GNLU LAW AND SOCIETY REVIEW

Gujarat National Law University

Volume II

Algorithms, Law and Democracy

2020

GNLU LAW & SOCIETY REVIEW

VOLUME II

"ALGORITHMS, LAW & DEMOCRACY"

ISSN: 2582-2446

Disclaimer:

The views expressed in the articles, essays, case comments, reviews, published in the GNLU Law & Society Review are those of the individual author/s and do not represent the views of the editorial board of the Review. No part of this work may be produced and stored in a retrieved system or transmitted in any form or by any means, electronic, mechanical, photocopying recording or otherwise without written permission from the publisher.

Published by:

Centre for Law & Society,
Gujarat National Law University,
Attalika Avenue, Koba, Gandhinagar – 382436,
Gujarat, India
E. glsr@gnlu.ac.in

ADVISORY BOARD

EDITORIAL BOARD

CONTENTS

EDITORIAL NOTE

-Apoorva Patel[*]

It gives me immense pleasure to present the reader with the special issue (Volume II) of the GNLU Law & Society Review on the theme '*Algorithms, Law, and Democracy*'.

The Law & Society Review was established in 2018 under the GNLU Centre for Law and Society with an aim to provide a platform where socio-legal issues are studied, analyzed, scrutinized, and criticized using interdisciplinary and multidisciplinary methods. The Review is published annually, and is a collaborative faculty cum student led, double-blind peer reviewed to maintain highest quality and standards in publication. The first volume was published in 2019 and it saw contributions from a diverse set of scholars writing on topical issues concerning status of Israeli-Arabs, ethical states, proposals for studying the impact of judicial decisions, need for social boycott laws to name a few. Our aim for the first issue was to provide alternatives to an 'isolationist' study of the law that remains blind to pretext and subtext of law which more often than not is wedded to social realities of the day. Using interdisciplinary and multidisciplinary tools in legal research allows us to see law not as an end in itself but a culmination of socio-economic and political processes.

Volume II of the Review is yet another step towards this objective. Algorithms have changed governance fundamentally. Its full-scale consequences are yet to be seen in laws, legal regimes and legal processes, gig-economy and in democracy in general. The special issue comprises of articles written by a host of scholars from India and around the world and students on this emerging but not adequately studied domain, especially in India. We hope the dearth in literature on issues arising from the use of algorithms finds address through this special issue - a small and modest contribution to an ever expanding realm.

On behalf of the Review, I would like to take this opportunity to extend a deep gratitude to Prof. (Dr.) S. Shanthakumar (Director, GNLU) for facilitating a very encouraging ecosystem and for his constant guidance and support. I am also grateful to Dr. A. N. Rao (Dean, Research & Publications Division), Dr. Saurabh Anand (HoD Social Sciences), Dr. Hardik Parikh and Mr. Sanjeev Choudhary, faculty members of the Centre, for their continued support and encouragement. In the same vein, I would like to express my thanks to Dr. Naveen Thayyil (Associate Professor of Law & STS, IIT-Delhi) for his immense contribution to the special issue

[*] Director, GNLU Centre for Law and Society; Editor-in-Chief, GNLU Law & Society Review.

as the Guest Editor. I would be remiss if I do not mention my special thanks to the Centre's Student Convener and Review's Managing Editor, Mr. Zaid Deva for his commitment and hard work, along with expressing gratitude to the student editorial board for anchoring the journal ably and making it a real possibility to bring out the Review in time and to realize the goals and objectives of the Review as well as the Centre for Law and Society.

RIGHTS, CITIZENSHIP AND SELFHOOD: THINKING THROUGH ALGORITHMS, LAW AND DEMOCRACY

- Naveen Thayyil[*]

Even as commonsense treats law as an algorithm for (liberal) democracy, the "foreseeably pervasive" deployment of high-tech algorithm,[1] is met with concerns for law and about democracy. The ostensible algorithmic turn in ordering society, sometimes in fundamental ways, raises questions about if and how law can respond to the social and technological changes that are attendant to the development and deployment of these technological trajectories. Can the attention on algorithms and law, including a silent facilitation of the pervasive deployment of algorithms in the public, help us understand the values inhabiting modern law better?

Oxford dictionary traces the etymology of the term algorithm to the Arabic source, "al-K̲wārizmī 'the man of K̲wārizm', a name given to the 9th century mathematician Abū Ja'far Muhammad ibn Mūsa, author of widely translated works on algebra and arithmetic.[2] While algorithms are generally recognized as a 'precise, step-by-step procedure that requires, in and of itself, no human intuition or guesswork',[3] we are concerned here with algorithms – in a reportative sense – with respect to the development and deployment of highly complex and advanced algorithms through the explosion of computing power, digital technologies, as well as big data and machine learning techniques. Much of the current attention on algorithms is connected to the supposed inexorable move towards (general) artificial intelligence, through machine learning algorithms in the 'frontier' fields of robotics, computer vision, speech recognition and natural language processing that is seen as inextricably altering social and human agency, while affecting law and its practice in fundamental ways.[4]

Through the last decade, there has been growing attention around how such algorithms, and the attendant technological trajectories that are generally referred to simply as AI, have grave effects on rights in various realms of traditional law - be it its effects on free speech, on discrimination,

[*] Associate Professor (Law and STS) at IIT Delhi.

[1] Paul Nemitz, *Constitutional Democracy and Technology in the age of Artificial Intelligence*, 376(2133) PHIL. TRANS. R. SOC. A. 1-14 (2018) [hereinafter Nemitz].

[2] OXFORD LEARNER'S DICTIONARY, https://www.oxfordlearnersdictionaries.com/definition/english/algorithm (last visited June 2, 2020).

[3] JOHN MACCORMICK, NINE ALGORITHMS THAT CHANGED THE FUTURE 3 (Princeton University Press 2012).

[4] See generally JERRY KAPLAN, ARTIFICIAL INTELLIGENCE (Oxford University Press 2016). Also see chapter 5, '*AI and the Law*'.

privacy and data security, control and surveillance, personality, causality, responsibility and liability, IPR, labour laws, and even the delivery of justice. Justice Sales aptly encapsulates these concerns here:

> *"How should legal doctrine adapt to accommodate the new world, in which so many social functions are speeded up, made more efficient, but also made more impersonal by algorithmic computing processes? At least with computer algorithms, one still has human agency in the background, guiding processes through admittedly complex computer programming. Still more profoundly, however, how should legal doctrine adapt to processes governed without human agency, by AI."*[5]

A quick sweep of law journals gives a glimpse of the different aspects that the legal world has identified as matters of concerns with respect to the pervasive deployment of such algorithms, and include threats to existing notions and rights about privacy, personal anonymity and individual autonomy as preconditions and levers of democracy, including threat to decisional and informational privacy,[6] and behavioral patterning as predictive technologies employing algorithms; legal protection against discrimination through algorithms and by AI,[7] including algorithmic bias in predictive policing and the insurance sector;[8] considerations about technological autonomy,[9] and challenges to asserting social control over technology, including the recognition of non-human algorithmic entities as legal persons and 'the virtual impossibility of governmental regulation of algorithmic control of regulated entities';[10] the implications of algorithmic decision-making to free speech protections,[11] including in realms like algorithmic editing,[12] implications of personalized pricing using algorithmic models,[13] and protection of well recognized rights for workers.[14] Even as grave concerns about the detrimental effects on the

[5] Philip Sales, *Algorithms, Artificial Intelligence and the Law*, THE SIR HENRY BROOKE LECTURE (Nov. 12, 2019) https://www.supremecourt.uk/docs/speech-191112.pdf (last visited June 19, 2020) [hereinafter Sales].

[6] Karl Manheim and Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J. L. & TECH. 106 (2019).

[7] Frederik J. Zuiderveen Borgesius, *Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence*, INT. J. HUM. RIGHTS. 8 (2020). That 'Algorithmic Decision Systems can lead to discrimination has been extensively documented in many areas, such as the judicial system, credit scoring, targeted advertising and employment'. See *Opportunities and Challenges in Algorithmic decision making*, EUROPEAN PARLIAMENT SCIENTIFIC FORESIGHT UNIT (STOA) (Panel for the Future of Science and Technology) https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf (last visited June 8, 2020) [hereinafter STOA].

[8] *Id.* STOA.

[9] Alex Ingrams, *Big Data and Dahl's Challenge of Democratic Governance*, 36(3) REV. POLICY RESEARCH 357-378 (2019).

[10] Lynn M. LoPucki, *Algorithmic Entities*, 95(4) WASHINGTON UNIV. L. REV. 887, 901 (2018).

[11] John O. McGinnis and Steven Wasick, *Law's Algorithm*, 66 FLORIDA L. REV. 991-1050 (2014).

[12] Stuart Minor Benjamin, *Algorithms and Speech*, 161(6) UNIV. OF PENN. L. REV. 1445-1494 (2013).

[13] Joshua A Gerlick and Stephan M Liozu, *Ethical and Legal considerations of Artificial Intelligence and Algorithmic decision-making in personalized pricing*, 19 J. REVENUE & PRICING MANAGEMENT 85–98 (2020).

[14] See the special issue on *Automation, Artificial Intelligence, & Labor Law*, *in the Comparative Labor Law & Policy Journal*, Volume 41(1), 2019.

realization of fundamental rights are being raised, some commentators also point out the possibilities of using algorithms, AI and machine learning as catalysts to revive and achieve 'democratic ideals',[15] calls for 'structuring the emerging market for AI justice' since it offers efficiency and 'at least an appearance of impartiality', fostering a turn towards codified justice that favors a paradigm of standardization above discretionary moral judgment.[16] These different responses and articulations also reveal the deep differences that are prevalent in society about public values regarding the development and deployment of such technologies.[17]

Accompanying these aforementioned responses are attendant questions about the ways in which law should foster regulation of algorithms and AI technologies, and how much legal regulation is desirable, and indeed possible. Fitsilis focuses on the major cases where regulation has currently been imposed upon algorithms, involving eight areas of law - competition, labor, environment, IPR, data-protection, consumer protection (and EU internal market).[18] The norms, principles and modalities in these cases may then become an early framework around which legal responses to these technologies as also democratic concerns about such responses may be deliberated upon. The limitations of these legal responses itself then may become starting points for analysts. For instance, administrative law principles of transparency, accountability and nondiscrimination have been employed in much of these legal interventions to assert public regulation. What are the limitations with this approach require further analysis and consideration.[19]

Importantly, further, concerns include possibilities of the slippery slope that these technologies will make obsolete even the very architecture through which fundamental rights and values are recognized as valuable. This is related to the concern and 'possibility that emerging technologies may change our moral and ethical considerations' – sometimes articulated as the habituation argument 'that although at present the new technology is in conflict with established morals,

[15] Emmanuel Letouzé & David Sangokoya, *Leveraging Algorithms for Positive Disruption: On Data, Democracy, Society and Statistics*, DATA-POP ALLIANCE WORKING PAPER (Dec. 2015) http://datapopalliance.org/wp-content/uploads/2016/03/DataPopAlliance_LeveragingAlgorithms.pdf (last visited June 8, 2020).

[16] Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242 (2019).

[17] See further, the special issue titled *Justice in Algorithmic Robes*, *International Review of Law, Computers & Technology* Volume 31 (2), 2017.

[18] FOTIOS FITSILIS, IMPOSING REGULATION ON ADVANCED ALGORITHMS (Springer 2019). The cases – Microsoft cases (OS separated from media player and browser), Volkswagen emissions case (illegal algorithmic switch to sense and adjust gas emissions signal in the vehicle), Axel Springer AG v. Eyeo (ad-blocking, German cases), Google Spain SL and Google Inc. v. AEPD and Gonzales (right to be forgotten) and sharing economy in AirBNB and Uber, algorithmic financial instruments in GDPR.

[19] For more on this see, Maayan Perel and Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, 69(1) FLORIDA L. REV. 181 (2017).

there will be reconsideration of the morals when people become used to the new technology and its possibilities and limitations. In time, morality will adapt'.[20] The breadth of concerns that are articulated brings our focus to the challenges (and opportunities) in the realization of disparate rights in isolated areas of law, as also towards the continuing recognition and strengthening of the very values that constitute contemporary societies and polities. Nemitz focuses on the challenges for law in responding to the concentration of digital power with the development of AI, and argues for incorporating principles of democracy, rule of law and human rights in the very design of the algorithm, and calls for Technology Impact Assessments to govern these technologies.[21] The Scientific Foresight Unit of the European Parliament identified a number of opportunities and challenges in the use of algorithmic decision-making. These included 'risks related to health, quality of life and physical integrity' and serious threats that undermine 'the fundamental principles of equality, privacy, dignity, autonomy and free will'.[22]

The question of agency – the presupposition that it is possible for human and social actors to influence the development of new technology– itself is a central concern about algorithmic and AI technologies. Within STS debates, the question of agency 'leads into a long-standing discussion about technological determinism, and its more recent counterpoint, social determination of technology. In the technological determinist view, emerging technologies will materialize anyhow, independent of what people think deliberate or decide.'[23] Given that we are talking about technologies that are termed as Artificial Intelligence, or autonomous technologies, these longstanding concerns generally about technology and human agency become sharper and even more prominent.

It is still early days for algorithmic trajectories to take roots in society, where the deep difference in public values about the introduction of these trajectories can and should impact the development and deployment of these technologies. What legal and regulatory principles can help in the articulation and institutionalization of these variegated concerns?[24] What visions,

[20] See Tsjalling Swierstra and Arie Rip, *Nano-ethics as NEST-ethics: Patterns of Moral Argumentation about New and Emerging Science and Technology*, 1 NANOETHICS 3-20 (2007) [hereinafter Swierstra].

[21] Nemitz, *supra* note 1.

[22] STOA, *supra* note 7; Similarly, Justice Sales argued for the constitution of a public agency for upstream scrutiny of algorithms, coupled with subsequent legal challenge in courts, Sales*, supra* note 5.

[23] Swierstra, *supra* note 20, at 8.

[24] Even as the relationship between law and regulation can be viewed as unclear, law can set the normative cornerstones for regulators, and legislations are a species of regulation. See ROGER BROWNSWORD, RIGHTS, REGULATION AND THE TECHNOLOGICAL REVOLUTION 6 (Oxford University Press 2008). See the statement of the UNESCO's Director General: '*a firm commitment to international solidarity in scientific progress, and to safeguarding human*

foresight and horizons are important here,[25] and how do we imagine the role of the public sphere to envision, deliberate, and assess these technologies. Given that these technologies also (re)produce 'social order' and the controversies pertaining to surveillance, bias, agency and transparency recur across multiple sectors 'ranging from the public sector to labour management and ordering digital communication', demonstrating the need for a broader conversation across and beyond specific sectors. This points to the contingent and contested nature of the issues at hand, with deep differences traversing sectors and contexts that are "shaped by interests, power, and resistance".[26] These debates ought to, then, regard foundational choices and constitutive values relating to the 'if and how' of developing and deploying these technologies. This special issue is seen as a contribution to the constitution and performance of such a public sphere on law, technology and society,[27] which takes the prospect of the 'algorithmic turn' seriously. It is in this context that the interventions in the current issue become significant.

The various contributions in this special issue offer valuable accounts and perspectives regarding the introduction of algorithms in multiple realms, be it in the so-called gig economy, health care, social media and hate speech, and its implications. The contributions focus on various concerns about risks to human rights, including dramatic increases in surveillance and erosions of privacy, loss of fair working conditions, and exacerbation of hate speech, arising from the introduction of algorithms in these realms, and the effects they have on democracy, citizenship and selfhood. Solange Ghernaouti dwells on the digital technologies that increase the potency of existing infrastructures of state surveillance, and how the digital ecosystem itself should then be viewed as a new risk for human rights. She presents a critical analysis of the detrimental effects of digital practices, digital ecosystem and the Internet on civil liberties, other related human rights, as also protection of the environment and biodiversity. She offers prescriptive and strategic recommendations based on the idea of the 'common good', and taking the concerns of cyber-security and multilateralism as part of the possible solutions. Siddharth Narrain focuses on the effects of online practices and platforms on an important fundamental right, the human right to

rights and human dignity from the misuse of science and technology', BIOTECH: INTERNATIONAL IMPLICATIONS 2 (UNESCO, Paris 2003); c.f. Brownsword above at 23.

[25] Visions and Visioning becomes regulatory and public tools to construct and examine coherent pictures of potential future socio-technological states that arise from the development and deployment of the identified technological trajectory, and can then become useful tools to anticipate, deliberate and change course if necessary. See further, the special issue of Economic and Political Weekly, Sekhsaria and Thayyil ed., *Narratives of Technology and Society Visioning in India*, 54(4) ECONOMIC & POLITICAL WEEKLY 39 (2019) – on the relationship between Visions, Visioning and other techniques of technology assessment with the development of technology.

[26] Christian Katzenbach, *Algorithmic Governance*, 8(4) INTERNET POLICY REV. 1-18 (2019).

[27] See further, Naveen Thayyil, *Claiming the Social: Beyond 'Law as Technology'*, 11(2) SOCIO-LEGAL REVIEW 1, 19 (2015).

freedom of speech and expression. He elaborates on the distinctive nature of hate speech online, its performative aspects, and the need for regulators to start focusing on the particularities of various platform infrastructures that makes the proliferation of such hate speech possible. Beyond the usual rhetoric of identifying what is hate speech in the quest for regulating online hate speech, he reminds analysts of the importance of taking the algorithmic materiality and infrastructure of hate speech in India seriously.

Two articles focus on the detrimental effects on rights via a prominent site of digital reconfiguration of work in the algorithmic 'gig economy' - ridesharing and cab aggregators. Mishra and Thayyil discuss the implications for labour rights from algorithmic infrastructures in rideshare apps in India. Within the context of a longer history of normalization of informal (and disorganized) work in the country, they identify specific facets in attendant labour practices, viz., precarious living conditions from the temporary and unprotected nature of the work environment, sharp drop of income for the transport workers accompanied with (even) longer working hours, as also erosion of unionization and collective bargaining possibilities. The authors locate the role of algorithms in these disentitlements, and call for specific regulatory measures towards algorithmic accountability. Marston and Holtum discuss the future of work in the time of algorithms, through a focus on the advent of ride-sourcing platforms in Australia. The deterioration of working conditions, and the erosion of the security and rewards of paid work through the steady growth of digital labour platforms of gig-work is seen as emblematic of the de-standardization of work in Australia. The authors focus on the future of work during AI and algorithmic automation as an opportunity to think about structuring work at large and to map out a more equal society. Towards this, the authors argue for a more considered, broader regulatory and social approach that addresses the cause of social inequality in society.

Two further contributions examine the complexities in the balancing acts that regulators encounter in specific sectors, viz., algorithms in the health sector, and in state surveillance. Nishtha Bharti focuses on the regulatory conundrums related to the recent impetus for algorithmic initiatives in healthcare in India. She identifies key conundrums through a cursory examination of this landscape and underline the challenges that beset the 'panacean' promise that drive the narratives of development and deployment of algorithmic technologies in healthcare. Through such a focus, she emphasizes the need for contextualizing this technology in the 'fragile but unique model of Indian healthcare, and to move forward keeping in mind the fallibility of technological design' of the underlying architecture, 'however pioneering it might appear'. Snehil

Singh focuses on the need for additional judicial safeguards against panoptic state control and surveillance, through Section 69 of the Information Technology Act, 2000 in India. Towards effective protection for the fundamental right to privacy, he argues for the institution of prior judicial scrutiny ('*ex-ante ex-parte*') so as to ensure that the specific executive actions of surveillance are well within recognized constitutional limits and to 'prevent unjust surveillance and protecting the right to privacy'. Whether the judiciary will have the infrastructural capacity and intellectual resources to engage in such 'technical matters' in what may well be a whole slew of such executive requests would however require further consideration.

The final two contributions offer a glimpse of the fundamental nature of the societal debates around the deployment of algorithmic infrastructures, in law, technology, and society spheres. Anubhav Banerjee seeks to move beyond an anthropocentric frame of citizenship by seeking to analyze the state capacity and legal consequences in conferring citizenship rights to artificially intelligent entities. He examines the limits of existing jurisprudence in accommodating the determination of AI entities as citizens, the feasibility of conferring such rights to AI entities, and its legal consequences. In contrast to these sanguine expectations about such technological trajectories, Upendra Baxi locates the new practices of self-making in artificial general intelligence as an integral feature of neoliberal governance, and the resultant commodification of subjective experience as an active usurping of freedom. In his brief postscript, he engages with the implicit notion of algorithmic selfhood, and offers a glimpse of the breadth of thinking that is necessary to approach the fundamental issues that are involved in this algorithmic turn. In taking into account the blurring lines of 'artificial' and natural legal personalities, he underlines the importance of engaging with the problem of 'how fictionality of legal person could ever transcend the will of those who created them in the first place'. Significantly, he argues for a vision of law as a social technology that seeks to impose some discipline of rights and responsibilities on the making of AI and such hard technologies.

Deliberating the desirability, possibility and the legitimate contours of legal regulation of new technologies, and identifying the relevant normative principles would be indeed key parts of the task at the supposed wake of this algorithmic turn. Many would agree that existing human rights frameworks could be seen as a baseline to start this conversation to seek accountability and assert responsibility from the making of the algorithms. Is it possible for law to aim for a broad agreement on the use of Algorithms and AI in general, or at least in specific sectors, and if not, what are the normative principles useful for law to recognize disagreement, characterize the contours of social disagreement, and act despite such disagreement? Further, what regulatory principles and legal doctrines are to be invented and employed to facilitate law to democratically act despite deep differences and disagreements, even as we find ways to identify knowledges and expertise that are (un)available for lawyers and regulators to govern the development and deployment of these technological trajectories? It is apparent to anyone but the willfully blind that these algorithmic technologies are deployed within existing social cleavages, sometimes exacerbating them, and sometime producing new inequalities. Attention on the regulatory aims and goals that the law recognizes as legitimate, and the regulatory tools that the law finds as efficient, effective and legitimate in responding to these technological trajectories, can make visible the disagreements and resistance about law itself. What is currently considered self-evident about law, then, is suddenly recognized as contested, and the inevitability of such contestations in democratic law would then also make a supposed radical technology a handmaiden for democratizing law. Through the fostering of new public arenas of discussion and deliberation on putative algorithmic selves, then, helps assert the inevitability of contestation and reassertion of what law ought to be, also as an algorithm for law and democracy.

