPMENTS (Thompson Reuters 2020).

[13] Mira T Sundara Rajan, MORAL RIGHTS 313 (Oxford University Press, 2012)

intervention. The concept of "authorship" lacks a precise definition within the framework of the Berne Convention. Nevertheless, it is highly recommended that the convention exclusively pertains to human producers, hence excluding content generated by artificial intelligence from its purview. Consequently, the Berne Convention's minimum requirement exclusively pertains to creations produced by individuals. The Indian Copyright Act 1957 is mostly modelled after the UK statute and applies to India, a country that was formerly a British colony. According to the Indian statute, the individual responsible for the creation of a computer-generated literary, dramatic, musical, or artistic work is referred to as an author in Section 2 (d)(vi). Any copyright application must only include the details of natural individuals as authors, according to the Practice and Procedure Manual (2018) of the Indian Copyright Office. It is possible to argue, according to section 2(d)(vi), that AI production that involves talent and creativity and has been heavily influenced by humans may be qualified for copyright protection. On the other hand, India's legal system for copyright ownership is inadequate when it concerns AI-generated outputs that do not involve any human involvement.

An artificial intelligence tool known as *Raghav* was awarded registration by the Indian Copyright Office in the year 2020. This accolade was given in honor of the fact that the tool and its human developer had jointly developed an artwork. Nevertheless, in the aftermath of the event, the Copyright Office issued a notification for the purpose of rescinding the registration award and mandating that the co-author and human developer submit communication to the Copyright Office of Raghav's legal position. This move appears to be the outcome of the Copyright Office's eventual understanding that the Indian statute does not provide any provision for the ownership of copyright by authors who are not human. This is a chilling reminder of the pitfalls of chattel slavery days where human used as slaves had no control or acknowledgement of work done by them. Is history on the path to repeat itself, whereby certain section of self-entitled creators with superiority complex might trump the rights of AI?

In all fairness, a criteria has to be devised as a benchmark for a work to be worthy of being granted ownership/authorship. The authors suggest that it may be devised as follows:

- If AI DOES NOT EXCEED the parameters set by the copyright law then it is NOT creativity.
- If AI EXCEEDS the parameters set by copyright law then it IS CREATIVITY.

Then even though work created by AI can be ascribed to the human making use of it, work created by the AI at par in terms of creativity and originality of human beings has to be acknowledged. As a result of the authorship issue surrounding the work of the AI, discussions that are centred on different narratives have emerged, which are currently causing a stir in the legal academia in a number of different countries.

The jurists arguing against AI authorship/ownership put forward the argument that, even though the work of AI may be creative and even meet the standards of originality in some cases; but can never be authentic. This argument is based on the reasoning that all AI work is derivative and somewhat mechanical in nature, wherein any AI software produces work on command given by a human and does not use its own intellect. AI is merely a software that compiles information after fetching it from the data already fed to it. And this data is almost always based on works previously done by humans. However, to counter this argument, it may be submitted that isn't all research based on some previous research? Isn't that what we human's also do? And isn't the primary objective of copyright law to 'allow *others* to build freely' upon previously done work? Then on what basis are we excluding AI from the purview of this term 'others'?

The fundamental basis of this matter resides in the manner in which generative AI systems undergo training. Similar to the majority of other machine learning models, their functioning involves the identification and replication of patterns within the data. In order to produce an output such as a written sentence or picture, it is necessary for the system to acquire knowledge from the actual activity of human beings. If an artificial intelligence (AI) picture generator generates artwork that bears resemblance to the artistic creations of Georgia O'Keefe, it can be inferred that the AI was taught using the authentic artwork of Georgia O'Keefe. Likewise, in order for an AI content generator to emulate the writing style of Toni Morrison, it must undergo training using Toni Morrison's own written words.[14] One growing application of Deep Learning techniques is the creation of synthetic data that may be utilized for many purposes, such as AI training and testing For example, DL systems are capable of producing copious amounts of data when they engage in what is known as "self-play," which is very useful when applied to reinforcement learning. Such information records the AI system's personal encounters with the actual world, be they virtual or

[14] Ellen Glover, *AI-Generated Content and Copyright Law: What We Know* (Mar.18, 2024, 7:38 AM) https://builtin.com/artificial-intelligence/ai-copyright

tangible. The importance of synthetic data generation and reinforcement learning in artificial intelligence is growing rapidly.[15]

3.4 Fair Use or Just an Excuse: The Moral Rights Debate

At the core of contemporary copyright legislation, it serves as the fundamental basis that enables scholars, journalists, filmmakers, and writers to utilize their creations without violating any rights in specific situations. The criteria used to ascertain the applicability of fair use encompass the purpose of utilization, the inherent characteristics of the copyrighted work, the extent of usage relative to the entirety of the work, and the influence on the potential market. Fair use acknowledges that the majority of works are derived from pre-existing works and seeks to avoid impeding future creativity and innovation.[16]

