Peer - Reviewed & Refereed Journal

The Law Journal strives to provide a platform for discussion of International as well as National Developments in the Field of Law.

DISCLAIMER

No part of this publication may be reproduced or copied in any form by any means without prior written permission of Editor-in-chief of White Black Legal – The Law Journal. The Editorial Team of White Black Legal holds the copyright to all articles contributed to this publication. The views expressed in this publication are purely personal opinions of the authors and do not reflect the views of the Editorial Team of White Black Legal. Though all efforts are made to ensure the accuracy and correctness of the information published, White Black Legal shall not be responsible for any errors caused due to oversight or otherwise.

EDITORIAL
TEAM

Raju Narayana Swamy (IAS) Indian Administrative Service officer

Dr. Raju Narayana Swamy popularly known as Kerala's Anti Corruption Crusader is the All India Topper of the 1991 batch of the IAS and is currently posted as Principal Secretary to the Government of Kerala. He has earned many accolades as he hit against the political-bureaucrat corruption nexus in India. Dr Swamy holds a B.Tech in Computer Science and Engineering from the IIT Madras and a Ph. D. in Cyber Law from Gujarat National Law University . He also has an LLM (Pro) (with specialization in IPR) as well as three PG Diplomas from the National Law University, Delhi- one in Urban Environmental Management and Law, another in Environmental Law and Policy and a third one in Tourism and Environmental Law. He also holds a post-graduate diploma in IPR from the National Law School, Bengaluru and a professional diploma in Public Procurement from the World Bank.

Dr. R. K. Upadhyay

Dr. R. K. Upadhyay is Registrar, University of Kota (Raj.), Dr Upadhyay obtained LLB, LLM degrees from Banaras Hindu University & Phd from university of Kota.He has succesfully completed UGC sponsored M.R.P for the work in the ares of the various prisoners reforms in the state of the Rajasthan.

Senior Editor

Dr. Neha Mishra

Dr. Neha Mishra is Associate Professor & Associate Dean (Scholarships) in Jindal Global Law School, OP Jindal Global University. She was awarded both her PhD degree and Associate Professor & Associate Dean M.A.; LL.B. (University of Delhi); LL.M.; Ph.D. (NLSIU, Bangalore) LLM from National Law School of India University, Bengaluru; she did her LL.B. from Faculty of Law, Delhi University as well as M.A. and B.A. from Hindu College and DCAC from DU respectively. Neha has been a Visiting Fellow, School of Social Work, Michigan State University, 2016 and invited speaker Panelist at Global Conference, Whitney R. Harris World Law Institute, Washington University in St.Louis, 2015.

Ms. Sumiti Ahuja

Ms. Sumiti Ahuja, Assistant Professor, Faculty of Law, University of Delhi, Ms. Sumiti Ahuja completed her LL.M. from the Indian Law Institute with specialization in Criminal Law and Corporate Law, and has over nine years of teaching experience. She has done her LL.B. from the Faculty of Law, University of Delhi. She is currently pursuing Ph.D. in the area of Forensics and Law. Prior to joining the teaching profession, she has worked as Research Assistant for projects funded by different agencies of Govt. of India. She has developed various audio-video teaching modules under UGC e-PG Pathshala programme in the area of Criminology, under the aegis of an MHRD Project. Her areas of interest are Criminal Law, Law of Evidence, Interpretation of Statutes, and Clinical Legal Education.

Dr. Navtika Singh Nautiyal

Dr. Navtika Singh Nautiyal presently working as an Assistant Professor in School of law, Forensic Justice and Policy studies at National Forensic Sciences University, Gandhinagar, Gujarat. She has 9 years of Teaching and Research Experience. She has completed her Philosophy of Doctorate in 'Intercountry adoption laws from Uttranchal University, Dehradun' and LLM from Indian Law Institute, New Delhi.

Dr. Rinu Saraswat

Associate Professor at School of Law, Apex University, Jaipur, M.A, LL.M, Ph.D,

Dr. Rinu have 5 yrs of teaching experience in renowned institutions like Jagannath University and Apex University. Participated in more than 20 national and international seminars and conferences and 5 workshops and training programmes.

Dr. Nitesh Saraswat

E.MBA, LL.M, Ph.D, PGDSAPM
Currently working as Assistant Professor at Law Centre II, Faculty of Law, University of Delhi. Dr. Nitesh have 14 years of Teaching, Administrative and research experience in Renowned Institutions like Amity University, Tata Institute of Social Sciences, Jai Narain Vyas University Jodhpur, Jagannath University and Nirma University.
More than 25 Publications in renowned National and International Journals and has authored a Text book on Cr.P.C and Juvenile Delinquency law.

Subhrajit Chanda

BBA. LL.B. (Hons.) (Amity University, Rajasthan); LL. M. (UPES, Dehradun) (Nottingham Trent University, UK); Ph.D. Candidate (G.D. Goenka University)

Subhrajit did his LL.M. in Sports Law, from Nottingham Trent University of United Kingdoms, with international scholarship provided by university; he has also completed another LL.M. in Energy Law from University of Petroleum and Energy Studies, India. He did his B.B.A.LL.B. (Hons.) focussing on International Trade Law.

ABOUT US

WHITE BLACK LEGAL is an open access, peer-reviewed and refereed journal providededicated to express views on topical legal issues, thereby generating a cross current of ideas on emerging matters. This platform shall also ignite the initiative and desire of young law students to contribute in the field of law. The erudite response of legal luminaries shall be solicited to enable readers to explore challenges that lie before law makers, lawyers and the society at large, in the event of the ever changing social, economic and technological scenario.