HOW DIGITAL ECOSYSTEM AND PRACTICES INCREASE THE SURVEILLANCE SYSTEM'S PERFORMANCE AND GENERATE NEW RISKS FOR HUMAN RIGHTS

- Solange Ghernaouti[*]

ABSTRACT

This article presents a critical analysis of how the internet, digital practices and the digital ecosystem affect human rights and civil liberties. It highlights the effectiveness of surveillance and control mechanisms and identifies aggravating factors related to the applications of big data and artificial intelligence. It proposes some strategic recommendations and possible solutions based on a multilateral approach to regulating a common good and on cybersecurity measures. Some perspectives related to the urgent need to act by considering the impacts of digital technology on the environment and biodiversity are given.

Key words: *Cyberspace, Digital dependencies, Artificial intelligence, Cyber risks*

I. THE DIGITAL MONITORING CONTEXT

Digital technologies now contribute to how we move through space and time, to how individuals are rated (rating of service providers, clients, doctors, patients, counting the numbers of friends or followers on social networks, etc.) and to how those individuals are monitored and controlled. Digital technologies facilitate mass monitoring because every activity, every connection, every use leaves a trace.

Users' digital fingerprints can be collected, stored, and cross-referenced with those from other sources (telephone, camera, web browsing, online payment, credit card, physical surveillance, etc.). Data provided by users are associated with data collected without their knowledge as well as data generated by computer processing operations applied to them. This makes it possible to create new data that helps build up profiles of people, particularly with regard to their behaviour, movements, tastes, feelings, commercial transactions, media consumption, etc. Classifying people, groups of people, associating a score, a rank, adding advantages (carrots – discounts) as well as restrictions (sticks – denying access to services) not only makes it possible to measure individuals but also to train them. Indeed, the customer experience that certain digital services or

[*] Professor at the University of Lausanne; Director, Swiss Cybersecurity Advisory and Research Group.

platforms provide and users' dependency on them contribute to making users "docile and useful", to subjugating and disciplining them.

A. THE ILLUSION OF FREEDOM AND SECURITY

In Europe, we are getting used to consuming digital services, to exposing ourselves on social networks without questioning too much the long-term impact that this can have on the preservation of our fundamental rights. We subscribe to the illusion of safety that can be provided by a surveillance system presented as a protection system. The illusion that security justifies surveillance is the product of the effects of language games deployed by the state and the media insisting on the urgent need to fight terrorism and crime with exceptional measures and sophisticated technology. Public policies bring after each dramatic event a reason for more severe surveillance and the population consents to this surveillance thinking that after such a drama the need for security justifies the violation of its privacy. This is notably what happened after September 11, 2001, with the adoption of the Patriot Act in the United States and in France after the terrorist attacks of 2015, a new law on internal intelligence was passed allowing wider surveillance of communications. It should be noted that these more invasive measures did not prevent criminal events from occurring, in 2015 there was one mass shooting every day in the USA,[1] in France, the Nice massacre in the summer of 2016 could neither be predicted nor prevented despite the existing surveillance. Insecurity therefore remains, the feeling of security is fading away while the surveillance continues. Some believe that the transparency of beings and their activities is synonymous with security, when in reality it contributes to their alienation (in a posture imposed by the surveillance system) and to their submission to the system. This is demonstrated by George Orwell in his novel "1984"[2] in which we are led to believe that human dignity and intimacy and privacy have gradually been destroyed, that social ties have broken down, and that individuals have isolated themselves from each other, because of the invisible control operated by the "telescreen", a true panoptic surveillance system.

In a similar way to the citizens of many cities and countries becoming accustomed to moving under the watchful eye of video surveillance cameras, we integrate, without question, the surveillance that our digital practices allow – practices imposed by the massive digital transformation of society where all traditional activities are now mediated by digital technologies

[1] Samuel Paulet, The Terrible Numbers of gun deaths in the United States, https://www.focusur.fr/societe/2017/10/03/les-statistiques-terribles-sur-le-port-darmes-aux-etats-unis/. (last visited June 8, 2020).

[2] GEORGE ORWELL, 1984 (Secker and Warburg 1949).

and information systems. In contradiction to the reality of surveillance and addiction, digital use has been sold as liberating to individuals. This was, moreover, a desire of the creators of the Internet during the hippie era. This freedom is an illusion. The system is designed in such a way that the individual believes that he or she is anonymous, that he or she is alone, that he or she is hidden, that there seems to be no legal framework, that he or she can access any content without being seen, that he or she can do anything without being judged. This feeling is in fact false, since all our behaviour is in the hands of companies, sometimes even influenced by marketing but also taken in the hands of States. The individual is in fact forced to live with digital services, there is no choice to do without, except to exclude himself from society.

The majority of people can no longer do without information and communication technologies, the adoption of which has been driven largely by constant entertainment and consumer practices. This self-alienation and self-subjugation are reinforced by the emotional bond developed by Internet users towards the services they consume and their connected objects. This contributes to preventing them from developing a critical view of the reality of the surveillance to which they are subjected and thus from being able to challenge it.

B. A WORLD OF MISTRUST BY DESIGN

The Internet makes it possible to communicate with potentially everyone, and therefore anyone. It is difficult, if not impossible, to verify who is hiding behind a screen or behind a virtual identity, fake identity or pseudonym. There is no "security" mechanism to guarantee the good faith of Internet users, let alone that of the software robots that feed communication platforms. It is up to Internet users to remain vigilant and to decide for themselves whether what they see, what they hear, is true or false.

How then is it possible to ascertain whether the content being accessed is advertising or false information? Is the information being disseminated an attempt to defraud, influence or manipulate? How much trust should we place in what we see or hear on the Internet? How can we decide if it is possible to interact with an entity trustingly and safely?

Moreover, a false sense of "trust and closeness", that cannot be reliably and lastingly justified or guaranteed, is cultivated by providers of social networks. The feeling of being "among friends" generates a false sense of security that can have disastrous consequences on people's lives in the cyber and physical world. Some of the main risks to which people are exposed include, for example, the following:[3]

- Harassment, intimidation, blackmail, anti-social behaviour, bullying, etc.;
- Defamation, damage to reputation;
- Exposure to malicious, offensive or unwanted content (viruses, spam, malware, ransomware, pornography, violence, incitement to racial hatred and xenophobia, propaganda, etc.), intrusive advertising, hoaxes, scams, fraud or abuse of any kind;
- Object of surveillance, traceability, excessive profiling, environmental monitoring, espionage;
- Data theft: personal data, confidential information, intellectual property;
- Theft of equipment (computers, telephones, tablets, USB keys, CD-ROMs, etc.);
- Identity theft;
- Disinformation, manipulation of opinion, influence, etc.

In addition, at the individual level, some people develop digital consumption behaviours and habits (social media, entertainment, gambling, sex, etc.) that could qualify as addictions. This raises the following questions in particular:

- How can we avoid the addictive consumption of digital technology?
- How can we avoid addictive behaviour that could harm people's physical and moral integrity?
- How can we protect people against their will?

From a societal point of view, in this age of fake news and information manipulation, how can we be a society without a common base of information and perception of reality?

C. THE HIJACK OF PRIVACY

A threshold was already crossed more than fifteen years ago when privacy and personal data were hijacked by actors in the digital economy. Indeed, the profitability of the digital economy has developed on the basis that personal data are considered to be marketable values provided free of charge by users. In addition, people's privacy in their homes is also affected by the intrusion of voice assistance systems such as those offered by Google, Amazon or Apple for example. Connected televisions or toys are also able to spy on their owners, as are connected watches or jewellery. The latter, as close as possible to the body of individuals, are able to record their slightest movement or biological functioning. Even once-private places like cars do not escape

[3] S. GHERNAOUTI, LA CYBERCRIMINALITÉ, LES NOUVELLES ARMES DE POUVOIR (Collection Le savoir suisse, PPUR 2017).

computerization, automation, decision support, which results in automated monitoring. Customers of such systems have most often more or less given their consent in exchange for less effort, personalized service, or facilitated interaction, to name but a few justifications.

Therefore, under these conditions, how can Article 12 of the Universal Declaration of Human Rights[4] – "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks on his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks" – be respected?

II. THE DIGITAL TRANSFORMATION OF SURVEILLANCE MEANS

'Discipline and punish' is the title of Michel Foucault's book, the subtitle of which is 'the birth of the prison', published by Gallimard in 1975 in France.[5] It is impossible not to think that increasingly intrusive video surveillance systems since they are increasingly intelligent and omnipresent, are the new invisible fortresses of the 21st century that imprison individuals in order to 'retrain' them, keep them on an electronic leash, force them to opt for behaviours that are not only socially acceptable, but normalized according to a specific political system. Moving under the constant eye of the cameras means accepting the internalized constraint of predefined behaviour, living locked in 'a social orthopaedic enterprise' in Foucault's words. When people know they are being observed, their behaviours change. Europe is not immune to the adoption of a system of total surveillance, skilfully justified to a population ready to accept it, or imposed by market forces serving a particular security policy supported by specific public-private partnerships.

The digital transition of society also concerns the transformation of surveillance means. These are implicitly integrated, not under this name and not as such, in all the personalised services developed with big data and artificial intelligence. Indeed, massive data collection and algorithms make it possible to deduce people's physical, psychological, sociobehavioural, sexual and other characteristics, even to possibly predict them and thus also to influence them.

[4] UN General Assembly, *Universal Declaration of Human Rights*, 10 December 1948, 217 A (III), http://www.un.org/fr/universal-declaration-human-rights/index.html., http://www.un.org/en/udhrbook/pdf/udhr_booklet_en_web.pdf. (last visited June 8, 2020).

[5] MICHEL FOUCALT, SURVEILLER ET PUNIR: NAISSANCE DE LA PRISON (Gallimard Parutions 1975), http://www.gallimard.fr/Catalogue/GALLIMARD/Bibliotheque-des-Histoires/Surveiller-et-punir.

These developments undermine human rights. If they are to be feared in Europe, they already exist in China with the 'social' control model (Social Credit System) that generalizes surveillance through technology. In this system, rights are granted according to the behaviour recorded via smartphone, web, video surveillance cameras, facial recognition systems, or by payment or access controls (train tickets, parking terminals). Depending on how well the person behaves in front of the surveillance technology, he or she will have the right to access private property, to a job, to move, to meet some people, to access some products. This is achieved by the automatic processing of the collected data and the automation of the decision by the machine. In the long term, such transformations could become a dictatorship of human rights by the algorithm which is the only master of judging behaviours and rights. For example, a robot geolocalized that you crossed at a red light, so you will not be able to buy a car because for the system you are a bad user of the public road.

A. HUMAN RIGHTS AT RISK

Digital technology favours a dual, physical and digital, surveillance of individuals, with the possibility of cross-referencing the data from each. This increases the surveillance system's performance, its automation, its deterrence and self-censorship effects as well as social pressure, while considerably restricting individual freedom and reducing the likelihood of risking not being in compliance with an average that has been artificially decided by algorithms that are more or less well designed and imposed by governments.

The protection of personal data is a prerequisite for self-determination and for protecting freedom of expression and human dignity.[6] At present, it is very difficult, if not impossible, to protect personal data, mostly considered as a commercial asset in an open competitive global market. We cannot ignore anymore that data has a hidden life beyond its owner's control and e-services have been designed to exploit them in the spirit of data and surveillance's capitalism digital economy.[7]

Deprived of our data, human is naked, transparent, invisible. The more people allow themselves to be dispossessed of their data, the more individuals can be considered as an information system to be improved. There are becoming data extraction entities (objects), under surveillance and

[6] S. GHERNAOUTI, CYBERPOWER, CRIME CONFLICT AND SECURITY IN CYBERSPACE (EPFL-CRC Press 2013).

[7] S. ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (Public Affairs 2019).

remotely controlled. In doing so, we are entering the era of programmed human obsolescence and in the same time we are adapting to this new technical-economic reality in order to exist.

Those who were born in the era of digital submission and who know nothing else, develop normalized behaviours through big data applications and algorithms. They are a kind of hybrids, perfected by software updates and nanotechnology grafts, forced to be measured and optimized, to perform even in the acts of intimate life. Born under the yoke of the digital age, they will be content to live as they were born. Subjected to the propaganda of a simplified, tailor-made and instrumental vision of the world. No need to vote, no need for democracy, it will be enough to adhere to a prefabricated way of thinking, shared by the mass of connected individuals. Manipulation, opinion management and fake news will be the new ways of expressing opinions as an alternative to public debates.

However, these transformations are not without impact, since standardization and submission to algorithms lead to the undermining of human rights. The human being subjected to targeted advertising, fake news, order reception by robots no longer has the freedom to act autonomously, to think freely. Some artificial intelligence system violates freedom of speech and freedom of the press by deciding what content can or cannot be communicated on the Internet. The fundamental right to privacy is also threatened in this environment where the user has no choice than to offer the content of his digital activity. It is possible to go even further in denouncing human rights through digital technology. The freedoms of worship and conscience will soon be non-existent since our beliefs could be manipulated by algorithms that decide what information people have to access. The right, to equality and the right not to be discriminated are also violated by digital technology since, without access, some individuals cannot benefit from knowledge, services or even employment. Finally, because of the impossibility of denouncing certain prejudices or other events linked to digital technology, the right to a fair trial, to a fair trial can also be called into question.

B. SPECIAL CASE OF BACKDOORS AND ENCRYPTION SOLUTIONS

In addition, backdoors imposed by manufacturers/suppliers, foreign or not, may be installed, sometimes at the request of a State, without the knowledge of customers and users, through the design of computer and telecom hardware and software. This increases the potential for abuse of surveillance and espionage.

The major problem is to be able to find the right balance between the need for trust and confidence in digital technologies and services on the one hand and the need for ease of intercepting data and controlling systems and users on the other. The need for surveillance may be justified on grounds of national security, the fight against crime, public safety and the defence of a State's interests. Reasons of competitiveness and economic warfare may motivate the interception, surveillance or espionage of data by private or public entities.

Verifying, auditing, testing, certifying products and services, offering a certificate for digital services and products 'without manufacturer backdoors' could be part of a solution. But in fact, it shifts the problem of trust to the entity that provides the certification. The problem of trust is therefore not solved, it requires a certain independence and autonomy of the certification authority and raises the problem of its recognition at national and international level. One could consider developing encryption solutions to be used on untrusted hardware and software. Generally, police forces do not appreciate not having easy access to data and that users have encryption techniques outside the police's control. A solution must then be found to implement reliable and trustworthy encryption techniques that allow the encryption of users' sensitive data and also allow judicial and police work when necessary.

Such solutions exist, and involve establishing a trusted authority (a kind of national electronic safe) with which users of encryption mechanisms would deposit their encryption keys, which could be issued, in the event of a judicial investigation, to authorized agents, upon a justified request validated by the competent authorities.

C. ARTIFICIAL INTELLIGENCE AS AN AUGMENTED FACTOR OF RISK

Data collected for surveillance purposes and entrusted to artificial intelligence do not pose any more of an 'ethical' problem than if surveillance is left in the hands of private or public agents. The word "ethical" is frequently misused or possibly refers to a certain 'code of conduct' for surveillance practices as defined and justified by an appropriate policy. Errors of artificial intelligence are large-scale automated human errors. Like computer programs, agents or mercenaries are corruptible. An artificial intelligence program that substitutes for a human in carrying out surveillance work raises the more general question of the right to work, the transformation of the surveillance professions, and even their disappearance.

The quality of artificial intelligence programs that predict potentially deviant behaviours before they are carried out, that identify 'dangerous' people, depends on the quality of the data collected, the know-how, the skills of the developers, the budget allocated to develop quality systems, and the control and validation systems for monitoring programs. With artificial intelligence, someone can be penalized even before a crime is committed on the basis of a simple presumption of guilt, which is determined according to non-transparent criteria and algorithms that can have bias.

D. THE BALANCE OF POWER AND MARKET FORCES

According to Solow's model, any economy reaches a point where any increase in production factors will not lead to an increase in per capita output. Output then becomes stationary. However, this point is never reached due to technical progress. The use of technology allows for constant progress, it is this technical progress that leads to ever higher production. Indeed, digital technology is seen today as one of the most powerful engines of growth not only for a company but also for the States. All sectors are concerned by digital technology. Similarly, all ways of consuming and producing have been transformed by digitalisation. This economy is dominated by the exploitation of data on a large scale, by network effects since the quality of service depends on the number of users, and by the automation of activities to ration production costs. As a result of increasing returns, the digital economy is naturally prone to market concentration (called 'natural' monopolies), even though innovation can at any time challenge a dominant position. In any case, today a company cannot survive without addressing the digital issue.

The power relations between all members of society, on a global scale, now require the mastery of information and data technologies. "Seeing and influencing without being seen" is a political, economic and techno-ideological project that stems from a desire for unlimited power. It is to be feared that some in Europe and elsewhere in the world will defend their interests by promoting computerised surveillance. It only takes one opportune moment (terrorist attacks, economic problems, immigration, etc.), convincing political speeches, a few repressive laws, and effective communication for the population to submit to such a system, or even request it.

The market exists, it is supported by manufacturers and solution providers who only want to sell. The digital economy, which can be described as an economy based on data, attention or interaction, is expanding rapidly, as is the security economy.

Moreover, the population is getting used to it without resistance, as it has become accustomed to the a priori painless use of credit cards, online payments or tax return, but also to personalized advertising. However, the traceability of electronic payments allows for monitoring and social control. In addition, there are many incentives to stop paying in cash. Personalised advertising is a form of surveillance that is initially commercial in nature but can evolve towards other purposes.

III. The Difficult Issue Of Personal Data Protection

The protection of personal data is a prerequisite for self-determination and for protecting freedom of expression and human dignity. At present, it is very difficult, if not impossible, to protect personal data and all technologically generated metadata because they have a hidden life beyond their owner's control and because services are not designed to take into account individuals' need for digital privacy.[8] The table in Figure 1 illustrates the different types of data derived from a consumer of digital services.[9] In addition, all the principles underlying data protection laws are based on the explicit consent of the individual, for a known and precise collection purpose and for a fixed period of time. With the phenomenon of big data, where data are collected indiscriminately for further and a priori unknown processing, it is difficult to control, with current regulations, these new practices for extracting information, whose profitability is based on the unlimited exploitation of this raw material.

Data granted by the user	Contents of messages, photos, completed forms.
Data from the observation (monitoring) of user practices	Behavioural data related to the use of resources (websites visited, pages visited, time, geolocation, access terminal, products studied, videos viewed, research carried out, online purchases, articles read, etc.)
Combined data (unified user profile generated by cross-referencing several sources)	Data gathered from several online sources (messages, photos, blogs, friends, contacts, etc.) and real-world sources (video surveillance

[8] Metadata is data concerning other data (e.g. Data: Name / Address of the website consulted; Metadata: identification of the time of consultation, geolocation, client terminal).

[9] S. Ghernaouti, Guide Pratique De La Cybersécurité Et De La Cyberdéfense. (Organisation internationale de la francophonie 2017), https://www.francophonie.org/IMG/pdf/cybersecurite-web.pdf (last visited June 8, 2020).

	cameras, credit card usage, health, insurance etc.)
Data deduced from inferences, algorithms, computer processing	Deduction of new user data via computer processing (mathematical and statistical models)

Figure 1: The various sources of big data

A. A NEW DIMENSION OF CYBER RISKS GENERATED BY THE INTERNET OF EVERYTHING

Connected objects constitutive of the Internet of Things (IoT) are de facto potential targets of cyberattacks. Anything connected to the Internet can be hacked. We should never forget that it is not just an object that we connect to the Internet, it is our lives that we connect, via mobile phones, digital assistants or self-quantification devices.[10]

Self-driving cars or connected refrigerators have already been taken control of remotely. All embedded objects and sensors, RFID (Radio Frequency IDentification) chips have the ability to collect data and transmit it to service providers, which can compromise the protection of personal data and the privacy of people connected to such objects. In addition, their lack of security or robustness, the possibility that they may be manipulated by malicious or unfair entities, can have detrimental consequences for the physical security of infrastructures and the people who depend on them. In the field of personal assistance with everyday activities, more or less sophisticated robots are beginning to share the daily life of some people. Able to influence our behaviour and our environment, their takeover by deviant or malicious entities can have a direct or indirect impact on human rights and civil liberties.

The world of electronics, whose components may be on the Nano scale,[11] and the biological world are increasingly converging, as various sensors, prostheses and biomedical electronic equipment is implanted in the human body to compensate for some of its failures (insulin pump, pacemaker, etc.). If the trend is confirmed to integrate more and more electronic equipment into biological bodies and thus to consider an evolution of the human being towards a hybrid being, a

[10] Self-quantified objects such as connected watches and activity bracelets, for example, make it possible to measure all kinds of parameters related to the functioning of the human body (heart rate, blood pressure, sleep, weight, etc.), lifestyle (number of steps taken, etc.), nutrition (calories consumed), etc. This makes it possible to develop the market for connected well-being and health, but also the market for monitoring employees or insured persons, for example.

[11] Nanotechnologies refer to devices made at the manometer scale (1nm = 10-9 m - notion of intelligent dust).

person with increased capacities through digital technologies, as the trend of transhumanism[12] prefigures, then security requirements become essential and must be taken into account right from the design-phase of technologies on which humanity's long-term survival would depend. Otherwise, the notion of augmented humans could translate into fallible humans at the mercy of those who have the power to control them remotely and to destroy them.

Neural interfaces already exist that allow interaction with computers through thought. If this makes it possible to contribute to the well-being of certain people, particularly people with disabilities, in the long term the integration of biology and electronics could lead to attacks, including the hacking of human body and thought.

B. ARTICLES OF THE UNIVERSAL DECLARATION OF HUMAN RIGHTS THAT COULD BE IMPACTED BY DIGITAL SERVICES

The Internet, digital transformation, electronic surveillance, traceability of activities, data collection, the exploitation of that data, business models based on personal data or cloud computing paradigm are not necessarily compatible with respect for fundamental rights and civil liberties. The consumer's stored, processed and exploited data are generally stored in a foreign country where the local law allows it but where the law of the State of which the Internet user is a national does not necessarily apply.