Utilising copyrighted content to train AI models is typically regarded as fair use, provided that the owner's consent is not necessary. The legal doctrine of fair use permits the restricted utilization of copyrighted content without obtaining explicit authorization. The considerations that establish the applicability of fair use include the purpose of use, the nature of the copyrighted work, the quantity used, and the impact on the potential market. Fair use acknowledges that the majority of works are derived from pre-existing works and seeks to avoid impeding future creativity and innovation. In December 2023, a lawsuit was initiated by The New York Times against OpenAI and Microsoft, alleging that ChatGPT and Copilot were improperly trained using the extensive New York Times article archive. There is concern among The Times and other organizations that these platforms may emerge as rival news agencies and become the primary source of news and information for consumers. This is the most exhaustive and detailed legal case pertaining to generative AI to date.

The practice of web scraping and crawling has been in existence for several decades and is generally considered to be lawful. Nevertheless, the manner in which the information is utilized can ascertain the presence of any copyright violation. Numerous websites possess terms of service that explicitly exclude web crawling. In the event that a web crawler disregards these restrictions,

---

[15] Data Governance Working Group: A Framework Paper for GPAI's Work on Data Governance 2.0 November 2022 - GPAI Tokyo Summit at page 9.

[16] Frank Palermo, 'AI Copyright Infringement Quandary: Generative AI on Trial (Mar.18, 2024, 7:38 AM) https://www.cmswire.com/digital-experience/ai-copyright-infringement-quandary-generative-ai-on-trial/

it may be deemed a violation of the contractual agreement, rather than a violation of copyright laws or infringement.

At this juncture it is pertinent to touch upon the clash between moral rights of the original author *versus* moral rights of generative AI and to determine whose rights are superior. The comprehension of the correlation between artificial intelligence (AI) and moral rights remains limited. According to several thinkers, the existence of consciousness, intellect, autonomy, creativity, or emotions in AI systems could potentially grant them moral rights. Nevertheless, there is a lack of agreement over the precise definition or quantification of certain abilities or traits. The notion of AI as a potential proprietor is based on the premise that AI is not merely an automated system, but rather has the potential to be an autonomous system. The supervision of AI system growth, encompassing data mining, machine learning, and training procedures, often involves human oversight. However, recent advancements in AI technology have facilitated the ability of AI to acquire knowledge from other AIs, a phenomenon sometimes referred to as "kickstarting".

Given this context, we can only express our viewpoint on whether moral rights for an AI align with the evolution and justifications of moral rights, as well as their legal basis. The moral rights concept is intricately linked to human inventions. Consequently, the application of moral rights to an AI would entail an expansion of the fundamental principles of moral rights, necessitating further justification.[17]

4. THE IMPENDING DATA WAR

Data serves as the building blocks of knowledge or intelligence, be it human or artificial. A person leaves a trail of data behind him; blissfully but riskily unaware how this data is disposed of or used or in some cases sold for a price or stolen. An order placed on an e-commerce platform, a booking made online for reservation of a hotel suite, payments made through mobile wallets; every act generates so much data, that a digital self of an individual would not be hard to construct. Sadly, this statement by the authors is not cynicism, paranoia or fiction talking, rather it is a haunting reality where humans are ceding more and more decision-making to digital systems neglecting the supposedly mundane nitty-gritties of complex interpersonal relations. A mental paralysis in the wake of AI and emphasis on machine learning is paving the way for a societal neurosis which is

[17] Miernicki, M., Ng (Huang Ying), *I. Artificial intelligence and moral rights* 36 AI & SOC 319, 324–329 (2021).

emblematic of the contemporary hypermodern era or industrial capitalism: the fixation on efficiency and effectiveness in one's use of time.

4.1 Understanding Machine Learning (ML)?

Machine Learning (ML) is the process whereby a machine is trained over datasets to deliver certain outcomes. These outcomes include classification of data into categories, use of past data to predict future outcomes and use past data to recommend new things. Looking at ML this way, it seems that it is nothing more than a sorting machine or a coffee dispenser—you put grounded coffee beans and hot water/milk into it and you get coffee. But it not as simple as it sounds. Tasked with constructing a sentence, AI has to decide the arrangement of words in a matter of seconds or even faster. Even if it uses statistics to judge the probability of what succeeding words would be, as per the 'Godfather of AI' Geoffrey Hinton, it has to be precise. This precision can only come with experience. This opens up the prospects of a self-perpetuating ML system, perfecting itself to the extent that it supersedes human intelligence. This may be far removed from reality at present, but if there is some possibility of this scenario materializing, then human have a lot to come to grips with.