With this thought, we hereby present to you

ARTIFICIAL INTELLIGENCE (AI)- THE LIABILITY ANALYSIS

AUTHORED BY - SHAURYA BHADAURIA

ABSTRACT

This paper focuses on the shortcomings of the existing legal systems across the globe besides due acknowledgment of the milestones Artificial Intelligence has achieved in fields including research, medicine, management, analysis, etc. No wonder, Artificial Intelligence (AI) Algorithmic bias and misleading information inspiring most often than not, grave consequences, with zero accountability owing to AI operating as "black boxes" cannot be overlooked. The paper underscores that despite regulations advocating AI and automated system ethics and responsibility besides human supervision yet remains largely unregulated. Additionally, the authorities cited provide an overview of the attempts at ensuring ethical legal frameworks. However, merely persuasive power underscores the spirit of avoiding right violations in the hands of automated systems and AI in consonance with making necessary changes to the existing legal frameworks. However, the paper duly takes into account the nuances arising out of fastening liability akin to humans and those arising out of absolving humans of complete liability for the simple reason of legal violations meted out through automated systems and AI.

KEYWORDS: *Artificial intelligence, algorithmic bias, misleading information, black boxes, legal violations.*

INTRODUCTION

Artificial Intelligence (AI) through active utilization of inputs and algorithms aimed at stimulating human intelligence could be understood to have come to dominate the digital world in our ever-expanding approach to digitization. It has massive implications in technology and vice-versa besides limited discharge in foolproof research as well as law databases.

Although we have come to terms with the wide outreach of AI and its ability to provide immediate and more often than not, reliable information, there lie ample ambiguities as to its accountability under circumstances of violation of legal, constitutional, and fundamental rights besides, to whom the said liability could be traced back to. No matter the need to navigate through various legal systems from across the world could prove fruitful in assessing whether the world at large has been able to fasten accountability and if yes, to what extent, the dimensions of AI need to be appraised carefully and the possible resolutions, if any, which can be adopted to bring in a sense of responsibility having due consideration towards harmonization of the identified legal provisions with the existing legal framework.

AI'S ECONOMIC IMPACT: A DEEP DIVE

While certain sources ascertain a market growth of AI to reach US$6.26bn by the end of 2024 with an annual growth rate (CAGR 2024-2030) of 28.63% as the US records a market size of US$50.16bn by the close of this financial year i.e. 2024[1], certain others estimate AI market growth of US$ 7.8bn by 2025 with a lead by the AI services market per se at a CAGR of 35.8%. Also, the AI software market is expected to grow at a CAGR of 18.1% by 2025.[2]

AI UTILITY

AI has achieved milestones in Disease diagnosis such as in spotting cancer[3], medication research and personalized medicine are three areas where artificial intelligence is finding more and more applications in healthcare. Additionally, chatbots and virtual assistants powered by AI are becoming increasingly popular for enhancing customer support. The processing of AI applications is becoming more efficient and powerful thanks to developments in AI processors and edge computing. And lastly, it is anticipated that the AI industry will experience increased innovation and growth as a result of AI's integration with blockchain and the IoT.

---

[1] 'Artificial Intelligence - India | Statista Market Forecast' (*Statista*, March 2024) <www.statista.com/outlook/tmo/artificial-intelligence/india> accessed 10 July 2024;

[2] Geetika Sachdev, '"India's AI market to reach USD 7.8 billion by 2025," says IDC's latest report on AI' (*IndiaAI*, 31 October 2021) <https://indiaai.gov.in/article/india-s-ai-market-to-reach-usd-7-8-billion-by-2025-says-idc-s-latest-report-on-ai> accessed 10 July 2024;

[3] Martin Stumpe, Technical Lead, and Lily Peng, Product Manager 'Assisting Pathologists in Detecting Cancer with Deep Learning' (*Google Research - Explore Our Latest Research in Science and AI*, 3 March 2017)<https://research.google/blog/assisting-pathologists-in-detecting-cancer-with-deep-learning/> accessed 10 July 2024;

Growth in the AI business is being propelled by multiple sources. To begin, data is essential for AI algorithms to train and improve; so, big data is increasing the number of uses for AI. Second, processing power and the availability of cloud computing are enhancing AI application processing. Thirdly, optimizing and automating processes in industries like transport, finance, and manufacturing is a key driver of AI adoption. The fourth factor driving growth in the AI market is the use of AI in consumer-facing applications such as chatbots and virtual assistants. Lastly, there has been an uptick in investment and collaboration between governments, research institutions, and tech companies to create new AI products and services.

Rising investment in AI R&D, improved AI algorithms and infrastructure, and widespread adoption of AI technologies in industry are all factors that are expected to drive the Artificial Intelligence (AI) market to new heights by 2030. The industry is expected to experience growth as AI becomes more incorporated into both consumer and commercial applications.