A level of cybersecurity adapted to the needs of individuals and public and private organizations is difficult to ensure. This can contribute, in particular, to a threat to economic competitiveness, freedom of expression, association, mobility (freedom of movement, Internet browsing), the right to access information or even knowledge, the right to secrecy of correspondence or the right to privacy, as recognized in the 1948 Universal Declaration of Human Rights,[13] as shown in the table in Figure 2.

Although the culture of human rights and civil liberties dear to democratic countries can be perceived and implemented very differently throughout the world, it is nevertheless necessary to

[12] For a history of transhumanism, see for example COURANT PHILOSPHIQUE, http://www.histophilo.com/transhumanisme.php (last visited April 25, 2020).

[13] UN General Assembly, *Universal Declaration of Human Rights*, 10 December 1948, 217 A (III), http://www.un.org/fr/universal-declaration-human-rights/index.html (last visited June 8, 2020).

contribute to their prosperity and respect through cyberspace and to share this common value, which can be considered as a lever harmonious coexistence everywhere in the world.

The asymmetric surveillance procedures that potentially restrict our freedoms inherent in the design of cyberspace and of digital technologies, and the many cases of abuse of power and invasion of privacy by state and non-state Internet actors (private companies, criminals, terrorists) suggest that the slogan 'Big Brother is watching you', where Big Brother is an invisible dictator exercising totalitarian control through an asymmetric surveillance system, is in fact, the Internet.

Article 1: « *All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one other in a spirit of brotherhood.* » Have the digital infrastructures and services on which human activities have become dependent been designed in a spirit of fraternity?
Article 3: « *Everyone has the right to life, liberty and security of person* ». Endangerment of freedom and security through the misuse, misappropriation or criminal use of information technologies and the surveillance possibilities inherent in digital technologies. The consequences of certain cyber-attacks can affect human life. The freedom of individuals can be undermined by some of the surveillance, censorship, control and manipulation practices promoted by Internet use.
Article 4: « *No one shall be held in slavery or servitude; slavery and the slave trade shall be prohibited in all their forms.* » Can we guarantee that individuals who are permanently connected or who provide unpaid data to service providers are not held in slavery and digital servitude?
Articles 6: « *Everyone has the right to recognition everywhere as a person before the law.* »

Article 8: « *Everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him as recognized by the constitution or by law.* » The transnational dimension of the Internet is often an obstacle to the application of the law and does not promote an effective remedy before national courts.
Article 12: « *No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.* » The protection of personal data and digital privacy are not guaranteed. Attacks on honour and reputation can be facilitated by the Internet. The application of national laws or European regulations, such as the GDPR for example, in an international space, raises many problems with prosecuting offending actors, reparations for victims, mutual legal assistance and international cooperation. The application of the law, even if laws exist, requires considerable resources, which is often an obstacle to the implementation of justice.
Article 19: « *Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.* » Freedom of opinion and expression and the right not to be disturbed for one's opinions can be considerably undermined by the use of information technologies. Censorship, filtering, blocking, profiling, surveillance, espionage, intimidation, etc. are realities of cyberspace.
Article 20: *« 1. Everyone has the right to freedom of peaceful assembly and association.* *2. No one can be compelled to belong to an association. »*

Freedom of assembly and association in cyberspace can be significantly restricted due to digital technologies and the possibilities of traceability, surveillance and censorship. By extension, ICT service providers should not have the possibility to impose general conditions obliging the user to accept a set of clauses, services that could be comparable to the obligation to be part of an "association".
Article 21: *« 3. The will of the people shall be the basis of the authority of government; this will shall be expressed in periodic and genuine elections which shall be by universal and equal suffrage and shall be held by secret vote or by equivalent free voting procedures. »* Disinformation, the dissemination of false and fake news, manipulated or altered information, the production of "tailor-made" and personalised information, and their widespread dissemination (in number and intensity) all disrupt the functioning of fair elections. In addition, any electronic voting or remote voting system may be hacked.
Articles 23: *« 1. Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment.* *2. Everyone, without any discrimination, has the right to equal pay for equal work.* *3. Everyone who works has the right to just and favourable remuneration ensuring for himself and his family an existence worthy of human dignity and supplemented, if necessary, by other means of social protection.* *4. Everyone has the right to form and join trade unions for the protection of his interests. »* **Article 24:** « *Everyone has the right to rest and leisure, including a reasonable limitation of working hours and periodic holidays with pay. »* The notion of the right to work is challenged by the substitution of certain tasks and functions initially performed by individuals by computer systems. The transformation of jobs through information technologies and artifical intelligence is faster than the creation of new jobs and the transformation of skills that this requires. Fair and satisfactory working conditions and protection against unemployment cannot be guaranteed through cyberspace, this is also true for points 2, 3, 4 of Article 23 and for Article 24.

Article 28: « *Everyone is entitled to a social and international order in which the rights and freedoms set forth in this Declaration can be fully realized.* » The weakness of international regulatory mechanisms and the lack of universal Internet governance limit the expression of this fundamental right.

Figure 2: The main articles of the 1948 Universal Declaration of Human Rights that may be affected by the use of digital technologies

C. LIMITS TO UNLIMITED POTENTIAL

These new risks force us to reinvent security in order to best control them and preserve the values to which we attach importance and which are endangered by the technological development of society.

Thus, in this new context of all this computerized information and the many dependencies and interdependencies of living beings and activities that our survival depends on – such as supply chains, agriculture, the chemical industry, the health industry, transport, energy, or even access to crucial resources such as drinking water – the challenges of the fight against cyber threats go far beyond the fight against economic crime and the control of financial risks. For all these reasons, it is crucial to know and be able to identify incidents and their perpetrators, in order to deter them, and to put in place the necessary measures to minimise digital risks of criminal origin, those linked to failure to take adequate account of security needs or to the misuse of technologies. Indeed, as long as there is a feeling of impunity, that there are no convincing limits, that the risk is low and the potential profit is high, there are opportunities for crime, but also for abusive uses coming from completely licit actors and entities. These abuses occur because the entity goes beyond the strict framework of its competence. This is what is going on continuously with the right to privacy: as long as companies are not unmasked, prosecuted and sanctioned, violations continue.

The European General Data Protection Regulation (GDPR)[14] was created in part to stop these abuses, but although awareness is growing, the data market continues to flourish unhindered, while privacy violations are not diminishing. For a State it is the same problem, there will abuse acts of surveillance, carried out without legal basis but as long as there is no denunciation, or effective counter-power, it will continue. In Europe, for instance there is currently a subject of debate on this point about the tracing contacts application of Covid-19. Many states want to impose on citizens an application that, thanks to collected personal data, could define which person has been in contact with which other person in order to contribute to stop the chain of transmission of a biological virus. Although this application may have an advantage in some cases, its development poses a problem in terms of monitoring and legal basis used. There are no grounds in State constitutions or laws that allow widespread surveillance to stop a pandemic. However, the states give priority to health security over human rights and liberties and impose their plans. There are many protests in Europe, but legal bases could be created to specifically authorize such surveillance. The framework of such a system could be hijacked: how long would it be in place? to whom would the data be communicated? could we be punished for travelling? do we have an obligation to use such tracking? So many issues that need to be resolved to protect the human rights. Google has already enforced the tracking of individuals around the world and provided statistics on travel during the Covid-19 pandemic[15]. Have we been informed? On what legal basis did Google use customers geo-location data? What does it then do with these reports? There is a risk of violations of the right to privacy (art. 11 UDHR), to the freedom of movement of individuals (art. 13 UDHR). Other violations could occur as discrimination and exclusion based on one's movements could be determined. This could lead to exclusion from the right to property (art. 18), the right to a fair trial (art. 10), the right to freedom of thought (18), expression (19), assembly (20), to take part in public affairs (21) or the right to work (23). Even if it seems remote, it is already happening in China and in some sectors of our democracies (e.g. access to insurance, bank credit, certain jobs).

IV. FOR A DIGITAL URBANIZATION THAT RESPECTS HUMAN RIGHTS

Under the guise of entertainment, decision-making assistance, or illusory promises of well-being, e-this and smart-that's condition their users, who are forced to agree to their terms and conditions, to submit to the power of algorithms and therefore to that of their manufacturers and

[14] Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 relating to the protection of natural persons with regard to the processing of personal data and on the free movement of such data, 2016 O.J. (L 119) 1.

[15] GOOGLE, https://www.google.com/covid19/mobility/ (last visited April 27, 2020).

masters. As Natacha Polony and the Orwell Committee point out in the book 'Welcome to the worst of worlds, the triumph of soft totalitarianism', the entertainment industry is 'much more effective than any system of coercion'.[16] The seductive and attractive power of electronic interfaces, combined with the strength of the ideological discourses served up by the hegemonic actors of the Net based on neoliberal market forces and digital capitalism, present technological progress as inevitable and unquestionable. The slogan TINA, 'There is no alternative', echoes the misconception that technological change is part of the theory of the evolution of species and natural selection put forward by Darwin in the 19th century.[17]

A. SOME PROSPECTS FOR RESOLUTION: CONSIDER CYBERSPACE AS A COMMON GOOD TO BE SHARED AND REGULATED BY A MULTILATERALISM MECHANISM

The cyberspace with which we interact could be considered a common global space whose borders were linked to the geographical location of the individuals who use it and of the physical support infrastructure. The notion of ownership and the existence of an applicable legal framework can also help to define cyberspace. However, its international coverage, its extensive use by an ever-increasing number of people throughout the world, particularly through mobile telephony applications, and the increasing dependence of societies on digital infrastructures, require us to think collectively and globally at the national, regional and international levels about its development, use, sharing and regulation. Indeed, it would be detrimental to the stability of countries if cyberspace were only an economic and military battleground, reflecting all kinds of economic and political conflicts or competitions.

Although cyberspace is the result of technological evolution in a particular geopolitical context, it cannot be considered a natural evolution and is not completely comparable to the Earth, the sea, air and space. However, like these natural elements, it must be shared and therefore regulated. Cyberspace requires coordination, cooperation and legal measures that are applicable at the local level and that are effective and compatible at the international level. Thus, a supranational control instrument within the framework of the United Nations should exist and be applicable.

This would help to specify acceptable practices and prosecute offences wherever they occur. In addition, a supranational framework should ensure that crimes against peace and security through

[16] NATACHA POLONY & THE ORWELL COMMITTEE, WELCOME TO THE WORST OF WORLDS, THE TRIUMPH OF SOFT TOTALITARIANISM IN EDITIONS PLON 23 (Editions J'ai lu 2018), https://www.lisez.com/auteur/natacha-polony/90348 (last visited June 8, 2020).

[17] CHARLES DARWIN, THE ORIGIN OF SPECIES (1859).

the Internet become punishable under international law, even if they cannot be prosecuted at a national level.

As the Internet has global coverage and cyber threats and crimes do not stop at the borders of a country or region, regional or bilateral agreements are not enough. Broader legal tools are therefore needed.

Thus, an international agreement would facilitate:

- A common understanding of all aspects of cybersecurity;
- Harmonisation and complementarity of strategic and operational measures;
- The development of a comprehensive strategy to prevent, deter and respond to cyber-attacks;
- Information sharing and early warning of incidents.

Supported by the international community, it should be possible to develop a Cyberspace Treaty[18] within the framework of the United Nations, to reflect strong political and economic will and a real commitment by all countries. The development of such a treaty is being discussed within the international community, and will take time. Only an open and honest international dialogue around a sharing of common interests will help to establish the rules of harmonious coexistence in cyberspace and the real world, to identify the limits of acceptable practices and those that are not acceptable in cyberspace and to denounce abusive or criminal uses and their perpetrators. Aware that a treaty is not a brake on crime or the expression of various forms of power and conflict, it is nevertheless necessary to raise awareness that it is no longer a question of imposing the law of the strongest actors in cyberspace. Like an international climate protocol, there could be guidelines for the development of cyberspace that we will leave to future generations and together mark the way forward, taking into account all relevant actors, for a certain ecology of cyberspace.

This process is a lengthy one, but also makes it possible to raise awareness of cyber risks among the population and economic and political leaders and to remind them of their responsibilities. We are all aware of the limits of international treaties, of the problems associated with their ratification and compliance, but they still have a passive utility and effectiveness in setting a frame of reference. Acting together, in full knowledge of the risks, to seize technological opportunities in order to build an inclusive information society, to build a cyberspace and a trusted and

[18] STEIN SCOLBERG & S.GHERNAOUTI-HÉLIE, A GLOBAL TREATY ON CYBERSECURITY AND CYBERCRIME: A CONTRIBUTION FOR PEACE, JUSTICE AND SECURITY IN CYBERSPACE (E-dit, Oslo 2009), http://www.cybercrimelaw.net/Cybercrimelaw.html (last visited June 8, 2020).

sustainable digital ecosystem, to guide the changes brought about by digital technology, should mobilize us around a common objective and shared societal values. Perhaps it could be summarized by the desire to live together in complete security and stability in an online and offline world while respecting human rights.

B. CONSIDER CYBERSECURITY AS PART OF THE SOLUTION BUT ALSO AS A RISK FACTOR

The purpose of cybersecurity measures is to minimize the negative impacts of risks generated by the use of digital technologies and services. Cybersecurity is about managing and controlling risks. It is divided into a strategic and an operational plan. It's about being able to:

- Protect at-risk assets and protect against potential threats through appropriate protection and prevention measures;
- Limit the intensity and extent of damage caused by risks;
- Respond to incidents through incident management, crisis management, business continuity and possibly prosecution measures.
- All IT security measures controlled by the owners or managers of IT and telecom infrastructures require monitoring of users' activities.

There must be a security culture that is respectful of human rights so that all actors develop coherent and responsible behaviour. This security culture must take into account the need to respect human rights and must be translated into legal, organizational and technical measures that allow it. In addition, mechanisms must be put in place to facilitate the transparency of security measures and their communication to users. The first mechanisms to be put in place for the protection of human rights concern first of all the transparency of the systems for the enlightened information of persons, without this information, the individual cannot be free, dignified and exercise his reason and conscience (art. 1 UDHR). Secondly, it concerns the possibility of choosing the destiny of one's personality in the digital environment. This means that an individual could choose not to use technologies without being excluded from society (principle of equality: art. 1 UDHR). There should be a right to use digital technology without violating one's privacy (11 UDHR), without endangering one's dignity (art. 1 UDHR). Security solutions should make it possible not to be captured by a system, a robot, an algorithm, a company or a state.

C. SOME STRATEGIC RECOMMENDATIONS

Dematerialization and generalised Internet, computer and digital platform use have pushed society into an era of change. Changing ways of communicating, consuming, moving, entertaining, working, and training all influence our relationships, whether real or dematerialized. Nowadays, the 'customer experience' imposed by service providers and digital platforms, impose conditions of the interactions between people and institutions. The digital services economy mimics business models based on the exploitation of data under the control of the hegemonic actors of the Net.[19] This cloning of the models deployed prevents the development of alternative models that could better take into account the needs of personal data protection, the preservation of digital privacy and respect for fundamental rights. Technological breakthroughs do not preclude the continued need for protection and respect for fundamental human rights. They are strategic disruptions that require strategic responses above all.

We propose some strategic recommendations for a better consideration of human rights:

1. Define and publicize fundamental rights. Illustrate how fundamental rights can be undermined by digital technologies. Make this knowledge accessible, educate people.
2. Educate for global citizenship in the spirit of UNESCO[20]
3. Educate those that contribute to the development technologies, infrastructure or services about human rights.
4. Recognize that fundamental rights are non-negotiable, that they must be protected, preserved, defended, demanded in the physical world as well as in cyberspace.
5. Boycott, refuse, as far as possible, interactions with entities that do not respect fundamental rights, penalize them.
6. Do without certain services or technologies.
7. Develop a culture of respect for human rights in all areas of activity and at all stages of life.
8. Have a scientific research policy that promotes the development of solutions that respect human rights (incentives) and is based on a coherent innovation and investment policy.
9. Have an ethics and validation committee for AI solutions with an appropriate organizational structure and AI certification processes
10. Denounce abuses, cases where fundamental rights are violated (name and shame).

[19] GAFAM: Google Amazon, Facebook, Apple, Microsoft (USA); NATU: Netfix, Airbnb, Telsa, Uber (USA); BATX: Baidu, Alibaba, Tencent et Xiaomi (China).

[20] Education for global citizenship. While the world may be increasingly interconnected, human rights violations, inequality and poverty still threaten peace and sustainability. Global Citizenship Education (GCED) is UNESCO's response to these challenges. It works by empowering learners of all ages to understand that these are global, not local issues and to become active promoters of more peaceful, tolerant, inclusive, secure and sustainable societies. UNESCO, https://en.unesco.org/themes/gced (last visited Apr. 29, 2020).

11. Never consider technology to be an end in itself.
12. Understand that digital capitalism is not necessarily compatible with respect for human rights.

V. PERSPECTIVES

The 21st century is the century of ubiquitous electronic technologies and the extensive use of computers and telecommunications in all areas of life, in all activities and on a global scale. The development of cognitive sciences and their application to the computerization of society opens the door to new possibilities. In addition to the infinite potentialities and the hopes of a better world transformed by Technology, it is profoundly changing our reality on a scale never before known. Techno sciences are also at the service of the expression of new forms of power and violence.

The digital giants, especially the social networks, impose a conception of life in society, a culture and a morality that are by no means universal. These applications enclose users in bubbles that facilitate profiling and informational targeting, but also psychological manipulation, as various elections have shown all around the world. Social networks restructure the way we are in the world, the way we behave, the very dynamics of our societies.

There is an asymmetry between the digital multinationals, about which we know almost nothing, and their users (public and private organizations, individuals, including political and economic leaders), about which they know everything. They are opposed to any State and international regulation, preferring the fiction of self-regulation. More often, these multinational companies communicate intensely about their inescapable character and their social and philanthropic roles. By appropriating the humanitarian discourse and the discourse on the evolution of techno sciences, they confiscate the ability to challenge their actions, which allows them to impose the rules that suit them. We need to ask ourselves about the long-term impact of this dynamic as well as our voluntary submission. The connectivity provided by the platforms does, of course, have a positive impact. But their negative impacts are here to stay. That is particularly true when considering the ecological risk associated to cyber practices.

At a time when civil society is taking up the theme of ecological transition (preservation of resources, sustainable development, etc.), the relationship between information technology and the climate remains little explored. The climate risk has become an international global emergency, and digital technologies could help to control it under the condition that its environmental impacts are limited. But we should not forget that the all-digital world is increasingly energy-intensive, consumes non-renewable resources and generate electronics and digital toxic waste. The digital ecosystem is a powerful accelerator of climate change. In order to halt climate degradation, global warming, do we need to change our lifestyles radically, including in our digital practices? Climate change is therefore a potential threat to freedom (the right to live in a balanced environment that respects health), so what do we have to give up in order to live safely in the face of climate danger? Because of the impact of digital technologies on climate change we have to understand in this context, how does the digital world increase the deprivation of freedom and free will? If we don't have a healthy environment, we won't have freedom! What about our freedoms whether global warming, the erosion of biodiversity and cyber generate risks are not under control?

Let's take advantage of the great climate momentum to be aware of our responsibilities towards future generations, to develop exponential awareness for safe digital practices. We have to accept the fact that if the planet is in a state of emergency, it is perhaps to remind us that we humans are also in a state of emergency, state also generated by extensive uses of information technologies. We have a great opportunity to rethink our state of nature. Let us live in harmony with nature and stop destroying it through futile digital activities. Biological life is not the digital life that some entities want to impose on us.

More than ever, it is urgent to reaffirm the values of human dignity, free consent, freedom of thought and movement, fraternity and peace. Perhaps it is time to think about the digital world in terms of co-responsibility and digital downsizing. Let us innovate so that values such as empathy, tolerance, caution and non-violence can be expressed through new technologies. Let us innovate to ensure that traditional human rights are respected in cyberspace and in the physical world. Let us innovate so that new human rights specifically related to digital technologies are recognized, such as the right to disconnect (and to be disconnected), the right to not be under computerized surveillance or the right to know whether the 'person' you are interacting with is actually artificial intelligence.

DIGITAL INFRASTRUCTURES AND INEQUALITIES: POLICY AND POLITICAL CONSIDERATIONS ASSOCIATED WITH RIDESOURCING PLATFORM WORK IN AUSTRALIA

- Greg Marston & Peter Holtum[*]

ABSTRACT

Platform work has received considerable critical attention in the media, in policy debates and increasingly in academic literature. Until recently much of this attention has been focused on the personalized service industry, namely Uber and other ridesourcing platforms. Uber's arrival in Australia in 2012 underscores a diffusion of risk and responsibility in a weakening labour market. In this article, we argue that these problems are not coincidental, but that they are emblematic of the growing de-standardization of work in Australia and other countries. We explore literature on platform work, the de-standardization of work, and Uber in an Australian context, and discuss socio-legal and regulatory responses. We argue that the legal and policy considerations about the future of the personalized transport industry need to address the broader socio-economic landscape, that is eroding the rewards of paid work.

Keywords: *Inequality, Platform Work, Policy, Precarity, Regulation, Rideshare, Uber*

I. INTRODUCTION

Australia, like other advanced economies, has experienced rapid socio-technical change in recent years. A substantial shift from manufacturing industries to service industries since the 1970s has helped transform Australia's economic market around the production and intermediation of knowledge and information.[1] Rapid advances in digital technologies in the 21st century have accelerated the production and dissemination of knowledge and information. While this digitization of labour has been welcomed by some as a way to reinvigorate labour markets without the ardour of bureaucratic regulations, there has been much conjecture around this shift in labour processes and its effects on patterns of work and individual well-being. Issues related to the scope of technological unemployment, skills and retraining, and future demand for education

[*] Greg Marston is the Head of School, School of Social Sciences at the University of Queensland. Peter Holtum is a postdoctoral research fellow at the University of Queensland.