Leaving the ruminations about the future aside and delving into the history of mankind, we see that in a feudal society, land was the resource which defined the relation between lords and serfs. Then came the age of capitalism with the Industrial Revolution in which financial capital determined the class system. Variously dubbed as the Fourth Industrial Revolution (4IR) or digital capitalism is an age in which data is the key resource.

4.2 What is the value of data and who are the players fighting over it?

A unit of data taken individually is of very little value. But when many such units of data are combined to form an aggregate, something of value is generated. This something is data intelligence i.e., the trends and patterns derived from the aggregate data using computational skills. The kind of orders one has placed or the kind of goods one has in his wish-list/basket on an e-commerce platform can help the AI underlying the platform recommend more and more goods to whet the consumerist appetite of the buyer. Knitting the bits of data together, such platforms can construct a 'digital self' and can manipulate consumer choices. As brought out in the data justice

primer of the Global Partnership on Artificial Intelligence (GPAI), e-commerce platforms can increase the prices for an impulsive buyer.

Since, the value of data lies in its aggregate, there would inevitably be a tendency towards data hoarding. Data platforms are not the only ones with a knack for data hoarding, even State agencies gather data to selectively reward or punish its citizens. An ordinary citizen, with no pretence of power would keep on wondering why his credit rating has taken a dip while that of his fellow citizen is intact or has risen, even though both of them are placed in similar circumstances. This is the Orwellian Big Brother manifesting himself through surveillance capitalism.

4.3 Is data Property? If yes then who is the owner?

There are competing arguments as to whether data is or is not property. Those like Professor Ben McFarlane of the University of Oxford consider data, even though intangible, property the same way bank deposits are property.[18] Without going into the merits of these rival arguments, if we take data to be property, then who has the primary rights over it? In other words, who is the owner of data? This leads to another question i.e., who all are the players in the data value chain. There is the person to whom the data relates, called the data subject. Then there is the person or entity holding the data of/for data subject, termed the data holder. Although there are those who gather data belonging to data subjects and bring it to data holders, called data workers (for instance, gig or platform workers); it would be safe to exclude them from the data proprietorship debate as there is a very negligible position in the value chain.

For the sake of convenience, leaving the question whether a waiver of data rights is possible aside, the more pertinent question becomes in what circumstances can the waiver of data rights be effected or be implied? Riddled with tardy terms and conditions set into a standard format, user of a digital platform can either board the platform or leave it. Thus, there appears a tendency not to give consent or withdraw the consent from sharing of data. This data fearing makes the data available under-representative as those who have not consented or have withdrawn their consent are missing. Thus, neither data hoarding by data platforms nor data fearing by users bode well for a vibrant data-driven economy. There is the need of a Digital Public Infrastructure (DPI) as well as the need for regulation over tech giants. Again, taking a leaf out of historical economic relations,

[18] Data Trusts and Defining Property, Ben McFarlane, Faculty of Law Blogs, University of Oxford (2019).

it was the State which provided the private players with the infrastructure over which they could carry out their marker-based activities. Thus, there was distinction in the actors who provided infrastructural services and those who provided super-structural services. This separation has blurred with the rise of digital platforms. E-commerce giants purporting to be a marketplace for suppliers are competitors in the provision of commodities as well. To add insult to injury, in addition to competing with the very suppliers to whom they are providing the digital infrastructure, they charge platform fees. In a recent episode, certain service providers from India were delisted from providing their services over platforms belonging to tech giant Google as they did not concede to the demand for platform fees.[19]

India boasts of a DPI and rightly so. But, in a segment of this DPI concerned with cashless transactions and financial inclusion, PayTM (short for Pay-through-Mobile) once considered one of the forces behind the DPI has faced the heat of the central bank lately over its infractions of the KYC norms (Know Your Customer norms, which require a service provided to confirm the identity of its consumers using identity markers assigned by the State agencies).[20] The economic justice primer of the GPAI urges the State to reclaim its position of infrastructure provider.[21] But that is no guarantee that the likes of *PayTM* who benefit from the DPI would behave in the best of ways. Thus, a strong regulatory framework is needed.

4.4 United We Stand

Given the asymmetry in the data ecosystem, where tech giants from a certain part of the world are extracting more and more data from data subjects; there is an urgent need to bring some sort of balance between data holders and data subjects. Since, individual data subjects cannot stand up to the tech giants even if they will, it is necessary that they have a collective voice. Pleasant to the ears of Marxists though it may be, the clarion call of data subjects must be, 'Data subjects of the world, unite!' Still, it is not as communist as it sounds because such ideas are coming from the Free World. The idea is to have a data commons. Its realization would depend on varied jurisprudence of different nation-States.