NAVIGATING AI LIABILITY CHALLENGES

The black box problem arises in connection to Artificial Intelligence with ambiguity as to the intent of its creator. Moreover, AI is designed to render expert human reason and intuitive judgment inspired by everything including data, pattern evaluation, history, etc. but personal experience in the absence of forthcoming instructions. Thus, the issue arises when we hold it against the notion of holding a speaker liable for providing such information, which in its own opinion it did not believe to be true. The same is problematic since opinion statements not only include facts, sometimes indicating values but also provide opinions as to things and people which in the usual course of action imply there's a basis for believing in it. To fasten liability requires evidence as to the speaker stating something, it otherwise doesn't itself believe to be true and the same can be evaluated through 'intent-based heuristics'- most notably 'scienter' (i.e. recklessness). However, the extent to which the aforesaid device holds utility is quite obscure. Pg 5

Before we fasten liability on a human entity, we more often than not, need not venture into the neural functions taking action unless 'mental unsoundness' is established. On the contrary, AI although designed with cognitive abilities akin to humans with internalized data points to make sound experience-based opinions, cannot match the complexities of neural activity as that of

humans rendering the intent-based heuristics almost inefficacious in determining liability in the case of AI.

ENFORCING TRANSPARENCY: THE LEGAL IMPERATIVE FOR EXPLAINABLE AI

Machine learning algorithms or precisely AI operate as "black boxes," resulting in difficulty in identifying the justifications as to particular decisions, recommendations, or predictions which makes AI subject to challenge.

UNVEILING FALSE FACTS

Fact liability questions such as questions as to intent to mislead are herein dealt with, in the light of Heuristics such as scienter, materiality[4] and reliance[5] concerning the speaker and the context and whether there was enough basis for rendering such an opinion or fact that had a deterministic effect on the person seeking information.[6]

[7]Additionally, as per NIST in its comments upon the element of "Cognitive Bias" in its report "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence (NIST Special Publication 1270" has remarked *"human and systemic institutional and societal factors are significant sources of AI bias as well, and are currently overlooked. Successfully meeting this challenge will require taking all forms of bias into account. This means expanding our perspective beyond the machine learning pipeline to recognize and investigate how this technology is both created within and impacts our society."[8]*

A study published in 2018 by the ACLU (American Civil Liberties Union) highlighted Amazon

---

[4]Wendy Gerwick Couture, 'Materiality and a Theory of Legal Circularity' (2015) 17 University of Pennsylvania Journal of Business Law 3, 453, 455;

[5]Daniel B. Dobbs, 'The Place of Reliance in Fraud' (2006) Vol. 48 Arizona Law Review< https://www.arizonalawreview.org/pdf/48-4/48arizlrev1001.pdf > accessed 11 July 2024;

[6] Yavar Bathaee, 'Artificial Intelligence Opinion Liability' (2020) 35(1) Berkeley Technology Law Journal 113, 122 <www.jstor.org/stable/10.2307/26954421> accessed 11 July 2024;

[7] IBM Data and AI Team, 'Shedding light on AI bias with real-world examples - IBM Blog' (*IBM Blog*, 16 October 2023)<www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/> accessed 11 July 2024;

[8] Reva Schwartz and others, Towards a Standard for Identifying and Managing Bias in Artificial Intelligence (NIST SP 1270, U.S. Department of Commerce 2022)<https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf> accessed 12 July 2024.

placing reliance upon the Reuters report discovered that the "Automated Hiring Tool" equipped with the ability to review resumes enabling automated hiring showed signs of systematic discrimination against women applying for technical positions such as software engineer jobs since a majority of Amazon Software engineers comprised of males and the algorithm was inspired the same demographic of the existing employment structure so much so that it out rightly rejected resumes that included the words "women's." as in "women's chess club captain." Amazon for the same reasons has sought to scrap the same. And what has been noted is *"These tools are not eliminating human bias — they are merely laundering it through software."*[9]

Despite such discriminations, employers' enthusiasm to use AI for hiring is highlighted by other organizations' forging ahead. Automation is allowing companies to explore options outside of the traditional recruiting networks, according to Kevin Parker, CEO of the Salt Lake City-area business HireVue. To cut down on resumes, his company studies video interviews to see how prospects speak and react. *"You weren't going back to the same old places; you weren't going back to just Ivy League schools,"* said Parker. Among his clients are Hilton and Unilever PLC. A new resume analyzing tool developed by Goldman Sachs attempts to place applicants in departments where they would be a "best fit," according to the firm. LinkedIn, owned by Microsoft Corp., the biggest professional network in the world, has gone even further. Companies can use its system to find the best candidates for open positions. But LinkedIn Talent Solutions VP John Jersin insists the platform can't oust human recruiters. *"I certainly would not trust any AI system today to make a hiring decision on its own,"* added the CEO. *"The technology is just not ready yet."* Concerns regarding AI transparency have been voiced by certain groups. The ACLU is presently contesting a statute that permits the criminal punishment of journalists and researchers who conduct tests of employment websites' algorithms for discrimination. According to Rachel Goodman, a staff attorney with the ACLU's Racial Justice Programme, *"We are increasingly focusing on algorithmic fairness as an issue."* It may be very difficult to sue an employer for automatic hiring, though, as Goodman and other AI naysayers pointed out: Prospective employees could not even realize it was being utilized. Amazon, on the other hand, was able to glean some useful information from its botched AI effort. According to one source familiar with the project, some basic tasks,

---

[9] Rachel Goodman, 'Why Amazon's Automated Hiring Tool Discriminated Against Women' (*ACLU* 12 October 2018)<https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against> accessed 11 July 2024;

such as removing duplicate candidate profiles from databases, are now handled by a "much-watered down version" of the recruiting engine. Someone else mentioned that a new group in Edinburgh is trying out automated hiring processes again, but this time they're emphasizing diversity.[10]

ESSENTIALS OF ARTIFICIAL INTELLIGENCE

The aforementioned questions are causing scholars and policy analysts great difficulty. We have developed and publicly reviewed numerous high-minded guiding principles about AI design, development, and usage as a result of this continuing discussion:

- The trust and transparency principles developed by IBM[11]: AI should supplement human intellect, not replace it; trust is essential for adoption; and data policies should be open and easy to understand.