[1] E. Connolly & C. Lewis, *Structural Change in the Australaian Economy,* RBA 1, 6 (2010).

and professionalization are hotly debated in the public realm.[2] Paramount to these debates is the ability of this unprecedented pace of socio-technical changes to lead to a greater social and economic equality. While many commentators are optimistic that Artificial Intelligence (AI) and automation may create more leisure time, better quality work and more enriching lives,[3] others caution that these technological advances may lead to a pooling of economic risk at the bottom of the income scale, with race, age, class and gender differences compounding employment precarity and existing patterns of disadvantage.[4]

The political erosion of worker rights is a problem that is compounded by a range of factors, including offshoring of jobs, the digitization of management, and the continual exclusion of workers groups (like unions) from the organization and regulation of these spaces. The ramifications of these factors suggest that these political, rather than technological dimensions are vital to understanding the inequality of experiences in contemporary workplaces. For these reasons we focus solely on the political context of which Uber has emerged in Australia and other countries, rather than the technological 'novelty' of its implementation. We argue that addressing social inequality is not a matter of regulating technological change in any single particular industry, but rather addressing broader social and economic patterns that allow problematic technologies in the first case.

Uber's recent expansion into India, for example, makes it an excellent and timely case-study for the expansion of digital labour platforms. While Uber is often understood to be the second largest ridesourcing[5] company in India, behind Ola,[6] the decision to move its national headquarters from the Netherlands to India demonstrates Uber's desire to invest heavily in the Indian market. Uber's Chief Executive has frequently expressed interest in India as a 'growth

[2] ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT, The Future of Work: OECD Employment Outlook 2019, http://www.oecd.org/employment/Employment-Outlook-2019-Highlight-EN.pdf (last visited Oct. 20, 2019).

[3] TIM DUNLOP, WHY THE FUTURE IS WORKLESS (University of New South Wales Press Ltd. 2016); MARTIN FORD, THE RISE OF THE ROBOTS: TECHNOLOGY AND THE THREAT OF MASS UNEMPLOYMENT (Oneworld Publications 2015).

[4] VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (St Martins Press 2018); JENNIFER RAYNER, BLUE COLLAR FRAYED: WORKING MEN IN TOMORROW'S ECONOMY (Black Inc. 2018).

[5] We use the term ridesourcing, rather than the misnomer, 'ridesharing', in recognition of evidence that platforms like Uber provide capitalistic, for-profit business models, rather than alternate 'postcapitalistic' business models that are more emblematic of the 'sharing economy'. Peticca-Harris, Nadia C. DeGama & M.N. Ravishankar, *Postcapitalist Precarious Work and those in the 'drivers' seat: Exploring the motivations and lived experiences of Uber drivers in Canada,* 27 ORGANIZATION 36 (2020).

[6] A. Surie & J. Koduganti, *The Emerging Nature of Work in Platform Economy Companies in Bengaluru, India: The Case of Uber and Ola Cab Drivers*, 5(3) EJICLS1 36 (2016).

market' for Uber.[7] Nevertheless, evidence from other markets, like Australia, where Uber is the preferred personalized transport service,[8] demonstrates that despite the profitability of the ridesourcing industry, very little money is returned to local economies by way of driver remuneration,[9] or state taxes.[10] The social and economic effects of platform work is perhaps the most contentious aspect of the ridesourcing industry, in part, because of the central role that employment has historically played in redistributing income, dignity and social connection. Jobs matter because they are more than just a source of income. They provide valued forms of social identity and the skills built into jobs are the drivers of increased productivity.

Our aim in this article is twofold. First, we explore the growing trend of non-standard employment in the form of platform work, in Australia in a global context. We examine the history of this growth, and the manner in which it has affected labour dynamics in Australia. We draw specifically from a case study of the arrival of Uber in Australia to help identify the major effects of digital labour platforms on the broader Australian economy. Second, we examine the development of public policy and address some political considerations that impact government's capacity to deal with traditional and emerging forms of non-standard employment in the digital age. We hope that the Australian experience of Uber — both its rewards and challenges — will be of use to legal scholars and policy makers in countries like India where Uber, and the ridesourcing industry more generally, are experiencing growth.

II. 'GIG-WORK', INSECURE WORK, AND DIGITAL LABOUR PLATFORMS

Despite much conjecture about Uber's innovative prowess,[11] much academic literature suggests that its business model is simply another form of gig-work. Rosenblat and Stark,[12] for instance, demonstrate considerable asymmetries of power between Uber drivers and the Uber platform, which supports claims from the International Labour Organization, that the "gig economy" is

[7] REUTERS, https://www.reuters.com/article/us-india-uber/uber-ceo-expects-to-ride-developing-market-growth-in-next-decade-idUSKBN1X11PPX (last visited Oct. 23, 2019).

[8] ROY MORGAN, http://www.roymorgan.com/findings/8098-uber-overtakes-taxis-june-2019-201908260239 (last visited Aug. 26, 2019).

[9] Jim Stanford, *Subsidising Billionaires: Simulating the New Incomes of UberX Drivers in Australia,* TAI 1, 7 (2018).

[10] AUSTRALIAN FINANCIAL REVIEW, https://www.afr.com/technology/uber-australia-makes-785m-pays-8-5m-tax-20190501-p51j62 (last visited May 2, 2019).

[11] Geoffrey Dudley, David Banister, Tim Schwanen, *The Rise of Uber and Regulating the Disruptive Innovator,* 88(3) THE POLITICAL QUARTERLY 492, 499 (2017); Yanwei Li, Araz Taeihagh & Martin de Jong, *The Governance of Risks in Ridesharing: A Revelatory Case from Singapore,* 11(5) ENERGIES 1277 (2018).

[12] Alex Rosenblat & Luke Stark, *Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers,* 10 INTERNATIONAL JOURNAL OF COMMUNICATION 3758, 3778 (2016).

not a new way of working, rather it is simply another reiteration of outsourcing.[13] 'Gig-work' is often defined as work that utilizes odd jobs (i.e. gigs). In contrast to the workers who are employed by an organization, gig-workers ideally "support themselves as flexible, free independent suppliers, moving seamlessly from one job (or 'gig') to another".[14] Stanford[15] defines gig-work as characterised by any one of five aspects: (i) work is performed on an on-demand or as-needed basis, (ii) work is compensated on a piece-work basis, (iii) producers are required to supply their own capital equipment (*i.e.* tools, cars, and even a place of work), (iv) a triangular relationship between the producer, the end-user, and the intermediary, and (v) some form of digital intermediation is utilised to commission the work, supervise it, deliver it to the final customer, and facilitate payment. Despite the recent interest in gig-work as a factor of technological change in labour markets, many have pointed out that gig-work is no new phenomenon, and can be linked to a political, rather than technological shift in society.[16] For instance, vocations in the farming, postal, and construction industries have relied on forms of gig-work for centuries. Moreover, as Stanford points out, of the five dimensions of gig-work, only *digital intermediation* has any novelty. Subsequently, we argue — alongside a host of others[17] — that the growth of gig-work in the 21st century emerges as a product of political changes (*i.e.* in regulations, policy, and culture), rather than technological developments in the workplace.

Paramount to the re-emergence of gig-work in the 21st century has been its effect on the destabilization and deterioration of job security and labour practices. Gig-work affects more than workplaces. In Australia's work-centric society the growth of gig-work renders social, economic, and cultural dimensions of humanity insecure. This disruption of the working model exposes individual workers to ontological insecurity, whereby the temporal dimensions of work (*e.g.* work time, skills acquisition, mastery of craft) fall to the realm of the individual or private sphere, rather than the organizational sphere of an employer. While we recognize that insecure work has been used in a myriad of contexts,[18] we understand insecure work in this article simply as work that is considered risky from the point of view of the worker. This risk factor can be defined by

[13] THE WORLD BANK, http://www.worldbank.org/en/news/feature/2015/12/22/regulating-the-gig-economy (last visited Dec. 22, 2015).

[14] Jim Stanford, *The resurgence of gig work: Historical and Theoretical Perspectives*, 28(3) ELRR 382, 384 (2017) [hereinafter Jim Stanford].

[15] *Id.*

[16] DAVID PEETZ, THE REALITIES AND FUTURES OF WORK (Australian National University Press 2019); GUY STANDING, THE PRECARIAT: THE NEW DANGEROUS CLASS (Bloomsbury 2011).

[17] *Id.*; Jim Stanford, *supra* note 14.

[18] EMILIANO ARMANO, ARIANNA BOVE, & ANNALISA MURGIA, MAPPING PRECARIOUSNESS, LABOUR INSECURITY AND UNCERTAIN LIVELIHOODS: SUBJECTIVITES AND RESISTANCE (Routledge 2017); Iain Campbell & Robin Price, *Precarious Work and Precarious Workers: Towards an Improved Conceptualisation,* 27(3) ELRR 314, 317 (2016).

more than the lack of continuing tenure in a position, but also the declining quality/status/ ability of a job to provide for its workers.[19]

Much confusion about the extent of insecure work in Australia has emerged from this conflation of insecure job tenure with job status insecurity. For instance, the then federal Minister for Small and Family Business, Craig Laundy argued in a radio interview that there was no work crisis[20] in Australia because there has been no increase in the amount of casual employment in Australia. While it is true that the Australian Bureau of Statistics[21] reports that 25% of the workforce do not have leave entitlements in Australia, the Australian Council of Trade Unions (ACTU) points out that if we include gig-workers, labour hire workers and casual employees, then at least 40% of the Australian workforce is engaged in insecure work (2018).[22] While it is important to recognize (as the ACTU does) that there is no necessary causal relation between insecure work, and insecure workers,[23] this significant shift in the dynamic and experience of work has a significant effect over the role that work plays in social structure, status and meaning. Determining how significant these emerging trends are on the individual and social level depends on the extent of social regulations, protections, and worker support in each country.

III. THE DE-STANDARDIZATION OF WORK IN AUSTRALIA

As highlighted above, gig-work is not a new phenomenon in Australia. Gig-work has for a long time been part of how labour is contracted in Australia. Australia Post — Australia's nationally owned post office agency, for example, has been employing independent contractors (or gig-workers) to deliver mail for at least 200 years. Similarly, labour hire agencies have been a vital feature in the growth of the service industry, since the 1970s.[24] More recently, statements from farmers during the 2016 Queensland Inquiry into Labour Hire made it abundantly clear that the

[19] Duncan Gallie, Alan Felstead, A Francis Green, & Hande Inanc, *The Hidden Face of Job Insecurity. Work, Employment and Society*, 31 OECD Statistics Directorate 1 (2016); Arne Kalleberg, *Precarious Work, Insecure Workers: Employment Relations in Transition*, 74 AM. SOCIOL. REV. 1 (2009).

[20] ABC RADIO NATIONAL., https://www.abc.net.au/radionational/programs/backgroundbriefing/background-briefing-05.08.2018/10067234 (last visited August 5, 2018).

[21] AUSTRALIAN BUREAU OF STATISTICS, https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/6333.0Media%20Release1August%202018?opendocument&tabname=Summary&prodno=6333.0&issue=August%202018&num=&view (last visited November 29, 2018).

[22] AUSTRALIAN COUNCIL OF TRADE UNIONS, https://www.actu.org.au/media/1033868/insecure-work_final-18052018-final.pdf. (last visited Oct. 5, 2018).

[23] Iain Campbell & Robin Price, *Precarious Work and Precarious Workers: Towards an Improved Conceptualisation*, 23(7) ELRR 314 (2018).

[24] AUSTRALIAN GOVERNMENT, https://www.business.gov.au/~/media/Business/Independent-contractors/Independent-contractors-the-essential-handbook-pdf.pdf?la=en (last visited Nov. 5, 2016).

farming and horticulture industry had been built around the seasonal labour supply of gig-workers, without whom most farmers would not be able to operate.[25]

In the late 1980s a rapid growth in the provision of labour hire work[26] helped the rapid de-standardization of work throughout the 90s and into the 2000s. The de-standardization of this workforce is believed to have been exacerbated by the widespread establishment of workplace enterprise bargaining agreements, the introduction of Human Resource Managers, and historically low unionizations rates: all of which resulted in a more rapid uptake of Labour Hiring policies around Australia practices.[27] Bramble[28] reports that between 1988 and 1990 employers at 1/3 of all workplaces with more than 20 staff undertook major restructuring of work practices. The same number introduced 'major new plant or technology'. While in workplaces employing more than 500 people, the proportion rose to 50%. Moreover, in the period of 1984-1994 the share of total household disposable income as well as average real household disposable incomes declined for all but the highest quintile of Australian workers.[29]

Recent statistics about the labour force in Australia demonstrate that the share of workers in full-time paid employment with leave entitlements is now less than 50%.[30] This fact is complemented by a growth in part-time work for self-employed workers,[31] a growth in underemployment and overemployment.[32] The persistence of these facts against the backdrop of a largely stabilized rates of casualization in Australia lend weight to arguments about the growth of outsourcing, and triangular supply chains used by companies to avoid 'formal' avenues of employment. For example, David Peetz[33] argues:

[25] Queensland Parliament, *Inquiry into the practices of the labour hire industry in Queensland*, 55 Finance and Administration Committee 1, (2016), https://www.parliament.qld.gov.au/documents/tableOffice/TabledPapers/2016/5516T1028-.pdf.

[26] John Burgess, & Iain Campbell, *Casual Employment in Australia: Growth, Characteristics, A Bridge or a Trap,* 9 ELRR 31, 54 (1998).

[27] AUSTRALIAN GOVERNMENT, https://www.pc.gov.au/research/supporting/labour-hire/labourhire.pdf (2002).

[28] TOM BRAMBLE, TRADE UNIONISM IN AUSTRALIA: A HISTORY FROM FLOOD TO EBB TIDE 151 (Cambridge University Press 2008).

[29] *Id.* at 179.

[30] Tanya Carney, & Jim Stanford, *The Dimensions of Insecure Work: A Factbook*. THE AUSTRALIA INSTITUTE: CENTRE FOR FUTURE WORK, (2018), https://www.futurework.org.au/the_dimensions_of_insecure_work.

[31] *Id.*

[32] Daniel Zilfer, *Underemployment and overemployment problems leave few workers satisfied*, ABC NEWS, (Sept. 9, 2019, 10:07 AM), https://www.abc.net.au/news/2019-09-10/working-50-hours-week-you-might-be-destroying-economy/1149-4196; Greg Jericho, *Looking for the cause of low wage growth? It's underemployment,* THE GUARDIAN (June 20, 2017, 10:39 AM), https://www.theguardian.com/business/grogonomics/2017/jun/19/.

[33] DAVID PEETZ, THE REALITIES AND FUTURES OF WORK (Australian National University Press 2019).

"*A lot of variation between industries and periods is hidden by aggregate figures. Franchising has grown in retailing. Labour hire in mining. Outsourcing in the public sector. Second jobs in manufacturing. Spin-offs in communications. Casualization in education and training. Global supply chains send jobs overseas to low-paid, often dangerous workplaces in a number of industries.*"

The steady growth of digital labour platforms (like Uber, Airtasker, and Airbnb)[34] means that these insecure, non-standard forms of work have greater avenues in which to grow. The growth of such insecurity has typically created significant socio-legal issues for workers, worker groups, and regulators. For example, Stewart and Stanford[35] argue that "the growth of the gig-economy (…) poses fundamental challenges to traditional models for regulating work and setting minimum standards". Stewart and Stanford[36] are concerned about the ability of conventional legal regulations to stem the growth of insecurity in these jobs:

"*It is not clear that existing regulations apply to gig workers, let alone that they can be effectively enforced in the digital economy. In fact, in some cases, evading traditional regulations (not to mention taxes) appears to have been a key rationale for establishing digital businesses in the first place.*"

International bodies like the International Labour Organization and the World Bank[37] agree that enforcing labour market protections can be difficult because of the increasingly complex supply chains and sub-contracting involved in working arrangements. Instead these organisations suggest that "countries need to extend their social and labour protections to people who are working part time or in serial part time or temporary jobs".[38] In a similar vein, the call for broader social protections to *all workers* is a core feature of our discussion on socio-legal and regulatory responses to platform work in the final section of this article. Key to our argument is a concern with the ability of global organisations to continually use digital technologies to avoid national legal regulations. Instead of retrospectively responding to regulatory breaches or "technological innovators", we propose a system that proactively supports insecure workers. Before we explore such a system, however, we examine the regulation of workers under the Uber model in Australia, which has blossomed against this backdrop of insecurity in the Australian labour

[34] Paula McDonald, Penny Williams, Andrew Stewart, Damian Oliver, & Robyn Mayes, *Digital Platform Work in Australia, Victoria*, VICTORIAN DEPARTMENT OF PREMIER AND CABINET (2019).

[35]Andrew Stewart, & Jim Stanford, *Regulating Work in the Gig Economy: What are the Options*? 28(3) ELRR 420 (2017), https://pdfs.semanticscholar.org/d6ec/825f335ca4241b265f41972ef57d8d430ee3.pdf.

[36] *Id.*

[37] The World Bank, *Regulating the Gig Economy*, THE WORLD BANK (Dec. 22, 2015, 06:15 PM), http://www.worldbank.org/en/news/feature/2015/12/22/regulating-the-gig-economy.

[38] *Id.*

market. Specifically, we demonstrate three specific examples in which the ridesourcing platform creates and *capitalizes* upon social inequalities amongst workers and illustrate how a lack of vision with policy permits these issues to endure.

IV. THE GROWTH OF UBER IN AUSTRALIA AND ITS SOCIAL CONSEQUENCES

Uber Technologies Inc. is officially listed as a technology company that operates the Uber ridesourcing application which connects riders with drivers. By positioning themselves as an intermediary platform to connect riders with drivers, Uber clearly utilizes a triangular relationship that is exemplary of gig-work arrangements. As a company, Uber was first registered in 2009 and did not offer a service until 2010 when it enlisted some drivers in San Francisco. Since then, Uber has spread to 700 cities, 63 countries, and offers 14 million rides a day.[39] Despite this significant growth, Uber has registered annual economic losses around the world,[40] which has created concerns about the extent to which Uber invests into new markets, as well as Automatic Vehicles research, and the data 'assetization' of passengers and drivers.

Uber has significantly affected the transport industry since arriving in Australia in 2012. Its aggressive entry into the transport industry is well documented by its rivalry with traditional taxis. In August, 2019 Roy Morgan reported that Uber has officially overtaken traditional taxis as the preferred private transport service of Australians. Other effects of Uber's aggressive entry to the Australian market have been demonstrated by the drastic decline of Taxi licenses shares reported around the country in 2016 when Uber was rendered legal;[41] the effects of which drastically affected taxi drivers who had invested heavily in their licenses.[42]

Reasons for Uber's rapid growth are generally believed to be its ability to outcompete traditional taxis on a number of factors. On an average, Uber is: cheaper for users than taxis,[43] has a higher 'user experience rating',[44] and (according to research in the United States) is more efficient in

[39] UBER TECHNOLOGY INC., https://www.uber.com/en-AU/newsroom/company-info/ (Dec. 2, 2019, 05:15 PM).

[40] Geoff Thompson, *Uber X drivers working for half the minimum wage, new report shows*, ABC NEWS (March 17, 2018, 07:14 PM), https://www.abc.net.au/news/2018-03-06/uber-x-drivers-working-for-half-the-minimum-wage/9513250.

[41] THE BRISBANE TIMES, https://www.brisbanetimes.com.au/politics/queensland/value-of-taxi-licences-plummets-across-queensland-even-without-uber-20180220-p4z0yo.html. (last visited Feb. 20, 2018).

[42] ABC RADIO NATIONAL, https://www.abc.net.au/radionational/programs/backgroundbriefing/background-briefing-05.08.2018/10067 (last visited Aug. 5, 2018). 234

[43] Jim Stanford, . *Subsidising Billionaires: Simulating the New Incomes of UberX Drivers in Australia,* TAI 1, 7 (2018); CHOICE, https://www.choice.com.au/transport/cars/general/articles/uberx-vs-taxi-which-one-is-best (last visited Aug. 14, 2018).

[44] *Id.*

terms of passengers per mile travelled.[45] Despite being illegal until 2015-2017 (depending on each Australian State/Territory), Uber continued to operate and expand around Australia (a process of aggressive expansion that has been well discussed).[46] Its expansion has been attributed to its cheap rates for passengers (and drivers), a transparent user interface that promotes safety through real time passenger and driver surveillance, and its use of promotional packages to incentivize drivers in new markets.[47] In 2015, for instance, Uber assisted drivers in Queensland by paying $1.7 million in fines issued to driver-partners for operating illegally.[48] However, since regulation (and the subsequent influx of drivers), Uber has stopped offering many benefits to its drivers and has increased operating charges to both users and drivers.[49] Furthermore, legal rulings from Australia's Fair Work Commission have ruled that the working relationship between drivers and the platform is not an "employment relationship"[50] thus allowing Uber to continue operations with little interference from industrial regulations.

The effects of Uber's growth in the Australian market have been mixed; in part because of its market success with users, but also because of a myriad of controversial outcomes for its driver-partners. We will briefly explore three ways in which Uber's expansion has increased insecurities for workers. However, we emphasize that while Uber has created new problems for Australians and Australian workers, it has also helped tens of thousands of workers access an income source, of which many would not have been able to access without the emergence of this platform. The complexity of this situation embellishes the thrust of our argument for the need for policy to address structural inequality in the labour market, rather than chasing rapidly advancing technological processes.

There are a number of regulatory and socio-legal issues that need to be considered. First, Uber's infamous ability to classify its drivers as independent contractors rather than employees on a payroll has been instrumental in the underemployment of its workers. Recent data from

[45] Judd Cramer & Alan Krueger, *Disruptive Change in the Taxi Business: The Case of Uber*, 106 AER 177 (2015).

[46] ALEX ROSENBLAT, UBERLAND: HOW ALGORITHMS ARE REWRITING THE RULES OF WORK 144 (University of California Press 2018) [hereinafter ROSENBLAT].

[47] *Id.* at 70.

[48] BRISBANE TIMES, https://www.brisbanetimes.com.au/national/queensland/uber-queensland-drivers-fined-17-million-in-three-months-20160719-gq8zv3.html (last visited Jul. 19, 2016).

[49] THE CANBERRA TIMES, https://www.canberratimes.com.au/business/workplace/when-i-first-started-the-money-was-very-good-ride-share-drivers-feeling-the-squeeze-20181022-p50b90.html (last visited Oct. 24, 2018).; Although some of these incentives have since reappeared in Queensland following the emergence of rival company, DiDi in 2019.