---

[19] *Google Boots Out Matrimony, Streaming Apps Over Platform Fee Tiff* THE HINDU (New Delhi, Mar. 21, 2024).
[20] *KYC Issues, Money Laundering Concerns Said To Have Led To RBI Order On Paytm's Bank* THE HINDU (New Delhi Feb. 2, 2024).
[21] Data Justice: A Primer on Data and Economic Justice, Report, November 2022, Global Partnership on AI.

In common law countries, data can be pooled into data trusts. Trustees, called by whatever name (data intermediary in the EU or data fiduciary in India) stand in a fiduciary relation with the beneficiaries. In other countries, where trust law is absent or not mature, data banks or cooperatives may be established (for example, Sparkassen in Germany). Whatever the form data stewardship may take, it is essentially about there being some intermediary between the data holder and data subjects. It would ensure that people do not hesitate from sharing their data and that data holder does not keep such enormous data all to himself. This way the societal and economic value of data will be realized. Moreover, bias in the algorithm running the AI as well as the data is fed to the AI would be addressed as it becomes more representative due to more and more data pouring from people who hitherto shied away from giving their data. This widens the scope of data rights from a mere right to access one's data to a right to port it as well. Portability rights allow the data subject to shift his data from the data holder to a data intermediary of his choice. This intermediary will allow access to data only when it in the best interest of the data subject. Without portability rights there can be no meaningful realization of access rights.[22]

It must be noted that there are certain issues which need to be addressed as data commons are brought into execution. There is the issue of cloud interoperability without which there can be no portability rights. Then there is the issue of incentivization of data sharing. As people do not have any encouragement to shift their data to third party intermediaries, they will not budge. Then there are novel ideas of data exchanges which can be prone to speculation as stock or commodity exchanges are. India distinguishes personal and non-personal data. Personal data is which can be used to identify an individual. Non-personal data is which is not personal data. This is not very helping definition. Still, what it is taken to mean in view of the Non-Personal Data Framework is that it can be data which was originally personal data which has now been anonymized or data which never related to any natural person. The difference is one of sensitivity. A person may be more sensitive towards personal data than non-personal data. But AI feeds on both kinds of data. So, the remit of data intermediaries must cover both kinds of data.[23]

---

[22] *Enabling data sharing for social benefit through data trusts: An Interim Report for the 2021* GPAI PARIS SUMMIT.
[23] Anurag Vaishnav, *Report Summary, Non-Personal Data Governance Framework* (Jul. 20, 2020).
Saket Surya, *Report Summary, Revised Draft Non-Personal Data Governance Framework* (Jan.4, 2021).

5. CONCLUSION

In conclusion, the burgeoning global race in artificial intelligence presents the international community with intellectual property challenges of unprecedented complexity. While legislative and judicial bodies worldwide have begun the arduous task of formulating responses to the rapid advancements in AI, the historical trajectory of technological innovation outpacing legal frameworks suggests that definitive resolutions regarding authorship, ownership, and rights attribution remain on a distant horizon. This lag is particularly pronounced in the context of AI, where the speed of technological evolution has seemingly caught the global legal community unprepared. Consequently, stakeholders, including corporations driven by the "strive for more" ethos and individuals whose creative endeavours are increasingly intertwined with AI, must proactively explore and implement strategies to mitigate the uncertainties arising from this technological surge.

Within the realm of AI-generated output, the significance of fundamental legal rights, such as moral rights and the less explored concept of beneficiary rights, becomes acutely apparent. These rights raise foundational questions concerning the true attribution of creative authorship and the equitable distribution of economic and moral benefits. For instance, the determination of whether the original human creator who curated the training data, or the AI system itself, should be recognized as the author of a novel image generated through machine learning algorithms, poses a significant legal quandary. Similarly, the extent to which individuals depicted in AI-generated content should possess control over its use and derive benefits from its dissemination remains a complex and largely unaddressed issue.

The answers to these intricate questions will inevitably be contingent upon a confluence of factors, including the specific nuances of evolving legal frameworks across jurisdictions and the unique circumstances surrounding the creation and exploitation of AI-generated works. The pivotal question of whether artificial intelligence can be legally recognized as an author under existing copyright law is unlikely to yield a swift, universal answer. Instead, it will likely be established through a gradual, case-by-case adjudication process, evolving in tandem with the continuous development of both sophisticated technological capabilities and adaptive legal systems. This dynamic interplay between technological progress and legal interpretation

underscores the protracted and nuanced journey ahead for the global community in navigating the intellectual property landscape shaped by the relentless advancement of artificial intelligence.