- Google's AI principles[12]: AI should safeguard privacy while also being socially useful, equitable, secure, and answerable to humans.

- A.I. principles at Asilomar[13]: These 23 principles address research, ethics, and values in AI, as well as long-term challenges; they were drafted at the 2017 Asilomar Conference. Elon Musk and the late Stephen Hawking are among the 2,541 interested persons who have joined 1,273 researchers in signing the principles.

- The principles upheld by the Partnership on AI (PAI)[14]: Eight principles for a welcoming space to talk about AI ethics, trust, explainability, and the social duty of AI delivery firms. These principles must be acknowledged by every partner who wants to become a part of the PAI.

---

[10] Jeffrey Dastin, 'Insight- Amazon scraps secret AI recruiting tool that showed bias against women' (Reuters, 11 October 2018) <www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/> accessed 12 July 2024.

[11] Seth Dobrin Chief AI Officer and Christina Montgomery Chief Privacy Officer & AI Ethics Board Co-Chair, 'Principles and Practices for Building More Trustworthy AI' (IBM Newsroom, 21 October 2021) <https://newsroom.ibm.com/Principles-and-Practices-for-Building-More-Trustworthy-AI> accessed 12 July 2024;

[12] 'AI Principles Progress Update 2023' (Making AI helpful for everyone - Google AI) <https://ai.google/static/documents/ai-principles-2023-progress-update.pdf> accessed 12 July 2024;

[13] 'Asilomar AI Principles - Future of Life Institute' (Future of Life Institute, 11 August 2017) <https://futureoflife.org/open-letter/ai-principles/> accessed 12 July 2024;

[14] 'Ethical Principles and Practices for Inclusive AI' (Partnership on AI, 20 July 2022) <https://partnershiponai.org/wp-content/uploads/dlm_uploads/2022/07/PAI_summary_making-AI-inclusive.pdf> accessed 12 July 2024;

- AI4PEOPLE's guiding principles and standards[15]: Real suggestions for European politicians to help AI develop throughout the continent.

- Ethical AI principles put out by the World Economic Forum[16]: Five guiding principles that address the following: AI's intended use, AI's fairness and intelligence, data protection, everyone's right to benefit from AI, and the prohibition of autonomous weaponry.

- IEEE (Institute of Electrical and Electronics Engineers)[17]: a collection of guidelines that situate AI within a human rights framework; these guidelines touch on topics including ethical AI, corporate responsibility, value by design, accountability, and wellness.

CHARTING EUROPE'S AI STRATEGY AND APPROACH

[18]EU's AI strategy has been underway since April 10, 2018, with the signing of "The Declaration of Cooperation on Artificial Intelligence aimed at ensuring competitiveness and answering social economic, ethical, and legal questions, followed by the adoption of "Communication on Artificial Intelligence" towards ensuring ethical and legal framework being in place as part of regulatory norms; appointment of AI HLEG (High Level Expert Group) comprising experts acting in advisory capacity for implementation of AI strategy; presentation of a "Coordinated Plan on AI" with constructive attempts being underway to develop and strengthen AI Learning and training programs, laying guidelines upholding ethics and a legal framework that doesn't discourage innovation. The "Ethics Guidelines for Trustworthy AI" and "Policy and Investment Recommendations for Trustworthy AI" were published in 2019 by HLEG outlining seven criteria that ought to be met to ensure AI trustworthiness and over thirty-three recommendations directed towards a coordinated plan for making AI sustainable, growth-oriented and competitiveness friendly. EU also issued the "Communication on Building Trust in Human Centric AI" enforcing "human agency and oversight; technical promptness and safety; privacy and data governance;

[15] 'AI4 People's Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations' (AI4People – Assessing AI Risk, 28 November 2019 <https://ai4people.org/PDF/AI4People_Ethical_Framework_For_A_Good_AI_Society.pdf> accessed 12 July 2024;

[16] Rob Smith, '5 core principles to keep AI ethical' (World Economic Forum, 19 April 2018) <www.weforum.org/agenda/2018/04/keep-calm-and-make-ai-ethical/> accessed 12 July 2024;

[17] Changwu Huang and others, 'An Overview of Artificial Intelligence Ethics' (2023) 4(4) IEEE Transactions on Artificial Intelligence 10.1109/TAI.2022.3194503 <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9844014> accessed 13 July 2024;

[18] Eric Brattberg, Raluca Csernatoni, and Venesa Rugova, Assessing the EU's Approach To AI (Carnegie Endowment for International Peace 2020)<www.jstor.org/stable/resrep25784.7> accessed 13 July 2024;

transparency; diversity, nondiscrimination, and fairness; societal and environmental well-being; accountability," respectively. The same was required to be followed by all parties involved in the communication, including users, providers, and developers of AI. Fast forward to April 2021, the EU artificial intelligence regulation framework was suggested by the European Commission in April 2021. What this means is that potential AI systems are evaluated and categorized based on the level of risk they represent to end users. More or less regulation will be imposed depending on the level of risk. In March 2024, the European Parliament adopted the AI Act, the first regulation on AI also adopted by the EU.