[50] FAIR WORK OMBUDSMAN, https://www.fairwork.gov.au/about-us/news-and-media-releases/2019-media-releases/june-2019/20190607-uber-media-release (last visited June 7, 2019).

Australia[51] supports research from the United States[52] and the United Kingdom[53] to demonstrate the poor income rates of Uber drivers. Union commissioned reports from Australia confirm the poor rates of income and suggest that fewer than 40% of drivers (of whom predominantly work on the Uber platform) contribute to superannuation.[54] In Australia it was estimated that Uber drivers receive an average of $14.62 an hour — which was significantly beneath the statutory minimum wage at the time of $18.29.[55] This figure is less than half of the industry rate that would normally apply to workers in this sector[56] who are expected to be additionally compensated for working at unsociable hours. The lack of accountability to industrial regulations and labour laws presents the first and perhaps most troubling issue for policy makers to address with regards to future work. By contracting drivers through a digitally mediated employment system, Uber is able to legally distance itself from conventional organizational responsibilities for its drivers. This lack of organizational responsibility causes numerous troubles beyond just the low wages; the ramifications of which render "driver-partners" more insecure in broader social life.[57]

Second, Uber's unregulated expansion in cities appears to have adversely affected urban areas, and contributed to congestion, and pollution. Reports of increasing competition between drivers for customers, space, and parking spots[58] emphasizes research (from the United States) that "ride-hailing attracts Americans away from bus services and light rail services".[59] Similar research[60] found that at "least half of ridesourcing trips replaced modes other than taxi, including public transport and driving". Thus, while ridesourcing appears to be beneficial to users, it is not clear

[51] Jim Stanford, *Subsidising Billionaires: Simulating the New Incomes of UberX Drivers in Australia*, TAI 1, 7 (2018).
[52] Stephen Zoepf, Stella Chen, Paa Adu & Gonzalo Pozo, *The Economics of Ride Hailing: Driver Revenue, Expenses and Taxes*, MIT CENTER FOR ENERGY AND ENVIRONMENTAL POLICY RESEARCH (Feb. 5, 2018) https://orfe.princeton.edu/~alaink/SmartDrivingCars/PDFs/Zoepf_The%20Economics%20of%20RideHialing_OriginalPdfFeb2018.pdf; Revised findings were published in a public statement from Zoepf (2018).
[53] Berger, T., Frey, C., Levin, G., & Danda, S. R., *Uber Happy? Work and Wellbeing,* Working paper to be presented at the 68th Panel Meeting of Economic Policy in October 2018 (2018).
[54] The Rideshare Cooperative & Transport Workers Union, *Rideshare Driver Survey: The Rideshare Cooperative & Transport Workers Union* (Dec. 20, 2018) http://www.twu.com.au/Rideshare-Survey-Infographic/
[55] Jim Stanford, *Subsidising Billionaires: Simulating the New Incomes of UberX Drivers in Australia*, TAI 1, 7 (2018).
[56] *Id.*
[57] Peter Holtum, Greg Marston, *Flexibility and Insecurity: An Insight into the experiences of Uber drivers in Brisbane*, UQ 3, 7 (2019), https://social-science.uq.edu.au/article/2019/05/flexibility-and-insecurity-insight-experiences-uber-drivers-brisbane.
[58] THE CANBERRA TIMES, https://www.canberratimes.com.au/business/workplace/when-i-first-started-the-money-was-very-good-ride-share-drivers-feeling-the-squeeze-20181022-p50b90.html (Oct. 24, 2018); JP Morgan Chase & Co, *The online platform economy in 2018: Drivers, Workers, Sellers, and Lessors,* JPMORGAN CHASE & CO, (Sept. 18th, 2018, 07:30 PM), https://institute.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/institute/pdf/institute-ope-2018.pdf.
[59] Regina Clewlow & Gouri Mishra, *Disruptive Transportation: The Adoption, Utilisation, and Impacts of Ride-Hailing in the United States, UCDITS 1, 5* (2017).
[60] Lisa Rayle, Danielle Dai, Nelson Chan & Robert Shaheen, *Just a Better Taxi: A Survey Based Comparison of Taxis, Transit, and Ridesourcing Services in San Francisco*, 45 TRANSPORT POLICY 168, 178 (2016).

that its overall effects on urban environments is beneficial or not. Moreover, such an increase in the amount of car-traffic on the roads in urbane environments is "likely" to contribute to the growth of congestion in cities, and increase carbon emissions.[61] This suggests that ridesourcing companies like Uber may contribute to the patterns of congestion and greenhouse gas emissions that they often claim to mitigate through their existence, if unchecked/unregulated by local councils.[62] While ridesourcing technology offers incredible potential to develop urban spaces, it nevertheless needs to be carefully integrated to the social and political environment.

Third, the relatively low-bar of entry to become an Uber driver appears to have attracted a group of economically vulnerable drivers who readily rely on this irregular service to support their family. While demographic data from Australia does not yet exist,[63] reports from Australia support these assertions of driver vulnerability, suggesting that Uber drivers are likely to have irregular incomes, have restricted rights to work because of migration agreements,[64] and/or have families to support.[65] Similar research from the United Kingdom[66] demonstrates that Uber drivers are "overwhelmingly male immigrants" from "Black, Bangladeshi, and Pakistani ethnic groups". Most were married (70%), have children (64%), and 82% were low-skilled immigrants.[67] The research suggests that it is likely that Uber offers these populations a better form of work than they might otherwise be able to access, given their low socio-economic status.[68] Consequently, while such research does not establish a causal link between gig-workers and social vulnerability, it is clear that vulnerable populations are over-represented in this line of 'work' and need social protections.

[61] Regina Clewlow & Gouri Mishra, *Disruptive Transportation: The Adoption, Utilisation, and Impacts of Ride-Hailing in the United States, UCDITS 1, 5* (2017); Lisa Rayle, Danielle Dai, Nelson Chan & Robert Shaheen, *Just a Better Taxi: A Survey Based Comparison of Taxis, Transit, and Ridesourcing Services in San Francisco*, 45 TRANSPORT POLICY 168, 178(2016).

[62] Infrastructure Partnerships Australia, *Driving Change: Australia's Cities Need a Measured Response*, Sydney, IPA 1, 5 (2016), https://infrastructure.org.au/wp-content/uploads/2016/10/Uber-Policy-Paper-Update.pdf.

[63] It is understood that Uber have commissioned a report on driver demographics in Australia that will be forthcoming in 2020.

[64] Peter Holtum & Greg Marston, *Flexibility and Insecurity: An Insight into the experiences of Uber drivers in Brisbane*, UQ 3, 7 (2019), https://social-science.uq.edu.au/article/2019/05/flexibility-and-insecurity-insight-experiences-uber-drivers-brisbane.

[65] THE COURIER MAIL, https://www.couriermail.com.au/news/small-uq-study-finds-better-to-drive-uber-than-work-in-hospitality-or-some-other-industries (May 9, 2019).

[66] Thor Berger, Carl Frey, Guy Levin & Santosh Danda, *Uber Happy? Work and Wellbeing,* OXON 3, 42 (2018).

[67] *Id.* at 47.

[68] *Id.*

In addition, while the data suggests that Uber offers migrant populations a better form of work than they might otherwise be able to access given their low socio-economic status,[69] it highlights the prominence of vulnerable social groups engaging with this precarious avenue of work.

In addition, such reliance on insecure work from populations who are likely to have little fiscal, social, and cultural capital is problematic in light of the poor income, and lack of industrial support offered to Uber driver-partners.[70] Such vulnerability is exemplified in a recent safety report about drivers in the ridesourcing industry in Australia that suggests a significant proportion of drivers have reportedly been assaulted: 37% of drivers have received threats from passengers which included "death threats, threats of rape and racial abuse", 10% of drivers have been physically assaulted, and over 6% have been sexually assaulted.[71] Media reports suggest that while Uber has been settling some of these issues with their drivers, the settlements are performed on a case by case basis behind closed doors.[72] Without a transparent, and equitable process it is unclear how drivers – the majority of whom are likely to be immigrants with little understanding of the Australian legal system — can be expected to pursue grievances on their own time, money, and resources. Thus, the third consideration for policy makers is this diffusion of responsibility and personal safety in which gig-workers are forced to rely on their own capital to secure lawful and equitable treatment as workers.

These three issues present significant problems for policy makers and regulators as non-standard employment continues to grow in Australia thanks — in part — to digital labour platforms. As a case study, we suggest Uber is not the exception but is rather *emblematic* of insecurity in the gig-economy. The lack of secure and predictable pay for workers who are not guaranteed a legal status of employment, the lack of organisation between the public and private spheres in urban environments, the centrality of vulnerable migrant populations, and the erosion of workplace responsibility at the hands of a triangular employment relationship pose significant issues for the future of not just paid work, but society as a whole. To ensure a just transition of work into the future will require fundamental change, rather than tinkering around the edges of policy and legislation on industry specific problems. The case study of transport is a small story that reveals some major challenges for policy and regulation more generally.

[69] *Id.*

[70] ROSENBLAT, *supra* note 46; THE RIDESHARE COOPERATIVE & TRANSPORT WORKERS UNION, http://www.twu.com.au/Rideshare-Survey-Infographic/ (Oct. 20, 2019).

[71] *Id.*

[72] THE AUSTRALIAN FINANCIAL REVIEW, https://www.afr.com/news/policy/industrial-relations/uber-settling-unfair-dismissal-claims-for-deactivated-drivers-20170629-gx0z2e (June 29, 2017).

V. SOCIO-LEGAL AND REGULATORY RESPONSES TO PLATFORM WORK

Efforts from unions, worker associations, cooperatives, lawyers, and legal scholars to reclassify the working relationship between drivers and their employers in the gig-economy can and will continue to help current workers being exploited by organizations.[73] As Stewart and Stanford[74] argue: "without adjusting and strengthening" pre-existing "labour regulations and safety nets to reflect new practices of gig-work, the prospect of building an inclusive, fair labour market…will be set back all the more". Nevertheless, these regulatory approaches leave unchanged, the broader and underlying features of society that continue to create exploitative businesses that will affect *future* workers. In other words, we need a regulatory policy that addresses the broader social and cultural practices that all too often reward exploitative business models that create precarious working conditions.

For governments and regulatory bodies, there are at least three dimensions to consider when implementing new rules and guidelines: *technological infrastructure* which facilitates information transmission and distribution between different service providers, consumers, drivers, and other participants; *economic infrastructure* that enables logistic resources or finances to be transacted in a certain way, and thus develop sustainable business models; and *cultural infrastructure* which refers to different stakeholders or actors who may or may not share economic interests but which generate cultural and social norms in the course of the communal exchange and interaction.[75] While much of the public discussion about regulation focuses on technology and the economy, the social and cultural norms around how work is changing and the capacity for adaptation receive less attention. There is a certain comfort, for example, that comes from thinking that re-skilling the workforce for the jobs of tomorrow will keep humans ahead of the technological curve. However, there are a number of limitations with this approach, not least of which is social identity and purpose. As Jennifer Rayner explains:[76]

"*The suggestion that the blue-collar men being displaced in their tens of thousands by technological and economic change can simply transfer into new jobs in entirely unrelated industries is either an expression of embarrassing ignorance or a conscious bait and switch.*"

[73] Nicholas DeBruyne, *Uber Drivers: A Disputed Employment Relationship in Light of the Sharing Economy*, 92(1) CHI KENT L. REV 289, 310 (2017); Andrew Stewart & Jim Stanford, *Regulating Work in the Gig Economy: What are the Options?* 28(3) ELRR 420, 421 (2017); Valerio De Stefano, *The Rise of the Just-in-Time Workforce: On-Demand Work, Crowdwork, and Labor Protection in the Gig-Economy*, 37 COMP. LAB. L. & POL'Y J. 471, 501 (2016).

[74] Andrew Stewart, & Jim Stanford, *Regulating Work in the Gig Economy: What are the Options?*, 28(3) ELRR 420, 421 (2017).

[75] Janet Xue, Alex Chung & Ying Yu, *Legal and Regulatory Challenges of the Sharing Economy*, FLJS 2, 7 (2018).

[76] JENNIFER RAYNER, BLUE COLLAR FRAYED: WORKING MEN IN TOMORROW'S ECONOMY 9 (Black Inc. 2018).

This problem is significant in light of the disruption of the labour market and the situation gig-workers and independent contractors find themselves in; without wage, organizational, and even task continuity it is tough for these workers to adapt to social changes and developments. The effect of digital disruption has been evident in the stories and accounts of people trying to make sense of rapid change. Media coverage of the impact of Uber on the taxi industry demonstrates the significant personal toll that uncertainty about the future takes on the individual workers.[77] The result is a doubt about the efficacy of the self, about one's agency, which reinforces anxiety and insecurity; resulting in a vicious cycle to create more power over things. Other drivers organize, pursue class actions and vent their frustration about social media and its ability to turn products like cars into services for use, and transform employees into independent contractors with all the attendant risks. The terrain between capital and labour remains contested and it has become clear that regulatory attempts to make organisations like Uber compliant with Australian laws have been far from sufficient.[78]

VI. RESKILLING

Perhaps the most practical policy response to this growing divide between groups of workers has been to 'reskill' or 'upskill' workers who might otherwise be 'left-behind'. These arguments permeate the narrative that the impacts of AI and automation will be directed towards routine, low-skilled jobs. They feed the narrative adopted widely by global interest groups, and 'data analysts'[79] that reskilling the workforce through a program of 'lifelong education' is achievable and ideal. Nevertheless, as Rayner argues above — this suggestion mischaracterizes the deep connection between social identity, and culture with certain types of work. Moreover, even in the event of such policies, a massive inequality between groups of skilled, and non-skilled, workers would emerge.[80] Helsper and Reisdorf[81] conceptualize a 'digital underclass', similar to what Huws[82] refers to as a 'cybertariat', when they warn of the threat of a future society in which exploitation, class disparities, and social exclusion manifest in new forms.

[77] ABC RADIO NATIONAL, https://www.abc.net.au/radionational/programs/backgroundbriefing/background-briefing-05.08.2018/10067234 (last visited Aug. 5, 2018).

[78] Andrew Stewart & Jim Stanford, *Regulating Work in the Gig Economy: What are the Options*? 28(3)ELRR 420 (2017), https://pdfs.semanticscholar.org/d6ec/825f335ca4241b265f41972ef57d8d430ee3.pdf.

[79] DELOITTE ACCESS ECONOMICS, https://www2.deloitte.com/au/en/pages/building-lucky-country/articles/path-prosperity-future-work.html#contact (last visited Dec. 5, 2019).

[80] Jaz Choi, *Inquiry into the Australian Government's Role in the Development of Cities*, BRISBANE: QUEENSLAND UNIVERSITY OF TECHNOLOGY (2017).

[81] Ellen Helsper & Bianca Reisdorf, *The Emergence of a "Digital Underclass" in Great Britain and Sweden: Changing Reasons for Digital Exclusion,* 19(8) NEW MEDIA & SOCIETY 1253, 1261 (2016).

[82] URSULA HUWS, LABOR IN THE GLOBAL DIGITAL ECONOMY: THE CYBERTARIAT COMES OF AGE (New York University Press 2014).

Loi[83] in particular cautions that the growing substitution of human work with computer-driven automation may lead to "technological unemployment" on a large scale, which will eventually lead to "human disenchantment" with the realm of labour.[84] In fact, Loi points out,[85] automation and digitization have already caused deskilling and the "worsening of human individual abilities and expectations" in some cases. More specifically, Loi argues that digital technologies have the ability to emulate cognition and service standards that are often common in middle income jobs and professions. Loi suggests that "by eliminating predominantly middle-skills, middle-class jobs" these technologies have the ability to 'disenhance' more workers than they 'enhance'.[86] That is, Loi argues, digital technologies may force workers to find jobs that are less amenable to "automation, but which, paradoxically, may turn out to be less desirable than the jobs most humans could find in the past".[87] Examples from Australia suggest this kind of replacement of middle-class jobs is already underway. For instance, in the same way that Uber replaced a top-heavy taxi industry, in 2018 the National Australia Bank replaced six thousand employees with digital software technologies.[88] As such, reskilling workers will not protect workers if their industry is rapidly changing.

Reskilling workers is not simply about giving people *more* skills; instead we need to think about the *value* of these skills being taught, and *why* they may be worthwhile. Intelligent, creative, and emotional skills are still considered less replaceable by machines, and therefore still largely "secure" in the wake of the digital revolution for the foreseeable future. Non-routine, uniquely human skills that focus on care, creativity, and human consciousness will be critical. By combining technical and interpersonal tasks in their work, these individuals may become the "new artisans"[89] of the new age. Future workers who are able to use these skills and deliver specialised services based on these skills will likely to remain competitive in the job market as well as craft meaningful social and human interactions. However, as mentioned above — these abilities will not be possible for *all*. Moreover, as digital technologies increase in uptake, they risk de-humanizing work which may suppress the need for human qualities such as creativity, care, touch and communication. In sum, a major risk associated with the digital technological

[83] Michele Loi, *Technological Unemployment and Human Disenhancement*, 17(3) ETHICS INFTECHNOL 201, 207 (2015) [hereinafter Loi].
[84] Schumpeter (1975) had a similar argument about the *creative destruction* of human labour.
[85] Loi, *supra* note 83.
[86] *Id.*
[87] *Id.*
[88] Daniel Zilfer, *NAB workers latest to fall as automation transforms the economy,* ABC NEWS (Feb. 21, 2018, 10:45 PM), https://www.abc.net.au/news/2018-02-21/nab-robots-taking-over-white-collar-jobs/9465524.
[89] Lawrence Katz & Robert Margo, *Technical Change and the Relative Demand for Skilled Labour*, *in* HUMAN CAPITAL IN HISTORY 15, 19 (Boustan, C. Frydman, and R.A. Margo, eds., University of Chicago Press 2014).

revolution is the risk of polarization between highly-skilled 'future-proof' working elite and a substantial group of workers with obsolete skills who will be at risk of being excluded from the emerging labour market.

CONCLUSION

In this article we have argued that Uber is simply another iteration of gig-work that has been evolving in Australia since the 1970s. We have been critical of approaches to Uber as a "technological innovator", and instead view the growth of Uber in Australia as one that has successfully capitalized on the insecure working conditions of a highly segmented Australian labour market. The precarious state of Australian labour conditions for a significant number of workers provided Uber with a supply of workers who could be contracted without the need to guarantee employment, nor reasonable rates of pay. As such, we argue for a more considered regulatory approach that seeks to address the cause of social inequality, rather than attempting to remedy the conditions in specific industry circumstances. This broader regulatory and social policy approach is essential to create a coherent vision for a preferable, rather than a predictable future for low-income citizens.[90]

As the power of our technology grows so do our future possibilities. This potential increases the importance of having clarity about our goals. Machines are not very good at this large-scale planning and creativity, but humans are, as Schwab[91] argues in his account of how we need to shape a preferable, rather than a predictable future:

"*Neither technology nor the disruption that comes with it is an exogenous force over which humans have no control. All of us are responsible for guiding its evolution, in the decisions we make on a daily basis as citizens, consumers and investors. We should grasp the opportunity we have to shape the Fourth Industrial Revolution and direct it towards a future that reflects our common objectives and values.*"

Debating the future of work can actually provide us an opportunity to map out a more equal society.[92] Addressing the risks and embracing the opportunities presented by AI and automation will require public debate and political leadership about how best to share the rewards from productivity gains and profits. Instead of asking the question, 'what will technology do to us' we

[90] Jamie Morgan, *Will we work in twenty-first century capitalism? A critique of the fourth industrial revolution literature*, 48 ECON. SOC. 371, 378 (2019).
[91] Klaus Schwab, *The Fourth Industrial Revolution,* W.E.F. 174 (2016).
[92] THE CONVERSATION, https://theconversation.com/automation-has-the-potential-to-improve-gender-equality-at-work-96807 (June 11, 2018).

should start with the question 'what do we want to do with technology'? These questions are not new. The philosopher Bertrand Russell provided a similar provocation to policy makers in the 1930s:

"*Modern methods of production have given us the possibility of ease and security for all; we have chosen instead to have overwork for some and starvation for others. We have continued to be as energetic as we were before there were machines. In this we have been foolish, but there is no reason to go on being foolish forever.*"[93]

[93] 1 BERTRAND RUSSELL, EDUCATION AND THE MODERN WORLD 4 (W. W. Norton & Co 1932).

ALGORITHMIC CAB AGGREGATORS AND PROSPECTS FOR FAIR CONDITIONS OF WORK

- Mannika Mishra and Naveen Thayyil[*]

I. INTRODUCTION

Algorithms are changing workplaces and practices in myriad and important ways, and the manners and the very capacity of law to keep the hitherto recognised rights of workers intact within algorithmic platforms is seen as a serious challenge. Various digital and algorithmic technologies are being developed rapidly in the past decade which has facilitated the consolidation of networks in such a way that algorithms control an increasing portion of our lives. The deployment of such technologies is seen as resulting in the increasing ubiquity of gig-work platforms that employ significant numbers of 'independent' workers, especially in the wake of recent global economic crises.[1]

The proliferation of such algorithmic platforms that facilitates 'gig work', and its regulation to ensure fair conditions of work is the focus of this paper. Even as the additional challenges in striving for fair conditions in an algorithmic workplace is to be viewed within a general trend of drastic erosion of labour rights in the current century, how to foster and protect fair working conditions when workplaces are infused and transformed through algorithmic application is the concern articulated in this paper. This paper seeks to explicate the labour rights implications of algorithmic infrastructures for rideshare apps operating in India. It should be noted that these apps are addressed within the context of the Indian economy, where, unlike in the European and North American labour context, informal work has long been normalised.[2] This formal, app-based coordination of informal work and workers, therefore, is proceeding in a vastly different workplace environment in India and particularly vulnerable to knowledge gaps which inadvertently end up discriminating against economically and socially marginalised communities.[3]

[*] Mannika Mishra is an independent research scholar based in Delhi. Naveen Thayyil is an Associate Professor (Law and STS) at IIT Delhi.

[1] Irene Padavic, *Laboring under uncertainty: Identity renegotiation among contingent workers*, 28(1) SYMBOLIC INTERACTION (2005); Peter Cappelli, & James R. Keller, *Classifying work in the new economy*, 38(4) ACADEMY OF MANAGEMENT REVIEW 575 (2013); Aditi Surie, *Are Ola and Uber Drivers Entrepreneurs or Exploited Workers?*, 53(24) ECONOMIC & POLITICAL WEEKLY (2018) [hereinafter Surie, 2018].