Analysis of EU's AI Strategy

No wonder, despite regulations being in place, the majority of those have retained a mere directory character amidst a non-binding framework. However, the message is given out quite clearly wherein the communication has attempted to underscore AI-related threats, unlike other nations that have stayed more inclined towards developing and advocating an opportunity-friendly approach.

To facilitate and provide for a favorable regulatory environment, what's important is to strike a balance between regulating and providing enough incentives for development. Also with all the regulation in place, Europe appears to be well positioned in terms of vision and the way forward built upon acceptable global standards. However, there is a need for actual deterrence and not mere direction.

BALANCING INNOVATION AND RIGHTS: AI'S IMPACT ON HUMAN RIGHTS

AI accountability towards human rights becomes indispensable with a delegation of greater responsibility[19]. The technologies being actively employed for ensuring security against criminal conduct and otherwise by governments across the globe to track down terrorism, fraud, threats, etc, through biometric authentications and video surveillance are being well exploited to even

---

[19] Gary R Lea, 'Constructivism and its risks in artificial intelligence' (2020) 36(4) Prometheus 322, 336 <www.jstor.org/stable/10.13169/prometheus.36.4.0322> accessed 13 July 2024;

monitor and track the general populace, thereby posing a threat to privacy.[20]

With efforts being underway for the social legitimization of technological advancement wherein human rights and dignity continually inspire technology, legitimization instead of solely being viewed in the light of security or economic efficiency would not suffice. Equal consideration ought to be had for democracy and the dignity of people. The Universal Declaration of Human Rights (UDHR) as adopted by the United Nations General Assembly enshrines the rights and freedoms of all human beings. Moreover, 'The Magna Carta' was one of the world's first attempts the form a document comprising the sovereign's commitments towards protecting people against violations of certain legal rights, thereby, ensuring respect for all.

"Algorithm prejudices" inspire discrimination and bias and disrupt equality[21]. For instance, in the US, investigations resulting in criminal sentencing widely use the "COMPAS Algorithm"[22] that has come to be widely challenged. The journalism group, 'ProPublica' upon an independent analysis concluded its findings in a publication titled "Machine Bias".

It was noted that of the 18,610 defendants, persons likely to re-offend (i.e. at a score of 5 or higher) were calculated to be 61% and were further detained so as to prevent subsequent crimes within two years. However, it was only 20 percent of the people deemed likely to "commit violent crimes went on to do so." For the same reason the result was "remarkably unreliable" as were racially biased with black defendants deemed 4.5 percent more likely than white defendants to re-offend. And as can be noted, where purpose is compromised, justification to such discrimination cannot be relied upon either, and for legitimate reasons being in place, where disproportionate means are resorted to, the conduct yields illegitimate discrimination. The burden of proof shall lie upon the state to prove the legitimacy of discrimination which is hard to come by if the discrimination results from the conduct of Artificial Intelligence (AI).

---

[20] Cataleta and Maria Stefania, Humane Artificial Intelligence: The Fragility of Human Rights Facing AI (East-West Center 2020) <www.jstor.org/stable/resrep25514> accessed 13 July 2024;

[21] Silva, Selena and Martin Kenney, 'Algorithms, Platforms, and Ethnic Bias: An Integrative Essay' (2018) 55(1 & 2) Phylon (1960-) 9, 29<www.jstor.org/stable/10.2307/26545017> accessed 13 July 2024;

[22] Tafari Mbadiwe, 'Algorithmic Injustice' (2018) (54) The New Atlantis 3, <www.jstor.org/stable/10.2307/90021005> accessed 13 July 2024;

[23]The "Correctional Offender Management Profiling for Alternative Sanction" (COMPAS) algorithm utilizes 137 questionnaire items to assess criminal activity, relationships, personality, family, and social isolation.16 Critics argue that the algorithm discriminates based on race, resulting in unfair treatment of people of color.

Similarly, it produces a bias in crime probability, affecting people of color twice as often as those with lighter complexion. The Laura and John Arnold Foundation developed the Public Safety Assessment (PSA) to mitigate the discriminatory consequences of COMPAS. This technique eliminates unfavorable impacts based on gender, race, or economic conditions. This algorithm uses nine risk indicators to predict if an individual will attend trial and commit an offense if freed before trial.17 Reduced prejudice is possible when criminal convictions have a stronger impact than other assessments and factors. The PSA would evaluate race impartially and report to the judge.

The European Union came out with its first draft of the "Ethics Guidelines for Trustworthy AI"[24] in December 2018 followed by its coming into force on 8th April 2028. No wonder, the guidelines provided that trustworthy AI implies AI being "Lawful," "Ethical" and "Robust" while advocating for human agency and oversight; technical robustness; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being; and accountability, but continued to persist as a directory provision.

THE AMERICAN SYSTEM- CONTRACT AND TORT DISPUTES

[25]The American gradual, common law adjudication system is seeing the emergence of AI as the latter is gaining relevance. Designing a car with autonomous driving capabilities can lead to both familiar and unusual contract and tort conflicts.[26] New types of legal disputes, such as those

[23] Jeff Larson and others, 'How We Analyzed the COMPAS Recidivism Algorithm' (ProPublica, 23 May 2016) <www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm> accessed 14 July 2024;
[24] 'Ethics guidelines for trustworthy AI' (European Commission- Shaping Europe's digital future, 8 April 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> accessed 14 July 2024;
[25] Cuéllar and Mariano-Florentino, 'A Common Law For The Age Of Artificial Intelligence: Incremental Adjudication, Institutions, And Relational Non-Arbitrariness' (2019) 119(7) Columbia Law Review 1773, <www.jstor.org/stable/26810848> accessed 13 July 2024;
[26] Geistfeld and Mark A, 'A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation' (2017) 105(6) California Law Review 1611, <www.jstor.org/stable/44630767> accessed 14 July 2024;

involving the responsibility of a party for harm and the subtle differences revealed by digital evidence of neural network evolution, will arise as reliance on AI grows more commonplace, similar to the challenges courts faced when translating traditional concepts like chattel trespass to cyberspace.[27] A slew of new lawsuits in American courts will likely center on the validity of drivers' and organizations' decisions to trust driverless vehicles or to train neural networks to solve complex health and safety problems.[28] Thus, there seems to be an intersection between AI, common law, and society.

In this context, "artificial intelligence" refers to information technology that can carry out tasks that would normally be associated with human intelligence, and which can also produce results that laypeople may have faith in. This approach of framing the word comprises both domain-specific applications that execute tasks like financial analysis or autonomous driving and systems that try to emulate general intelligence through conversation or analytical capabilities across domains. Although computer science and statistics are certainly involved in the machine learning techniques at the core of certain AI applications, this explanation also centers on the presence of a separation between AI and ordinary statistical inference.

Anglo-American tort law, guided by more flexible concepts of proximate causality, foreseeability, and responsibility creates a more adaptive framework in contrast to numerous types of rigid regulation that aim to make individuals and organizations face the social costs of their choices and actions. The resolution of cases in this area often requires a balance between applying established principles and allowing the doctrine to adjust to new situations. This is because proximate causation and related foreseeability inquiries are designed to be flexible enough to account for changing social, technological, and economic conditions. Some complex flexibility-fidelity trade-offs will arise for our profession as a result of evaluating these decisions.

[29]The process of driving or the detection of suspicious transactions both include trade-offs. The

---

[27] IntelCorp. v. Hamidi, Supreme Court of California, 30 June 2003, S103781, P.3d, 71,296,308 (California)<https://law.justia.com/cases/california/court-of-appeal/4th/94/325.html> accessed 15 July 2024;

[28] Cuéllar and Mariano-Florentino, 'A Common Law For The Age Of Artificial Intelligence: Incremental Adjudication, Institutions, And Relational Non-Arbitrariness' (2019) 119(7) Columbia Law Review 1773, <www.jstor.org/stable/26810848> accessed 13 July 2024;

[29] Bryant Walker Smith, 'Automated Driving and Product Liability' (2017) (1) MICH. ST. L. REV. 32;

conventional doctrinal inquiry becomes more challenging when proximate causation evaluations rely on AI-driven decision rationales that are not always understandable or explainable. This is because, at the very least, it is not always apparent how justifiable it is for an individual to depend on a specific decision-making technology.[30] Also, accidents resulting from the conduct of autonomous systems defy the originally realized "fault and agency" in motor accidents.[31] The complex design decisions impacting AI systems' ability to communicate with people and derive "justifications" for decisions from systems like artificial neural networks will determine the relevance of reasonableness in a future where these systems are increasingly ubiquitous. This problem is comparable to the ones that arise when artificial intelligence is asked to evaluate vague laws, such as the Administrative Procedure Act, or fundamental constitutional principles, such as reasonable suspicion. One way to ensure that machine decision-making (or decision-support) systems align with our goals is to hold machine answers to a standard of "relational non-arbitrariness." This concept is connected to Ashley Deeks' xAI and could be used to illustrate Kate Strandburg's observation that human decision-making is often collaborative.[32]

Fair decision-making calls for the evaluation of both public organization as well as public institution stakeholders in decision-making since it is a primary concern that decision costs ought not to override the intended benefit of policies and law. Thus, it is taken into consideration whether first, there is a foundation for a decision taken by a human post extensive consultation with an AI system or by the system itself, in theory, so that we may claim the decision is not arbitrary. The second reason is that there are opposing values and complexities in the analysis, therefore it's important to examine if the human-machine connection reflects this. Thirdly, it prompts the question of whether decision-making processes encourage additional discussion of the decision by those community members who are associated with or impacted by it.

Optimal presentation of information to persuade the reviewing authority to accept the applicable reason and simultaneous reduction of an underlying cost function might occur in an AI system.

[30] Jack Boeglin, 'The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation' (2015) 17 Yale J.L. & Tech. 174<https://yjolt.org/sites/default/files/boeglin_the_costs_of_self-driving_cars_0.pdf> accessed 14 July 2024;

[32] Ashley Deeks, 'The Judicial Demand For Explainable Artificial Intelligence' (2019) 119(7) Columbia Law Review 1829 <www.jstor.org/stable/10.2307/26810851> accessed 14 July 2024;

However, the end goal ought not to be to weigh justifications for public and private action but to enable logic and justification in some way. And merely because "reasonableness" debates start in a different doctrinal setting in tort law doesn't mean they don't serve the same purpose: to let us evaluate how a citizen (whether or not using AI) explains her behavior about a more generally accepted norm of behavior and to let us think about how that norm should evolve.