[2] JAN BREMAN, FOOTLOOSE LABOUR (Cambridge University Press 1996); Mihir S. Sharma, *India's Coalitions for Change: Transforming the world's most dynamic economy*, FRIEDRICH EBERT STIFTUNG REPORT (2016).

[3] Reetika Khera, *Impact of Aadhaar on welfare programmes*, 52 ECONOMIC & POLITICAL WEEKLY 61–70 (2017); Silvia Masiero, *New routes to cashlessness? ICTs, demonetisation, and the Indian informal economy*, DSA CON. (2017); Mara

Both informality in employment – workers without access to social security - and the widespread nature of disorganised sectors[4] make the concerns about the effects of algorithms and the gig economy for fair working conditions particularly important in India, as elaborated later in the paper.

The second section introduces the changes that are ushered into contemporary formal workspaces through digitisation. The third section problematises the specific issue of rideshare apps to provide a broad look at both the functioning and background of these apps. The fourth and fifth sections identify specific labour practices – exploitation, unionisation, and collective bargaining – which are being quashed during algorithmic managerial intervention and suggest new avenues for regulatory measures.

II. DIGITISATION AND THE WORKPLACE

The structure of labour markets has been noted to be undergoing a gradual delocalisation across the world through the 1990s, away from its local histories, city and community-specific entryways to work, as in many ways also away from the sovereign control of the Westphalian nation-state, and instead becoming what Mark Graham and Mohammad Amir Anwar term a 'planetary labour market' where the forces of investment and capital are more starkly opposed to the demands of labour than in the past.[5] Present-day employment is being combined with and governed by real-time data collection at a scale unprecedented by regulators[6], and while regulators have a responsibility to regulate innovation and its effects within fundamental values and rights, regulatory attempts with respect to digitisation are seen as lacking in innovation.[7] As traditional organisational structures are argued by interested sections as being too flabby for the 'free market', the focus is instead shifting to large spreads of geography,[8] and an insular rush to monopolisation.[9]

Hvistendahl, *Inside China's Vast New Experiment in Social Ranking*, WIRED (Dec. 14, 2017); Pawan Singh, *Aadhaar and data privacy: biometric identification and anxieties of recognition in India*, INFOR. COMM. & SOCIETY (2019).

[4] NCEUS, REPORT ON DEFINITIONAL AND STATISTICAL ISSUES RELATING TO THE INFORMAL ECONOMY (2008).

[5] Mark Graham & M. Anwar, *The global gig economy: Towards a planetary labour market?*, FIRST MONDAY (2019) [hereinafter Graham & Anwar].

[6] Gerald Friedman, *Workers without employers: shadow corporations and the rise of the gig economy*, 2(2) REV. OF KEYNESIAN EC. (2014).

[7] YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM (Yale University Press 2006).

[8] Graham & Anwar, *supra* note 5.

[9] Lambèr Royakkers, Jelte Timmer, Linda Kool & Rinie van Est, *Societal and ethical issues of digitization*, 20(2) ETHICS & INFORMATION JOURNAL 127 (2018).

In the European and North American contexts, workers now find themselves in a position that they increasingly consider short-term employment opportunities with a variety of different employers, as opposed to the traditional trend of finding work with one or two employers lasting for years together.[10] This trend, however, has long been the norm in India, where informal economy is commonplace. However, access to temporary jobs almost exclusively governed via online gig-work platforms is a new development. Therefore, this informalization of work conditions also includes the increasing informalization of work relationship even in sectors which were hitherto formalized through the liberalization and structural neo-liberal reforms of the 1990s.[11] Food delivery intermediaries such as UberEats, Postmates, JustEat; websites offering an assortment of freelance work from illustration to translation such as Upwork, Freelancer.com, Fiverr; and rideshare apps such as Lyft, Uber, and in India, OLA, and Ride. Unlike formal workplaces, temporary workers are cut off from hierarchical organisational frameworks within which it is easy to locate one's identity.[12] Collectively, this translates to a nebulous relationship with their professional community which hampers representation.[13] Put differently, workers are in a general milieu where they are forced to sacrifice economic and professional security in the workplace and take on their employers' share of risk for a less constricted and 'more free' personal life.[14]

Introduction of gig-work fits well into this pattern of informalization and systemic avoidance of formal application of fair working conditions in the Indian economy, with a large percentage of workers holding down multiple jobs to earn a living – for instance, it is fairly common for security guards to work multiple shifts in different firms to make ends meets.[15] Digital platforms like Ola and Uber thus have a huge pool of temporary workers that can be recruited from in

[10] Kristine M. Kuhn, *The rise of the "gig economy" and implications for understanding work and workers*, 9(1) INDUSTRIAL & ORGANISATIONAL PSYCHOLOGY 157 (2016) [hereinafter Kuhn]; Lilly Irani, *Difference and dependence among digital workers: The case of Amazon Mechanical Turk*, 1 SOUTH ATLANTIC QUARTERLY 114 (2015) [hereinafter Irani].

[11] Kaushik Basu, *Structural Reform in India, 1991-93*, 48(28) ECONOMIC & POLITICAL WEEKLY (1993); R Nagaraj, *What Has Happened since 1991 - Assessment of India's Economic Reforms*, 32(44-45) ECONOMIC & POLITICAL WEEKLY (1997); Pulapre Balakrishnan, *Markets, Growth and Social Opportunity*, 52(2) ECONOMIC & POLITICAL WEEKLY (2017).

[12] Gianpiero Petriglieri, Susan J. Ashford & Amy Wrzesniewski, *Agony and ecstasy in the gig economy: Cultivating holding environments for precarious and personalized work identities*, 64(1) ADMIN. SCIENCE QUARTERLY 124 (2019).

[13] D. Scott DeRue, & Susan J. Ashford, *Who will lead and who will follow? A social process of leadership identity construction in organizations*, 35(4) ACADEMY OF MANAGEMENT REVIEW (2010); Blake E. Ashforth, Spencer H. Harrison, & Kevin G. Corley, *Identification in organizations: An examination of four fundamental questions*, 34(3) JOURNAL OF MANAGEMENT 325 (2008).

[14] ULRICH BECK, SCOTT LASH & BRIAN WYNNE, RISK SOCIETY: TOWARDS A NEW MODERNITY (1992).

[15] FRANK PASQUALE, THE BLACK BOX SOCIETY (Harvard University Press 2015) [hereinafter PASQUALE]; ROBYN CAPLAN ET AL., ALGORITHMIC ACCOUNTABILITY: A PRIMER (Data and Society 2018) [hereinafter CAPLAN].

India, and algorithmic systems ensure the management of larger and larger numbers of drivers, and wider territories across cities, and countries.[16]

Algorithmic management systems are increasingly used for the recruitment of workers, their monitoring and assignment of work to them, including by assigning conditions of thresholds regarding when work is to be assigned, prioritised or denied. However, meaningful oversight with respect to the creation and monitoring of such algorithmic assignment system is due to a variety of reasons mentioned later in the article, even though attempts at least as early as 2007 (for example the U.S. Federal Agency Data Mining Reporting Act, 2007) exist.[17]

The data that acts as the input to their functioning, the data which is collected as a result of their functioning, and the choices made by the algorithm are not even accessible to public bodies, much less properly regulated.[18] Consider that in rideshare apps especially, driver and passenger ratings determine whether drivers will get more rides or not, and a dip in performance may warrant arbitrary deactivation or suspension of driver accounts.[19] Given that the effect of such algorithmic system is the creation of a universal threshold of performance through ratings that can lead to even the termination of a person's job without even the possibility for the employee to explain herself, or the intervention of a single human manager who might grasp the finer nuances and details about a low rating.

The effects of algorithms in the workplace in such contexts, which are increasingly commonplace, may then at least have the potential to seriously affect existing labour rights for fair working conditions in various ways. In conventional frames, the function undertaken by these algorithmic platforms of providing skilled labour like restaurants, taxi drivers, caterers, plumbers and mechanics would be a labour company. And when such labour companies also own the parent business or vice-versa, it would be common-place for law to assume that there is an employer-employee relationship. What imaginations would allow such algorithmic platforms to be classified as intermediary facilitators and not employers, as it is often been the case, and what work conditions does this definition then lead to?

[16] Min Kyung Lee, Daniel Kusbit, Evan Metsky, & Laura Dabbish, *Working with machines: The impact of algorithmic and data-driven management on human workers*, PROCEEDINGS OF THE 33RD ANNUAL ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS 1603-1612 (2015) [hereinafter Lee].

[17] Nicholas Diakopoulos, *Algorithmic accountability reporting: On the investigation of black boxes*, TOW CENTRE FOR DIGITAL JOURNALISM REPORT (2014) [hereinafter Diakopoulos, 2014].

[18] CAPLAN, *supra* note 15.

[19] Diakopoulos, 2014, *supra* note 17; Nicholas Diakopoulous, *How Uber surge pricing really works*, THE WASHINGTON POST (2015) [hereinafter Diakopoulos, 2015].

III. BLOOD, SWEAT, AND UNRELIABLE INCOME: THE RIDESHARE APP MARKET IN INDIA

The landscape of Indian rideshare app market has brought forth concerns among commentators about fair working conditions and monopolistic consolidation. The market is dominated by two companies, OLA and Uber India which together occupy about 95% of the rideshare market, and OLA had held the upper hand at 56.2% at least until 2018.[20] OLA is an Indian rideshare company which operates in 110 cities and is expanding to overseas markets such as New Zealand and Australia.[21] OLA is owned by ANI Technologies Private Limited; a private unlisted company classified as 'limited by shares'. Major investments have been made by Ratan Tata, Hyundai and Kia Motors, Steadview Capital, Flipkart's co-founder Sachin Bansal, Singaporean Temasek, and an assortment of smaller Chinese funds.[22] A recent round of fundraising raised USD 11 million from the investors Swedish DIG Investment Ab, American Deshe Holdings and two founders of Dubai-based Jabbar Internet Group.[23] OLA Electric Mobility (OEM) has raised about USD 250 million from SoftBank, and in May, Tata Sons Chairman Emeritus Ratan Tata also joined the club, investing an unspecified amount into OEM.[24] OLA Electric Mobility has further secured INR 400 crore in funds invested by Tiger Global and Matrix India.[25] However, all this began with the Japanese Softbank Vision Fund when OLA had 33,000 cars in only 19 cities in India, and which still holds a significant stake of 25% (the founders have a 10% stake) in the company.[26]

Softbank also has a 15% stake in OLA's rival, Uber, which – in a recent IPO pricing - is valued at USD 82.4 billion to OLA's USD 6.2 billion[27], and counts India as one of its most crucial markets in the Asia-Pacific after withdrawing from China following a buyout by Didi Chuxing.[28] So does Ratan Tata via the Tata Opportunities Fund, which backed Uber's India operations with USD 100 million in funding in 2015, and Bennett, Coleman, and Company Limited; better known as the parent company of India's largest media conglomerate – The Times Group.[29]

[20] Sayan Chakraborty, *For Ola and Uber, India's shared taxi market is the next battleground*, LIVEMINT (June 6, 2017).

[21] *India's Ola forays into New Zealand in latest overseas push*, REUTERS (Sept. 18, 2018).

[22] Sandeep Soni, *India's 2nd most valuable startup: Ola valuation to hit $6.2 billion, new funding proposal by Hyundai, Kia Motors shows*, FINANCIAL EXP. (May 2, 2019) [hereinafter Soni].

[23] ET Bureau, *Ola's parent ANI Tech gets $11 m*, ET (Jul. 4, 2019).

[24] Press Trust of India, *SoftBank pumps in $250 million in Ola Electric*, ET (July 2, 2019).

[25] *Id.*

[26] Soni, *supra* note 22.

[27] *Uber's IPO joins ranks of Wall Street flops*, LIVEMINT (May 11, 2019).

[28] TNN, *Subsidised rides will go away as both cos get bigger: Uber's Travis Kalanick*, ET (Dec. 18, 2016).

[29] Douglas Macmillan & Talis Demo, *Uber Valued at More Than $50 Billion*, WALL STREET JOURNAL (July 31, 2015); Mark Scott, *Uber raises $100 million from Tata of India*, NEW YORK TIMES (Aug. 19, 2015).

This overlap of investors is accompanied by fierce competition between the two companies in India where both recognize that their current operations are not profitable, and yet both are burning considerable money seeking to capture the market.[30] For instance, during the height of their stand-off, OLA alone directed USD 300-400 million into subsidies, driver and rider incentives,[31] moves necessary to dominate its sole competitor Uber and the rideshare market.

The average number of rides booked on OLA each day stood at 1.5 million in 2018, when there were reportedly 450,000 vehicles affiliated to the platform (OLA report, 2019). OLA reports that it has seen a rise in both customers and drivers but it not clear from publically available data as to how many of those drivers are actually on the streets as opposed to a mere digital account.[32] There is also a dearth of data on the proportion of drivers who make a good living exclusively from work via OLA versus those who use a variety of other apps to get 'reliable' work. ToGo, Ridely, and Ride are other Indian rideshare ventures offering a 'social riding experience'.[33]

Shared rides, offered by OLA and Uber, make about 30% of all rides in the markets that they are present in, and reports of slashing prices of shared rides in a bid to attract more customers, have hit the worker-driver hard. OLA has been known to cut prices – cutting base fares, minimum fare per mile, etc. - and driver allowances without warning,[34] while Uber had capped its pool rides at INR 49 for the first 8 kilometres in three large Indian cities – Bangalore, Chennai, and Delhi, a figure unsustainable for any meaningful wage for the driver.

In the beginning of 2017, the incomes of rideshare app drivers fell drastically by anywhere between 30% to 45%,[35] leading to a series of strikes by OLA and Uber drivers, the casualties of the escalating competition between the two platforms for the larger chunk of market share. These protests led to a reduction of customer bookings (by 5%) and driver enrolment (by 10%), resulting in longer travel times and slower service culminating in a vicious cycle. For instance,

[30] Press Trust of India, *Uber to stay invested in India; looks to grow 10x: CEO*, ET (Feb. 22, 2018); Press Trust of India, *Ola losses narrow to Rs 2,842 crore for FY18; revenue zooms 61%*, ET (Jan. 31, 2019).

[31] Ananya Bhattacharya, *As Uber sputters, Ola is really stepping on the gas in India*, QUARTZ IND. (Feb. 15, 2018).

[32] Shashwati Shankar, *Undeterred by high attrition rate, Ola and Uber banking on drivers in their 20s*, ET, June 01, 2017; Shashwati Shankar & Madhav Chanchani, *Ola's revenue surges seven-fold, but loss widens to Rs 2,313.66 crore in FY16*, ET, May 1, 2017.

[33] Sanchit Khera, *The demand for ride-sharing apps in India and why more people are using them*, THE NEWS MIN., July 9, 2018.

[34] Surie, 2018, *supra* note 1; Ira Anjali Anwar, *Ola, Uber and the Precarious Future of Blue Collar Platform Workers*, THE WIRE, Mar. 26, 2018; Digbijay Mishra, *Ola, Uber cut driver pay by a third in 1 year,* ET (Sept. 18, 2017).

[35] RedSeer Consulting, *State of the Online Cabs Market*, Q1 (2017).

while booking services remained satisfactory for consumers, the ride experiences themselves deteriorated. As regards regulation, RedSeer found no evidence of regulatory effects enacted by the government, indicating that the existing regulations proved to be ineffective in curtailing the more exploitative labour practices adopted by these companies. For instance, the steady (and in some cases drastic) decline of incentives offered to drivers from 2017, initially in more 'mature markets' like New Delhi and Bangalore, and subsequently in other cities led to a sharp decline in the real wages of the drivers. In contradistinction to these concerns of fair working conditions, the RedSeer industry report viewed the drop of "incentives" not as a cause for concern, but merely a symptom of a larger process of industry stabilisation after a boom. Indeed, in the intervening months the revenues of both OLA and its main rival, Uber India, saw a steady uptick. Regulatory reports from the fiscal years of 2017 and 2018 show that OLA continues to grow, most promisingly for investors since it has halved its losses to Rs 2,676.7 crore, and its revenue also rose 44.6% to INR 1,860.6 crore in the same period.[36]

The global nature of these companies further complicates meaningful regulation of these algorithmic practices.[37] European countries such as Germany have attempted to penalise tech platforms for violating labour laws. Apart from the challenges about the possible reach of such regulatory steps in algorithmic shaping of labour practices, these regulations have rarely had a meaningful effect in regulating operations in countries where labour laws do not have deep social roots and the worker groups are more vulnerable and liable to exploitation.[38] The labour regulation in India has been silent on the issue of wages about the driver workers, implicitly accepting the platform logic that the drivers are not workers, and the company is only a digital platform for businesses (i.e. the individual drivers) to find customers. Even as the experience of most drivers working in the platform is one where work frequently starts off as relatively lucrative for the workers, wages subsequently drop as the supply of workers increase and further fluctuate according to the opaque parameters set by the algorithm. This usually means that workers have to progressively work for cheaper and cheaper wages and therefore for longer hours each day to make the same amount of money, and in some cases, this barely amounts to minimum wage. In what way can law recognise the reality of this experience of the driver, and regulate fair working conditions?

[36] PTI, 2019, *supra* note 24.

[37] Andrew Stewart & Jim Stanford, *Regulating work in the gig economy: What are the options*?, 28(3) ECO. & LAB. RELATIONS REV. 382 (2017).

[38] Graham & Anwar, *supra* note 5; Mark Graham, Isis Hjorth & Vili Lehdonvirta, *Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods*, 23(2) EURO. REV. OF LAB. & RESEARCH 135 (2017) [hereinafter Graham et. al, 2017].

IV. DISENTITLEMENTS OF WORKER RIGHTS IN ALGORITHMIC PLATFORMS

It is increasingly clear that the flexibility allowed by the algorithms in the apps, which is one of the attractions of temporary work, has become detrimental to the rights of the workers.[39] Further, as is elaborated in this section, law has been used in manners detrimental to worker rights. The flexibility through apps comes in part from the non-limiting nature of the contracts supplied by these platforms which are explicit about not specifying the minimum, binding rights that must be afforded to workers. Fluctuation around the constantly changing conditions of demand and supply, then guides the way in which workers are paid – and valued – at any given hour to produce the most efficient and cost-effective relationship between the worker, the app, and the consumer.

A significant part of algorithmic structuring of working conditions though is structured through the designation of the work to be assigned to workers. Instead of employees, the taxi workers are classified as independent contractors and entrepreneurs to whom no legal job protections need apply.[40] Put differently, the platforms are not liable for their welfare as they are contracted for short-term work – albeit on a fairly regular basis – on their own reconnaissance.[41] This has been accomplished by re-branding the label of 'entrepreneur' as something not only to aspire to, but as something which is easily accomplished - an independent worker who is their own boss. The fact that this comes with the loss of umbrella protections is conveniently ignored in the adverts and testimonials plastered around cities, speaking of workers who earn well and hold working hours which are convenient to their own needs. The rise of this gig economy model of designating employees as business partners is not limited to digital platforms but has steadily pushed into mainstream hiring practices; increasingly, work is outsourced and freelanced piecemeal in 'traditional workplaces.[42] It is important to recognise the detrimental effects of employing this label of entrepreneur and evading the employer-employee definition (cogently argued in Greg Marston and Peter Holtum in this special issue).

Further, exploitation occurs in other insidious ways on digital platforms, and consequently in areas which make use of their services. With the rise of automation, the expectation of labour availability has climbed ever higher so that response times for workers have shrunk dramatically,

[39] Kuhn, *supra* note 10; Irani, *supra* note 10; Diakopoulos, 2015, *supra* note 19.

[40] Graham & Anwar, *supra* note 5.

[41] Kuhn, *supra* note 10.

[42] Alek Feltstiner, *Working the crowd: employment and labor law in the crowdsourcing industry*, 32 BERKELEY J. EMP. & L. REV. 143 (2011).

especially for on-demand services such as getting a taxi or food delivery.[43] When hailing a taxi, one expects it to arrive instantaneously. There is a crucial distinction, however, as expectations of mechanised systems cannot be held to human workers and doing so puts vast stresses on the workers' abilities and on current expectations of protections.[44] A rapidly progressing environment of 'consumer-centric' work has therefore been recognised as eroding the working conditions of workers.[45]

In essence, workers are both governed and expected to perform up to the standards of codes and algorithms attuned to the hermetic conditions of the market of the platform rather than to wider considerations of benefits and rights.[46] Exploitation, therefore, occurs in a *hidden* way on digital platforms. It is facilitated by appraisal practices such as Key Performance Indicators, which can be measured in unprecedented detail now via monitoring processes built into workers' apps. The way this information and labour is used is almost always opaque, as labour data is used to further enrich the algorithm that runs these platforms.[47] Through what ways can and should the law respond to stem the disentitlement of worker rights in the current turn to algorithmic work platforms, and further find ways to protect rights through algorithms, become crucial questions in the contemporary workplace.

V. TOWARDS ALGORITHMIC ACCOUNTABILITY

Unionisation is a well-established and hard-won practice to protect employees from the arbitrary whims of their employers, ensuring that certain important working conditions are met with reliable certainty so that employees are not left in the lurch.[48] If one's working conditions are determined *fait accompli* according to a broader pattern of data – worker demand at a specific time, pricing conventions which change in accordance to real-time information – to which workers do not have access or recourse, then the lawfulness and fairness of their own working conditions is hidden from them.[49] Many companies – *e.g.* Deliveroo - also provide clauses in their contracts which permit termination in case of unionisation attempts. In some cases, existing union

[43] Lauren Weber & Rachel Emma Silverman, *On-demand workers: "We are not robots"*, WSJ 27 (2015).
[44] *Id.*
[45] Surie, 2018, *supra* note 1.
[46] Graham et. al, 2017, *supra* note 38.
[47] Irani, *supra* note 10.
[48] TOWARDS A FAIRER GIG ECONOMY (Mark Graham and Joe Shaw eds., Meatspace Press 2017) [hereinafter Graham & Shaw]; Hannah Johnston, and Chris Land-Kazlauskas, *Organizing on-demand: Representation, voice, and collective bargaining in the gig economy*, CONDITIONS OF WORK AND EMPLOYMENT SERIES 94 (2018).
[49] Graham & Shaw, *supra* note 48; Surie, 2018, *supra* note 1.

organisations have supported platform-worker groups in efforts to force policy changes in tech company practices and government regulation of worker's rights.