All in all, if one is to use an ideal to evaluate the use of AI in decision-making, it would be to prioritize the views of networks of decision-makers over those of individual decision-makers, thus addressing the often-implicit legal worry. In keeping with the strong emphasis on reason-giving and justification in both public law and common law traditions, it is crucial to prioritize justifications that can be supported by human networks that include principled and reasonable discussions. These networks are intended to determine whether certain reasons are valid enough to justify the use of force or the rejection of a presumed duty of care that members of a community have towards each other.

TACKLING CONCERNS: ARTIFICIAL INTELLIGENCE, DEEPFAKES, AND DISINFORMATION

[33]New developments in AI and computer science have given rise to a powerful new tool for spreading false information: deepfakes. According to Merriam-Webster (undated-a), "deepfake videos" are videos that have been digitally manipulated to make the subject look like someone or something else. A lot of people are worried that this technology will make fake news from outside and inside the country much more dangerous because these movies are getting more realistic. For numerous women, this danger has come to pass as a result of pornography sites that use artificial intelligence (Jankowicz et al., 2021). But in other respects, the mayhem-inducing possibilities have not yet materialized. Some analysts have even gone so far as to predict that a deepfake video may disrupt the 2020 election. The fact that the deepfakes did not materialize does not mean that future elections would not be vulnerable (Simonite, 2020). The spread of deepfake and similar AI-generated disinformation has come at a particularly delicate moment for the global community and the US in particular. Truth Decay: An Initial Exploration of the Diminishing Role of Facts and

---

[33] Todd C. Helmus, Artificial Intelligence, Deepfakes, and Disinformation A Primer (2022) <www.jstor.org/stable/resrep42027> accessed 14 July 2024.

Analysis in American Public Life (2018), a seminal report by RAND colleagues Jennifer Kavanagh and Michael D. Rich, presents four trends that, taken together, indicate that truth is losing its significance in American society. These trends include a growing divide between factual assessments and analytical interpretations of data and facts; a blurring of the boundary between fact and opinion; an increase in the relative volume and influence of personal experience and opinion over facts; and a decrease in trust in once respected sources of factual information.

So long as these tendencies persist, deepfakes will continue to target people who are easily fooled.

DETECTION OF DEEPFAKES

Detection of Deepfakes requires rapid implementation of automated systems such as the GAN system which can distinguish between actual and fake images besides the Media Forensics (MediFor) program and the Semantic Forensics (SemaFor) program. However, there's a fundamental limit to such detection by a particular detector. Additionally, Initiatives such as the "Deepfake Challenge Competition," with over 2,000 participants advanced tested models for such detection but with about 65 percent accurate detections against the "black box dataset" and 82 percent accuracy against a public data set of deepfakes.

To bolster detection efforts, social media platforms should make available their extensive image collection, which includes synthetic media. The training data stored in these repositories could be used to update detection programs on the latest developments in deepfake offspring. Another strategy is to make "radioactive" training data that deepfake makers could employ to have their content recognition programs notice it. Any "model trained on [these data] will bear an identifiable mark" since the data has been "imperceptibly changed" with radioactive elements. No wonder, a ton of strategies could be put into action to achieve greater AI functioning and reliance. Yet, the question as to legal backing remains unanswered across the globe.

ENVISIONING THE FUTURE: AUTONOMOUS SYSTEMS
AS LEGAL PERSONS

Is there going to be a practical legal framework for addressing liability issues if autonomous systems are given legal personhood? New entities can be established within the framework of the

law since it is adaptable. The report states the following as the question of the European Parliament regarding this issue:

(T.) , ultimately, robots' autonomy raises the question of their nature in the light of the existing legal categories –of whether they should be regarded as natural persons, animals, or objects –or whether a new category should be created, with its specific features and implications as regards the attribution of rights and duties, including liability for damage;[34]

European lawmakers are considering expanding the concept of corporate liability to include autonomous systems. We must exercise caution and consider the issues, such as misuse: brands utilizing autonomous systems may find it advantageous to hold the system accountable and then disappear without paying damages.

In one such instance on March 18, 2018, at around 9:58 a.m., a 49-year-old lady was killed in an accident on Mill Avenue in Phoenix, Arizona, while riding in an Uber test vehicle that had software from Volvo. The collision was fatal for the woman. A software glitch was most likely the cause of the fatal accident caused by the self-driving test car, according to the National Transportation Safety Board.[35]

Questions like these take on more substance when we apply the concept of legal persons to the Uber scenario that was previously mentioned. So, who's going to be the autonomous system's lawyer? Do you mean Volvo or Uber? The difficult question to ask is whether, if the self-driving car is considered a legal person with, in this example, Uber and Volvo representatives, this implies a distribution of control and culpability that extends to these actors. It is critical to address this issue from a practical and legal standpoint to close the gap between human control and the growing autonomy of systems.

---

[34] Mady Delvaux, DRAFT REPORT with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) (PE582.443v01-00, 2014)<www.europarl.europa.eu/doceo/document/JURI-PR-582443_EN.pdf?redirect> accessed 15 July 2024;

[35] Highway Accident Report- Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian Tempe, Arizona (NTSB/HAR-19/03 PB2019-101402, National Transportation Safety Board 2018)<www.ntsb.gov/investigations/AccidentReports/Reports/HAR1903.pdf> accessed 15 July 2024;

CORRUPT AI CANNOT ABSOLVE HUMAN LIABILITY

Corruption occurs when those in positions of authority, whether in the public or private sector, misuse AI systems. Corruption in artificial intelligence presents unique challenges because power is fundamental to the crime. Those who control the data and the code tend to have more influence in today's increasingly digital society. Therefore, AI has the potential to solidify and worsen preexisting power discrepancies. For instance, victims have little leverage to end corruption since they gain from it, while those in positions of authority typically lack the motivation to do so. For example, dishonest data scientists may manipulate the algorithms to benefit themselves, their clique, or the well-off if they were responsible for developing an AI-based system that could predict a patient's survival rate. On the other side, when powerful people get access to AI, this trend will only get worse. It takes the malevolent design of an AI system for it to be corrupted. It can also happen when existing AI systems have their weaknesses taken advantage of, even though the systems are generally useful. When AI gets involved in corrupt acts, the limitations on corruption are lessened because this corrupt behavior is less likely to be detected and sanctioned due to the diffusion of responsibility. A well-known occurrence in the field of behavioral study is the diffusion of responsibility.[36] Furthermore, it is extremely challenging to demonstrate unambiguous responsibility when AI is utilized for corrupt purposes.[37]

SUGGESTIONS

There appears to be a dearth of regulatory standards and frameworks describing various concepts of ethical AI, such as openness, justice, and responsibility, even though organizations like the OECD, UNESCO, and the High-Level Expert Group on AI of the European Commission have proposed such guidelines. Making the factors that impact an AI system's decision-making process visible to the appropriate parties is one way to implement data and code transparency through code audits that would make data and code publicly available.[38] Implementing thorough and impartial

---

[36] Darley and others, 'Bystander intervention in emergencies: diffusion of responsibility' (1968) 8(4) Journal of Personality and Social Psychology 377 <www.neshaminy.org/site/handlers/filedownload.ashx?moduleinstanceid=34368&dataid=51119&FileName=Article%2010-%20To%20Help%20or%20Not%20to%20Help.pdf> accessed 15 July 2024;

[37] Nils Christopher Köbis, Christopher Starke and Jaselle Edward-Gill, The Corruption Risks of Artificial Intelligence (Transparency International 2022) <www.jstor.org/stable/resrep43028.> accessed 16 July 2024;

[38] Niklas Kossow, Svea Windwehr and Matthew Jenkins, Algorithmic transparency and accountability. (Transparency International 2021)<www.jstor.org/stable/pdf/resrep30838.pdf> accessed 15 July 2024;

audits is another technique that shows promise.[39] Algorithms can be audited independently to make sure they follow the ethical standards in regulations. institutions like Algorithmic Justice League and Algorithm Watch are good examples of such institutions. They can create protections against the accidental repercussions of AI implementation in social settings and the deliberate abuse of AI for corrupt and other unlawful and immoral purposes. A further technical aspect that can lessen the likelihood of AI corruption is the facilitation of such code audits. Making machine learning programming more compatible with one another is one practical step in this direction. The two most popular languages used by data scientists to create machine learning models are TensorFlow and PyTorch.

In addition, code auditors and data scientists are now vital players in the AI ecosystem. This meteoric rise to prominence has occurred in the business and public sectors at a dizzying rate. Unlike more traditional professions associated with authority, such as law enforcement, medicine, and politics, there are no established standards of conduct, much less anti-corruption ones. Simultaneously, a widespread idea for ensuring responsible and ethical AI is to provide programmers and data scientists with ethics training.[40] A lot of times, people get supplanted by AI systems. The capacity of institutions to protect whistleblowers is dwindling as a result of firms' trend towards using fewer humans to supervise critical tasks. An individual must act independently to blow the whistle on their employer. Incentives for artificial intelligence algorithms are more likely to be in line with those of the organization or business that uses them. Consequently, there is currently no way for AI systems to report internally or blow the whistle. Here, two things are happening. To start, fewer people will be able to report if AI is used instead of humans. Second, the remaining whistleblowers may feel less secure and eager to come forward if AI technologies are included. Because people may not suspect that AI systems can rebel, they may have (very) positive assessments of their performance. No whistle-blowing capabilities are further diminished by the fact that AI algorithmic procedures are frequently opaque. If we want individuals to continue speaking out against (AI) corruption, we need to make sure they know about these two cutbacks in reporting and whistle-blowing skills.

---

[39] James Guszcza and others, 'Why We Need to Audit Algorithms' [2018] Harvard Business Review <https://hbr.org/2018/11/why-we-need-to-audit-algorithms> accessed 15 July 2024;

[40] Rob Reich, Mehran Sahami, and Jeremy M. Weinstein, System Error: Where Big Tech Went Wrong and How We Can Reboot (HarperCollins Publishers 2021);

CONCLUSION

With the introduction of the world to AI comes the promise of overreaching advancement goals enough to surpass human intelligence and with that a need is realized for the development of AI-specific legal framework thus, avoiding being bent upon long-existing laws and doctrines designed to assess human conduct. However, AI's ability to provide opinion statements creates a fiction as to its bona fide belief in the said information. Such information if found false suggests the inauthentic source of knowledge and underlying facts and could have disastrous repercussions. Thus, where a right is violated and injury is caused, liability arises in the usual course of action but the question as to accountability about the injury resulting from any act of AI has continued to stay unanswered across nations. Nevertheless, there has been implementation of guidelines, frameworks, and regulations governing AI over the past decades yet the mere directory nature of such rules and regulations without the actual law backed by the sanction of the state proves no good in the wake of ample instances of AI malfunctions resulting from the accumulation of huge data and auto interpretation to suit trends which are most often than not, discriminatory, inspired by bad faith, incorrect sources of information, so on and so forth.