Unionisation is often a key to a bigger bargaining chip for workers; in this case, preventing unionisation is only one hurdle to achieving means to engender collective bargaining. Platform's reluctance to let workers unionise makes for an uncertain situation for those whose entire means of livelihood depend on the work provided via the app, and who do not want to antagonise it and risk having their accounts deleted. The erosion of labour rights is of course not anything new to be attributed singularly to the ubiquity of algorithmic structuring of the workplace. And yet there are newer ways of disentitlement in and through algorithmic platforms that require special attention through the law, even as there are also ways in which algorithms can be employed to ameliorate such disentitlement.

There is an urgent need for establishing a new culture of scrutiny around algorithms: setting clear standards for labour and rights protection, the transparent collection of data and information, and the creation and implementation of regulatory measures which will largely stem from the impetus of the first two tasks.[50] The ways in which algorithms operate to control the experience and the behaviour of drivers – taking on the role of human managers – is often made difficult to access and parse. Private companies like Ola and Uber do not release the full breadth of the data they collect, since they are considered proprietorial and crucial to inform and run their algorithms. Put simply, an algorithm is a series of instructions to perform a particular task. These instructions are programmed and optimised by human beings, or sometimes even by other algorithms such as machine-learning algorithms, using data about human behaviour and trends relevant to the task. In other words, most of the information which goes into an algorithm is collected and selected by human programmers, and this can preserve and even enhance human biases in algorithms. Because of their opaque nature, they are not frequently subject to wide scrutiny as to its ethics and the tool of exacting accountability, regulation, can therefore not be applied with any great effect.[51]

Algorithms are primed to transpose the already precarious working conditions of the Indian worker further entrenching a temporary and unprotected work environment on digital platforms. By chipping away and slowly dissembling the traditional taxi system in India, rideshare companies

[50] PASQUALE, *supra* note 15; Caplan, *supra* note 15; Diakopoulos, 2014, *supra* note 19.

[51] PASQUALE, *supra* note 15; Danielle Keats Citron & Frank Pasquale, *The scored society: Due process for automated predictions*, 89(1) WASH. L. REV. 1 (2014) [hereinafter Citron & Pasquale].

like Ola and Uber are slowly creating a captive base not only of customers but also drivers. There are rules and criteria embedded in algorithms, and despite their seemingly reliable black and white nature they need constant tweaking to keep pace with the nuance and complexities of societal needs.[52] In short, algorithms are capable of making decisions and running autonomously but they are influenced by human operators and reflect the values and culture within which they operate.

Nick Diakopoulos (2014) has divided the work of algorithms in four broad categories: prioritising, classification, association, and filtering. Different combinations of these activities produce a highly curated experience which accounts for all things from the wait times for a taxi in a specific area to the determination of one's salary according to performance.[53] For instance, algorithmic implementation of pricing manifests itself in 'surge-pricing', which seeks to bring down wait times by re-allocating driver resources to areas where the algorithm sets higher prices for each ride.[54] Diakopoulos (2015) found that the Uber algorithm reacts 'temperamentally' to the data with the result that pricing changes every three to five minutes, too quickly to be of any use to drivers or consumers.

Transparency here is pushed forward as a tool to enact accountability, but this conception of accountability is susceptible to being rendered toothless. Even transparency and its offshoots must be regulated to form what Frank Pasquale terms 'qualified transparency':[55] a secretive system which makes sure that no information may be got out, but considering the amount of private data collected even with one's consent, complete transparency would not be desirable to even the most ardent enthusiast of full disclosure.[56] Transparency, and the sort of journalistic transparency that Diakopoulous (2015) espouses in particular, is an important tool for *collecting* knowledge and making algorithms less specialised and elite, but a meaningful expectation of accountability might only result from regulation. This kind of regulation would audit the information being processed, collected, and produced by algorithms, thereby making sure that important civil protections remain intact.

Most promisingly, the Association of Computing Machinery[57] makes the social impact of computing a crucial issue to the future of programming. The ACM is the pre-eminent industry

[52] CAPLAN, *supra* note 15; Lee, *supra* note 16; Diakopoulos, 2014, *supra* note 17.
[53] CAPLAN, *supra* note 15; Citron & Pasquale, *supra* note 51.
[54] Diakopoulos, 2015, *supra* note 19.
[55] PASQUALE, *supra* note 15.
[56] *Id.*
[57] ACM, SOFTWARE ENGINEERING CODE OF ETHICS AND PRACTICE (2015).

society for computing professionals which provides a forum for professionals, academics and students to conduct a dialogue about pressing issues in computer science. The ACM's code of ethics seeks to identify and provide broad guidelines for ethical programming and prioritises the 'do no harm' principle – however, these abstract guidelines do very little to provide actionable regulatory thresholds which ensure accountability. In this scenario, a direct line can be drawn between the precariousness of temporary work – in this case, the Indian rideshare gig-work – and the algorithms which govern it. Where is the ethical benchmark for an app which contravenes the notions of traditional protections of employment? The answer may lie with incentivising pro-labour algorithms and setting achievable benchmarks for apps to develop ethical and responsible worker-friendly algorithms. But then would the law discover susceptible political pressure to be able to incentivise the development and deployment of such pro-labour algorithms?

FROM THE RHETORICAL SOFTWARE TO THE 'HARDWARE OF THE LAW': REGULATING HATE SPEECH ONLINE IN INDIA

- *Siddharth Narrain*[*]

I. INTRODUCTION

Traditional debates around the regulation of hate speech have focused on the 'what' question – what should be regulated and shouldn't, and the limits of free speech in relation to hate speech. In India, much of this question has revolved around whether 'reasonable restrictions' in Article 19(2), that have circumscribed the right to freedom of speech and expression in Article 19(1)(a), can be applied to limit specific instances of hate speech, and aspects of hate speech that could be legitimately prohibited by the law. In this essay, I argue that legal scholarship around hate speech needs to factor in not just the question of 'what' but also the 'how' – how is hate speech online being regulated, and what are the specific sites where such legal regulation takes place? In order to do this, I argue that we need to refocus our attention from the rhetoric, to what the scholar Cornelia Vismann terms the "hardware of the law"[1] – the infrastructures, bureaucracies and mechanisms of legal regulation that are underpinned by specific technological affordances,[2] devices and networked publics.[3]

I begin this paper by examining whether it makes sense to distinguish between the conceptual categories of 'online' and 'offline' hate speech. I discuss this question in relation to legal debates in India around the regulation of hate speech online. I then examine theoretical debates around 'speech as thought' as opposed to 'speech as action', and the implication of these debates the legal regulation of hate speech online. I then move on to discuss the specific question of the infrastructures of hate speech, which I argue opens up an important dimension of the legal

[*] PhD Candidate, Faculty of Law, University of New South Wales (UNSW), Sydney.

[1] Cornelia Vismann, *Jurisprudence: A Transfer Science*, 10 LAW AND CRITIQUE 279, 284 (1999) [hereinafter Vismann].

[2] The term affordance was first proposed by the psychologist James Gibson to refer to that which an environment offers or furnishes an animal. JAMES GIBSON, THE ECOLOGICAL APPROACH TO VISUAL PERCEPTION (Routledge 2014). Technological affordances refer to action possibilities and opportunities that emerge from actors engaging with focal technology and is located within a relational ontology that gives equal importance to the material and the social. Samer Faraj & Bijan Azad, *The Materiality of Technology: An Affordance Perspective, in* MATERIALITY AND ORGANIZING: SOCIAL INTERACTION IN A TECHNOLOGICAL WORLD (Paul M. Leonardi, Bonnie A. Mardi & Jannis Kallinikos eds., Oxford Scholarship Online 2012).

[3] Networked publics are publics restructured by networked technologies. The scholar Dana Boyd states that the term refers to both a) the space constructed through networked technologies, b) the imagined collective that emerges as a result of the intersection of people, technology and practice. Danah Boyd, *Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications*, *in* NETWORKED SELF: IDENTITY, COMMUNITY, AND CULTURE ON SOCIAL NETWORK SITES 39-58 (Zizi Papacharissi ed., 2010).

regulation of hate speech online. I end the essay flagging the question of the appropriate research methodology for legal researchers, if the site of research is the infrastructure of hate speech, and not just its content.

II. WHAT'S DIFFERENT ABOUT HATE SPEECH ONLINE?

With the emergence of hate speech on social media platforms, and regulatory responses such as the Information Technology Act 2000 enacted in India, one of the key questions that has emerged is whether the standards that were used to regulate hate speech offline, should be applied online. This question was an important part of the arguments in the litigation around section 66A of the Information Technology Act, which was challenged by a number of individuals and organisations before the Indian Supreme Court in the *Shreya Singhal* case, after this particular section was repeatedly used by the central and state governments to target legitimate criticism of state action. In March 2015, a two-judge bench of the Indian Supreme Court, case struck down section 66A of the Information Technology Act, as violative of Article 19(1)(a)- the freedom of speech and expression.[4] Striking down section 66A as vague and arbitrary, the court distinguished between 'discussion', 'advocacy', and 'incitement', holding that only speech amounting to 'incitement' could be prohibited by the law.

While striking down this law, the Court observed that it disagreed with the argument put forward by the petitioners that the law violated Article 14 of the Indian Constitution, (the right to equality) as laws governing the regulation of the Internet should be on par with other media. The Court stated that many of the sections in the Information Technology Act were framed with the view that separate standards were needed for the regulation of the Internet.[5] This view is in line with international legal standards, as per which laws regulating media should take into account the specifics of these forms of media, while at the same time ensuring that the rights that individuals enjoy offline, are also protected online, and that limitations on these rights are justified with the same criteria.[6]

The *Shreya Singhal* judgment has been lauded for its commitment to free speech and striking down a law that was being used against dissenters, and artists, weaponised by the government

[4] Shreya Singhal v. Union of India, AIR 2015 SC 1523.

[5] Siddharth Narrain, *The Death of Section 66A and the Dawn of a New Era of Free Speech Jurisprudence*, KAFILA.ORG (Mar. 23, 2015), https://kafila.online/2015/03/24/the-death-of-66a-and-the-dawn-of-a-new-era-of-free-speech-jurisprudence-siddharth-narrain/#comments (last visited June 8, 2020).

[6] ARTICLE 19, SELF REGULATION AND 'HATE SPEECH' ON SOCIAL MEDIA PLATFORMS 8-9 (London 2018).

against those it was supposed to enable speech by. For some time after the striking down of the law, there has been talk of a gap in the law. This gap in the law is mainly with respect to hate speech online – when communities or individuals are targeted online in the name of religion or other identities, what protection do people have. As of now those targeted by hate speech online in India have to resort to using provisions from the Indian Penal Code such as sections 153A and 295A, which were never designed to address hate speech on online platforms, or to the community guidelines and internal policies of such platforms.

The Law Commission of India, responding to a 2014 Supreme Court judgment *Pravasi Bhalai Sangathan*,[7] a case dealing with hate speech against inter-state migrants, where the Supreme Court requested the body to examine the issue of the legal regulation of hate speech, has made a series of recommendations that seek to address this gap. In its 267th Report submitted in 2017,[8] the Commission has noted that hate speech has not been defined under Indian law, and the practices that we understand as broadly comprising hate speech are regulated by a number of laws that range from electoral laws under the Representation of People's Act, to existing Indian penal Code provisions, media-specific laws such as the Information Technology Act, and laws such as the SC/ST Prevention of Atrocities Act that regulate specific class of slurs and speech based on historical discrimination.[9]

The Commission in its report, specifically talks about the emergence of online hate speech in India as being a regulatory challenge. The report specifically mentions the 2012 exodus of thousands of persons from the Northeast from cities such as Bangalore and Pune after threats against them were circulated widely on social media.[10] While the Commission has not recommended enacting a specific law to deal separately with online hate speech, it has recommended that the existing sections 295 and 505 of the Indian Penal Code be amended to include sub clauses that will cover hate speech defined more in line with comparative and international law that uses the language of 'incitement to hatred and the incitement to discrimination standard.' The Commission endorsed the 'incitement' standard laid down in the Shreya Singhal case, and the distinction made between advocacy, discussion and incitement.[11]

[7] Pravasi Bhalai Sangathan v. Union of India, AIR 2014 SC 1591.

[8] LAW COMMISSION OF INDIA, REPORT NO. 267, *Hate Speech*, March 2017.

[9] *Id.*

[10] *Id.*

[11] *Id.* For an overview of the Law Commission's recommendations, see Siddharth Narrain, *Unpacking the Law Commission's Hate Speech Report*, SOCIO-LEGAL REVIEW FORUM, NATIONAL LAW SCHOOL OF INDIA UNIVERSITY (2017).

While the Law Commission report does not discuss this in detail, and does not address this in terms of recommendations, the discussion in its report seems to suggest that there is something specific about online hate speech that is different from hate speech on older forms of media. While it is true that the form of hate speech on platforms such as WhatsApp, and Facebook does not differ much from what ones sees on other media, and that there is a constant flow of information across online and offline media, there is something specific about online hate speech that legal scholars have begun to single out. For instance, the legal scholar Wolfgang Schulz, in his work on the German Networks Enforcement Act, 2018 (NetzDG), that punishes platforms for allowing for violations of the German law, identifies certain characteristics of online hate speech that make it different. These characteristics include the sheer magnitude of content online. For example, Facebook reported to have 100,000 content related decisions per month that were being reviewed under the NetzDG that penalizes online speech deemed illegal under domestic German law.[12]

Another aspect that Schulz identifies is acceleration, or the speed with which content is circulated online. Content has the potential to go viral and the response time for state authorities to address any potential impact of such speech has reduced dramatically.[13] Virality is a specific mode of transmission linked to what Manuel Castells termed 'information capitalism',[14] based on the logic of contagion and repetition,[15] and involving key nodes through which speech is transmitted.

These distinct features make hate speech online especially dangerous, and this is clear from the measures that platforms have begun to take to tackle this problem. The controversy over the way in which Facebook was used to target persons from the Rohingya community in Myanmar and the role that this played in the violence against them, prompted Facebook to take a series of measures to deal with the problem, and has been raised in a number of government enquiries against the company.[16] One way in which companies have responded is to use automated search

[12] Wolfgang Schultz, *Regulating Intermediaries to Protect Privacy Online: The Case of the German NetzDG*, at 3, DISCUSSION PAPER SERIES 2018-01, HUMBOLDT INSTITUT FUR INTERNET UNDE GESELLSCHAFT (HIIG).

[13] *Id.*

[14] MANUEL CASTELLS, THE RISE OF THE NETWORKED SOCIETY (Blackwell Publishers 2000).

[15] JUSSI PARIKKA, DIGITAL CONTAGIONS: A MEDIA ARCHAEOLOGY OF COMPUTER VIRUSES (Peter Lang 2007).

[16] A. Stevenson, *Facebook admits it was used to incite violence in Myanmar*, NEW YORK TIMES (Nov. 6, 2018). For an in-depth analysis of the factors that contributed to the attacks on the Rohingya community in Myanmar including the role of the state-owned media see Ronan Lee, *Extreme Speech in Myanmar*, 13 INTERNATIONAL JOURNAL OF COMMUNICATION 3203-3224 (2019). For a historical survey of digital media in Myanmar see Daniel Arnaudo, *Bridging the Deepest Digital Divides: A History and Survey of Digital Media in Myanmar*, *in* MAPPING GLOBAL DIGITAL CULTURES 96-121 (Aswin Punathambekar & Sriram Mohan eds., University of Michigan Press 2019).

functions along with alerts from users to weed out what they define as hate speech, as their reliance on user's flagging such content decreases.[17]

However even with such sophisticated technology, it is difficult to regulate hate speech, as compared to child pornography or copyright violations, for instance, given how subjective and contextual the determination of online hate speech can be.[18] The problem here is also related to the lack of response time and the massive volume involved, both factors that Schulz points out are specific to hate speech online.

The situation is even more complicated in countries like India where there are many regional languages and moderators who do not understand the nuances of these languages, and in the specific context will not understand why certain kinds of speech have a charge or effect in a local context. Further, some speeches are not direct attacks but coded messages that underlie statements referring to persons of certain out group. Recent examples of these include Home Minister Amit Shah's insidious references to Muslim refugees and immigrants as 'termites' during the 2019 General Election.[19] These kinds of statements are what are referred to as dog whistles or insidious references to groups – giving them an identity that is subhuman, or dehumanising them, and indicating that it is fine to discriminate and perpetrate violence against these groups, since they are not human.

In the next section I discuss the divide between those who view speech as thought and those who view speech as action, and the implications of these views for the legal regulation of hate speech online.

III. THE PERFORMATIVE ASPECTS OF HATE SPEECH ONLINE

When one looks at the scholarship around free speech jurisprudence more generally, they can be divided broadly into two camps – those who view speech as more in the nature of thought, and therefore less likely to support its legal regulation, and others who view speech as in the nature of

[17] Rachel Metz, *Facebook is Doubling Down on AI to Clean Up the Social Network*, CNN Business, https://edition.cnn.com/2019/05/01/tech/facebook-ai-f8/index.html (last visited Dec. 20, 2019).

[18] For a more detailed account of the mechanics of platform regulation see Kate Klonick, *The New Governors: The People. Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598-1670 (2018).

[19] Devjyot Ghoshal, *Amit Shah vows to Throw Illegal Immigrants into Bay of Bengal*, REUTERS (Apr. 13, 2019), https://www.reuters.com/article/india-election-speech/amit-shah-vows-to-throw-illegal-immigrants-into-bay-of-bengal-idUSKCN1RO1YD (last visited Dec. 20, 2019).

action, and therefore more likely to support its legal regulation.[20] For those who view speech as in the nature of an action, their focus is on what speech *does*, for them, just as in the case of speech on other forms of media, online hate speech is agent-driven and the materiality of online hate speech matters.[21] One of the most important scholars who influenced the view that speech is not just thought but *does* things, is the philosopher J.L. Austin. Through his series of lectures titled "How to do Things with Words", Austin helps us move away from the conception that speech is completely distinguishable from actions.[22]

Austin argued that some forms of speech function like actions and he called these performative speech acts, as distinguished from constative speech-acts which were more in the nature of statements of description. A performative speech-act in contrast derives its meaning from the force of its utterance, where to utter the statement is not just to describe doing it but to actually do it. Austin's arguments were mobilised by critical legal scholars such as Mari Matsuda and Richard Delgado in the United States, who argued that hate speech cannot be seen as purely speech, and that the impact of hate speech is disproportionate on persons from marginalised communities. They argued for the regulation of hate speech, and for a rethinking of the wide scope of First Amendment protections to many forms of hate speech.[23]

The German legal scholar and media theorist Cornelia Vismann took Austin's arguments one step further. Vismann argues that one has to move away from a purely phenomenological account of speech to the infrastructural conditions through which it operates.[24] In the context of hate speech online this would mean moving away from the focus on the purely rhetorical – whether content amounts to hate speech, to an account of the material infrastructures and bureaucracies of speech and regulation that mediate such speech. This infrastructure ranges from governmental departments that oversee the regulation of the speech online to institutions such as the police and the executive who administer hate speech regulations at the local level, and bureaucracies and administrative structures set up within platforms that are meant to detect and regulate hate speech. In the Indian context, this could range from examining the internal working

[20] DIGITAL DUALISM AND THE "SPEECH AS THOUGHT" PARADOX (Katherine Gelber & Susan J. Brison eds., Oxford Scholarship Online 2019).
[21] *Id.* at 13.
[22] J. L. AUSTIN, HOW TO DO THINGS WITH WORDS (Harvard University Press 1975).
[23] Mari J. Matsuda, Charles R. Lawrence, Richard Delgado & Kimberly Williams Crenshaw, *Words that Wound: Critical Race Theory, Assaultative Speech and the First Amendment* (Westview, 1993).
[24] Vismann, *supra* note 1 at 286.

of Facebook's community guideline mechanisms, to the changes in policy and design that WhatsApp implements to prevent hate speech from going viral.

Vismann's arguments have parallels in the work of the legal scholar Lawrence Lessig, who in his influential work Code 2.0, points to the infrastructures or architectures of the Internet as performing the role of structuring and constraining social and legal powers.[25] Lessig argues that code or instructions embedded in the software and hardware of cyberspace is seen as 'doing something'. He argues that code embeds certain values that perform the function of regulation, and that the expression of code is done in relation to its operating system. Vismann and Lessig are pointing to the crucial role of the infrastructure of speech – platforms, algorithms they employ, and the materiality of the digital ecosystem, in any debate around the legal regulation of speech online, including hate speech online. The scholar Jean-Christophe Plantin, while writing about algorithms, which is one such crucial infrastructure of speech, in the contemporary context, argues that algorithms form both the basis of how platforms track hate speech, but also more broadly form the basis by which they organize and create publics.[26] If one thinks of algorithms as technologies of representation that help codify the publics they are supposed to measure,[27] then these become an important aspect in the way algorithms 'do things' or have performative aspects.

Thus, if one were to think about How to do Things with Words, a 2.0 version, the way online speech does things is mediated through what Lessig calls code, through infrastructures such as algorithms, and through what Vismann would describe as the infrastructural conditions through which speech operates. In the next section, I discuss some of the relevant literature in this area, that will push the discussion from the performative aspects of online hate speech to its materiality, with a focus on the infrastructures of online hate speech regulation.

IV. THE INFRASTRUCTURES OF HATE SPEECH ONLINE

Platforms are what the scholar Julie Cohen terms 'core organizational forms of the emerging informational economy'[28]. Cohen argues that platforms serve as sites that are materially and algorithmically mediated in way the market served as a means of barter and exchange in the

[25] LAWRENCE LESSIG, CODE 2.0 (Basic Books 2006).

[26] Tarleton Gillespie, *The Relevance of Algorithms*, *in* MEDIA TECHNOLOGIES: ESSAYS ON COMMUNICATION, MATERIALITY, AND SOCIETY 167-193 (Kirsten A. Foot, Pablo J. Boczkowski & Tarleton Gillespie eds., Cambridge, MA, MIT Press 2014).

[27] *Id.* at 189.

[28] Julie Cohen, *Law for the Platform Economy*, 51 UC DAVIS L. REV. 133 (2017) [hereinafter Cohen].

industrial era economy.[29] She states that from the perspective of users, and advertisers, dominant platforms such as Facebook are beginning to resemble utilities (or infrastructures), while at the same time constituting vast and highly complex information ecosystems.[30] For Cohen, platforms have characteristics of both networks and infrastructures. She argues that platforms "represent infrastructure-based strategies for introducing friction into networks".[31]

In India it is these dominant platforms such as Twitter, Facebook, WhatsApp, YouTube, and Tik Tok, that have been at the centre of debates around hate speech online. In taking into account legal regulatory responses to hate speech on these platforms one has to recognise that these platforms perform multiple roles. These include a) hosting i.e. making content available to the public and storing third party content, where the platform performs the role traditionally associated with an infrastructure b) online distribution i.e. making content visible and easily accessible, which is done through algorithms, and c) performing functions associated with traditional media outlets such as publishing news and acting as a forum for public discussion.[32]

The references above to platforms performing the function of an infrastructure echoes the work of scholars such as Jean-Christophe Plantin who argue that platforms such as Facebook and WhatsApp are increasingly performing a dual function – that of a platform and that of an infrastructure. Plantin states that there are two main approaches to the study of platforms – 'platform studies' and 'infrastructure studies'. The former is centred around media studies, while the latter draws on disciplines such as Science and Technology Studies (STS), and information science. Plantin argues for the use of a combined framework.[33]

These frames through which emerging digital ecosystems have been characterised – as platforms, as networks and as infrastructures, will in turn determine the method of legal regulation. If Facebook is seen as being predominantly an infrastructure, that would make it crucial for it to be treated on a level playing field, a demand that was highlighted during the debate over Facebook's Free Basics offer that was blocked by the Indian Telecom Regulatory Authority (TRAI).[34] If one

[29] *Id.* at 136.
[30] *Id.* at 148.
[31] *Id.* at 143.
[32] ARTICLE 19, SELF REGULATION AND 'HATE SPEECH' ON SOCIAL MEDIA PLATFORMS 13-14 (London 2018).
[33] Jean-Christophe Plantin, Carl Lagoze, Paul N Edwards & Christian Sandwig, *Infrastructure Studies meet Platform Studies in the Age of Google and Facebook*, 20(1) NEW MEDIA & SOCIETY 293 (2018).
[34] Rahul Bhatia, *The Inside Story of Facebook's Biggest Setback*, THE GUARDIAN (May 12, 2016), https://www.theguardian.com/technology/2016/may/12/facebook-free-basics-india-zuckerberg (last visited March 30, 2020).

views WhatsApp to be a network which functions through a structure of hubs and nodes, then these sites become important for legal and policy intervention when online hate speech circulates rapidly during and preceding tense inter-group violence and clashes. This perspective also foregrounded the role of virality of hate speech, and the crowd that mediates its transmission, both themes that I will not touch upon here, but are central to understanding the legal regulation of online hate speech in the Indian context.

I will come back to the argument in this essay, which is the shift from the 'what' to the 'how' question in relation to the legal regulation of online hate speech in India. Cohen makes an important intervention by reshaping the question that has dominated in this field – she argues that the question that we should be asking "is not how law should apply to disputes over information but how disputes over information are reshaping the enterprise of law at the institutional level."[35] She states that law is already being written through "the uncoordinated but self-interested efforts of information economy participants and the lawyers and lobbyists they employ."[36] This is the site that she identifies as being in need of our attention.

In a similar vein, the scholar Kate Klonick in her work has focused on the 'what question'. Using detailed interviews with ex-employees of Facebook, YouTube and Twitter, and publicly available statements and archival records, Klonick writes about the nitty gritty of how these platforms moderate content.[37] While Klonick describes content moderation within platforms as a process of governance rather than legislation, she does emphasize the similarities between the decisions that content moderators take on these platforms to take down or keep content as very similar to the judicial process. She compares decisions taken by content moderators to those by judges as exercising professional judgment, and applying legal concepts such as relevancy, reason through example and analogy and applying multifactor test, to apply their standards.[38]

Klonick and Cohen are amongst two of the scholars cited in a recent piece by Richard Ashby Wilson, a legal anthropologist, who has mapped the direction that scholarship in the field of legal anthropology should take, to respond to the question of hate speech online. [39] Central to

[35] Cohen, *supra* note 28, at 136.

[36] *Id.*

[37] Kate Klonick, *The New Governors: The People. Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598-1670 (2018).

[38] *Id.* at 1647.

[39] Richard Ashby Wilson, *The Digital Ethnography of Law: Studying Online Hate Speech Online and Offline*, 3(1) JOURNAL OF LEGAL ANTHROPOLOGY 1-20 (2019).

Wilson's argument is the strong link between the conceptual questions around the legal regulation of hate speech online and the question of method. Wilson argues that the evolving normative framework of social media regulation has become 'law' in terms of its function and meaning – it sets standards of what is acceptable and what is not, what is legitimate speech, and thus it should be a site for the focus of legal anthropology.[40]

Based on Erving Goffman's call to study the 'backstage' of social life, Wilson outlines three levels of the digital ethnography of law that need attention – the backstage or the infrastructure of legal regulation; the onstage or social media discourse and the offstage, or the offline effects of online speech.[41] This is a very useful classification for those interested in exploring the legal regulation of online hate speech. For instance, Wilson identifies the policy and technical infrastructure of platforms as sites that need to be studied. He argues that the infrastructure of these platforms shapes how much agency actors on these platforms have, and highlights the role of algorithms, which influence social behaviour.[42]

Wilson, based in this schema, identifies some of the questions that digital ethnographers of the law should be asking. In the context of this essay, these questions include a) what are specific algorithmic programs that allow platforms to moderate content on a massive scale b) who sets the criteria for algorithmic filtering, and what are the economic imperatives which shape these decisions and c) what is the relationship between the content moderation policies that social media companies use and existing laws, administrative and executive powers that regulate hate speech online?

The approach that Wilson identifies here, could be useful to those interested in examining the algorithms in relation to the law, and will resonate with those reading this journal's issue, which is dedicated to this theme. Perhaps, one could also extend this argument beyond anthropological approaches to the law to other forms of qualitative, quantitative, doctrinal and mixed legal methods, to the extent that Wilson is arguing for a shift in the sites of investigation in this area, through the three stages that he identifies. One of the points that Wilson stresses on is that more work needs to be done in each local context in which these platforms operate. In the next section, I tie together the different strands that I have explored in this essay, and link this to the context of online hate speech regulation in India.

[40] *Id.* at 2-3.
[41] *Id.* at 4.
[42] *Id.* at 5.

CONCLUSION

In this essay I have argued that the specific aspects of hate speech online, that distinguish it from hate speech on older forms of media, necessitate a move from the focus of legal debates on the rhetoric of hate speech to its materiality, and this in turn links to opening up the question of the sites that it must be studied in. This is both a question of what we are studying and how we do it, or the methodological approaches we employ. I have argued that in order to take into account what online hate speech *does* one has to understand the infrastructures of online hate speech, including the use of algorithms to automate the detection of hate speech, and at the same time creating and codifying new publics.

In order to understand the regulation of hate speech online in India, one has to take into account what the scholars Aswin Punathambekar and Sriram Mohan highlight as the specific trajectories of digital media economies in South Asia, where there is a high degree of intermediality and media convergence.[43] They argue that in South Asia "we are still coming to terms with the postcolonial states' infrastructural dispositions, the aesthetics and affective power that infrastructures wield in public culture (e.g. dams, highways, cinema halls, cell phone towers) and crucially, the layering and convergences of different media infrastructures."[44] These specific conditions will impact the way in which the virality of online hate speech in India, and the way it is mediated through crowds, determines the kinds of legal infrastructures that are put in place for its regulation.

In terms of the specific role of legal regulation of online hate speech in India, the scholars Cherian George, William Mazzarella and Raminder Kaur, warn us about the nature of offense-taking in India, where hate speech is often spun[45] or manufactured through networks of individuals who see these controversies over speech as opportunities to magnify claims on the state as a form of publicity or 'profitable provocation'.[46] The motivations of these local actors play an important role in fanning controversies, and ensuring that hate speech *does* things – incites violence, which is often in the nature of inter-group violence. These actors now have at their disposal the ability to spread such motivated hatred with lightning speed, through easily

[43] MAPPING GLOBAL DIGITAL CULTURES: PERSPECTIVES FROM SOUTH ASIA 10 (Aswin Punathambekar & Sriram Mohan eds., University of Michigan Press 2019).

[44] *Id.* at 13.

[45] CHERIAN GEORGE, HATE SPIN: THE MANUFACTURE OF RELIGIOUS OFFENSE AND ITS THREAT TO DEMOCRACY (MIT Press, Cambridge, MA 2016).

[46] CENSORSHIP IN SOUTH ASIA: CULTURAL REGULATION FROM SEDITION TO SEDUCTION (William Mazzarella & Raminder Kaur eds., Indiana University Press, Bloomington 2009).

accessible platforms such as WhatsApp and Facebook. These platforms have been at the centre of a maelstrom in recent times, with pressures from government[47] and civil society[48] to make more effective the moderation of content in India. These platforms have gradually begun introducing measures to respond to the virality of hate speech in India, for instance, a limit to the number of forwards on WhatsApp,[49] which in turn have had global implications, as these measures were gradually standardised across the globe. These measures by platforms have been put even as the central and state law enforcement agencies in the country develop their own infrastructures to monitor speech online.[50]

These developments further highlight the importance of moving from the 'what' to 'how' question in relation to online hate speech regulation in India. Our understanding of legal standards of incitement to violence and causation, which are crucial to this area of law, will be incomplete if not placed in the context of the emerging legal infrastructures around online hate speech, and the ways in which platform, through infrastructures such as algorithms are already moderating content and shaping, creating and codifying publics.

[47] Chinmayi Arun, *On WhatsApp, Rumours, and Lynchings*, 44(6) ECONOMIC AND POLITICAL WEEKLY (2019).

[48] Chinmayi Arun, *Rebalancing Regulation of Speech: Hyper-Local Content on Global Web-Based Platforms*, Berkman Klein Centre for Internet & Society at Harvard University (Mar. 29, 2018), https://medium.com/berkman-klein-center/rebalancing-regulation-of-speech-hyper-local-content-on-global-web-based-platforms-1-386d65d86e32 (last visited June 8, 2020).

[49] *Facebook's WhatsApp Limits Users to Five Text Forwards to Curb Rumors,* REUTERS, JAKARTA, (Jan. 21, 2019), https://www.reuters.com/article/us-facebook-whatsapp/facebooks-whatsapp-limits-text-forwards-to-five-recipients-to-curb-rumors-idUSKCN1PF0TP (last visited Mar. 30, 2020).

[50] Siddharth Narrain, *Dangerous Speech in Real Time: Social Media, Policing and Communal Violence*, 52(34) ECONOMIC AND POLITICAL WEEKLY (2017).

ALGORITHMIC APPLICATIONS IN HEALTHCARE: KEY CONUNDRUMS

- Nishtha Bharti[*]

I. INTRODUCTION

The advent of big data and advanced analytics, combined with the rapid advancement of interactive digital technologies, has facilitated the envisioning of algorithmic applications in the Indian healthcare sector. This phenomenon has been especially accelerated by ubiquitous internet access and proliferation of consumer wearables that record and transmit metrics and track user activity in real time. With the steady move of digitizing electronic health records, there is presently an effusion of big data in the healthcare sector, which, when combined with enhanced computational power, has offered impetus to algorithmic initiatives in healthcare. The technology has garnered worldwide attention, especially in the discourse around improvement of quality and efficiency in healthcare delivery. Its deployment is conceived across various verticals – diagnostic decision support, patient monitoring, risk predictions, insurance and preventive medicine.[1] The alignment of artificial neural networks, natural language processing and computer vision are purported to facilitate healthcare interventions such as drug discovery, infectious disease surveillance, image analysis and virtual health assistants.[2] They have further paved the way for the development of completely new applications, products, services and business models, all with inherent opportunities and risks.

This essay seeks to identify the key conundrums that manifest through the unfolding of algorithmic infrastructures in the realm of healthcare, based on a preliminary mapping of this landscape in India. An effort is made, wherever possible, to situate these conundrums in existing scholarship. This is an exploratory exercise - it is not the ambition of the essay to propose definitive solutions. The intention here is to foster a discussion that must exist in tandem with the aggrandizement of algorithmic models.

In recent years, a certain matchmaking between digital platforms and computational algorithms has captured the imagination of Indian policy-makers, as a 'disruptive technology' that can

[*] PhD Candidate, Department of Humanities and Social Sciences, Indian Institute of Technology, New Delhi.

[1] Sandeep Reddy, et al., *Artificial Intelligence-Enabled Healthcare Delivery*, 112 JOURNAL OF THE ROYAL SOCIETY OF MEDICINE 22–28 (2019).

[2] *Id.* at 22.

address persistent socio-economic challenges and usher in 'good governance'. This optimism pervades through various budgetary announcements and position papers, key amongst them being the 'National Strategy for Artificial Intelligence'[3] - a Discussion Paper released in June 2018 by the NITI Aayog, the policy think-tank of the Government of India. Emphasizing on "how India can leverage the transformative technologies to ensure social and inclusive growth in line with the development philosophy of the government"[4], this document proposes healthcare as one of its first AI-driven focus areas.[5] After acknowledging the "challenges of quality, accessibility and affordability"[6] that plague the healthcare sector in India, the document goes on to make a case for how this very sector is an "obvious use case primed for intervention by AI driven solutions, as evidenced by the increasing activity from large corporate and startups alike in developing AI focused healthcare solutions".[7] It is this narrative around AI, imbued with the promise of panacean proportions that prompts this essay's effort to state the challenges that beset this promise. The following four sections elaborate on these challenges, framed by relevant literature and embedded within the ecology of a rapidly evolving technological assemblage.

II. OF DATA: ACCESS AND ACQUISITION

Emanating from the premise of equity, the *first conundrum* pertains to the concept of access - which has a bearing on various milestones in technological development and deployment. As with any algorithmic model, a key asset in its development in the domain of healthcare is the availability of and access to relevant medical data sets. This challenge can actualize differently for various developers: from a start-up to an academic collaboration to an industry giant. Based on their wherewithal to persuade medical establishments or research institutions to partner with them and provide them with training data, developers may end up either with first-hand medical records or with open access datasets available online, which might not be reflective of the target population for whom they are building the algorithm. This results in what Ryan Calo terms the 'Data Parity Problem',[8] where asymmetric access to sufficient training data can influence a proliferation of ML applications in the profit-driven private sphere, since big firms or institutions are better positioned to acquire appropriate databases. Such a situation has the potential to trigger the centralisation of power itself, as well as the coalition of particular interests whose values may

[3] National Strategy for Artificial Intelligence, NITI AAYOG, https://niti.gov.in/national-strategy-artificial-intelligence (accessed June 23, 2020).
[4] *Id.* at 5.
[5] *Id.* at 7.
[6] *Id.* at 24.
[7] *Id.* at 27.
[8] Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS LAW REVIEW 398-435 (2017).

not necessarily align with the government's purported noble objectives of introducing a technology like AI in healthcare.[9] The questions therefore, of who owns the data and who gets access to it, invariably connect to who eventually benefits from the algorithm and to what ends it is put to use.

The existing situation in the healthcare sector in India is that on one hand, it is largely financed through out-of-pocket payments and on the other hand, access to specialists, facilities and medication in a timely manner is mostly limited to urban centres.[10] Further, there is an inherent asymmetry in the constitutional division of power and responsibilities between the central and the state governments. As Sujatha Rao has helpfully elaborated,[11] while the states are responsible for delivering healthcare services, it is the centre that dispenses the funding and resources, resulting in inter-state disparities in health status. The economic sustainability of public health delivery system is further compromised due to bureaucratic protocols and procedures that routinely delay the release of funds. Frequent revisions in budgeting cycles and the arbitrary manner of imposing cuts - mostly a result of political misalignment - further generate uncertainty. Finally, the pervasive distortions in service delivery due to a weak public sector infrastructure has buttressed the private sector in the domain of healthcare.[12]

In a scenario where algorithmic applications potentially unfold without timely interventions from the government at the right junctures (as elaborated further) and certain players have privileged access to the initial (data) resources for algorithm development, it is likely that we will witness a preferential standing of the rights of those patients at private establishments who are able to overcome the socioeconomic, geographic or other obstacles that hinder access to physicians in the first place. It is pertinent to explore then, where and for whose benefit algorithmic systems are being deployed within the healthcare domain, since these variables determine access to healthcare and equitability of health outcomes.

[9] Soham H. Bhaduri, *Only a Strong Public Health Sector Can ensure fair prices and quality care at private hospitals,* SCROLL.IN (March 28, 2019, 2:30 PM), https://scroll.in/pulse/917578/only-a-strong-public-health-sector-can-ensure-fair-prices-and-quality-care-at-private-hospitals; Himani Chandna. *Fortis, Max, Medanta - Why Private Indian Hospitals are Selling Out to Foreign Players*, THE PRINT (July 1, 2019, 9:31 AM), https://theprint.in/economy/this-is-why-private-indian-hospitals-are-selling-out-to-foreign-players/255874/.

[10] Debasis Barik and Amit Thorat, *Issues of unequal access to public health in India*, 3 FRONTIERS IN PUBLIC HEALTH 1-3 (2015); Anamika Pandey et al., *Trends in Catastrophic Health Expenditure in India: 1993 to 2014*, BULLETIN OF THE WORLD HEALTH ORGANIZATION (Nov. 30, 2017), https://www.who.int/bulletin/volumes/96/1/17-191759/en/.

[11] SUJATHA RAO, DO WE CARE? INDIA'S HEALTH SYSTEM (Oxford Scholarship Online 2017).

[12] *Id.* at 40, 41.

III. OF ALGORITHMS: DATA CONTEXT AND CHARACTERISTICS

The *second conundrum* precipitated by algorithmic applications has acquired significant purchase in the literature on Critical Data Studies regarding the validity and representativeness of big data. It is argued that data sets which are used to train learning algorithms are often poorly representative of the wider population and, as a result, such algorithms could make unfair decisions that reflect the pervasive biases in society. The phenomenon has been succinctly described by Barocas & Selbst - "*Data is frequently imperfect in ways that allow these algorithms to inherit the prejudices of prior decision makers. In other cases, data may simply reflect the widespread biases that persist in society at large. In still others, data mining can discover surprisingly useful regularities that are really just preexisting patterns of exclusion and inequality. Unthinking reliance on data mining can deny historically disadvantaged and vulnerable groups full participation in society.*"[13] Kate Crawford similarly highlights the 'signal problems' in big-data, pertaining to "dark zones or shadows where some citizens and communities are overlooked or underrepresented."[14]

This implies that a decisive difficulty with big data is that the insights it furnishes are mostly dissociated from the complexity of experiences of those on the margins, whose circumstances are not sufficiently documented and therefore too easily sidelined.[15] As Lerman has pointed out, people "who do not routinely engage in activities that big data is designed to capture" remain at its periphery and "their preferences and needs risk being routinely ignored when governments and private industry use big data and advanced analytics to shape public policy and the marketplace".[16] The major source of big data in healthcare sector is consumer information from electronic health records, health insurance claims, pharmaceutical drug prescription and usage trends, as well as health applications and wearables. All of these only cover persons who have sought healthcare at some point or at least operate devices equipped with digital health assistants - a predicate which itself suffers from disparities. What is ironic here, especially in a country like India, is that individuals who are unable to integrate into the big data trail are the very people most in need of increased health research, intervention, and care.

Any demographic inconsistency in the data sets with which algorithms are trained implies that certain cohorts may be excluded from the purview of subsequent operationalisation of the

[13] Solon Barocas and Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671 (2016).

[14] Kate Crawford, *The Hidden Biases in Big Data*, HARV. B. REV. (Apr. 1, 2013), https://hbr.org/2013/04/the-hidden-biases-in-big-data.

[15] Sarah E. Malanga et al., *Who's Left out of Big Data?*, *in* BIG DATA, HEALTH LAW, AND BIOETHICS 98-111 (I. Glenn Cohen et al. eds., 2018).

[16] Jonas Lerman, *Big Data and Its Exclusions*, 66 STAN. L. REV 55 (2013).

algorithmic model, while others might be favoured. Further, biases embedded in these data sets can have a cascading effect, and reinforce longstanding health disparities. By designing systems at scale based on existing patterns in society, prevailing prejudices or structural inequalities may not only be reproduced, but also escalated. It is in this sense that exclusion from the algorithmic infrastructure might translate into a barrier to full citizenship rights and in solidifying existing socio-economic stratifications because "only people who matter - quite literally the only ones who count - are those who regularly contribute to the right data flows."[17]

Further, algorithms driven by data mining and visualisation programs are, as Kitchin has asserted, "very poor at capturing and deciphering meaning or context" of the information being processed.[18] Making abstractions from data without situating them in its context, delegitimizes the 'lived experience' of diverse social groups that might not render itself to quantification. The employment of big data and the deluge of information that constitutes it, brings about an undercurrent of indifference towards alternate forms of analysis and research that looks beyond the sheer volume of numbers to locate meaning in aspects of everyday life. The application of a trained machine learning system to an unanticipated patient context can then subvert the very rationale of replacing human agents with AI systems - that of efficiency and accuracy. When algorithms trained on specific datasets are deployed outside the context of operation of those datasets, they might not necessarily function according to appropriate standards, resulting in flawed performance. For instance, such 'transfer context bias'[19] could arise in administering a healthcare algorithm functioning on data garnered from urban, exposed-to-Internet of Medical things (IoMT) population to persons accessing a rural clinic. The same heuristic model cannot be applied to both settings as the amenities available to them to generate the initial data differ remarkably. Since the new setting will most likely have markedly different characteristics, the system would almost certainly have significant algorithmic bias. The assumption inherent in the design of such an algorithm - that of efficiently identifying disease incidence - would not hold, and would provide an inequitable service. The algorithm in this way might not capture the diversity amongst citizens, whose inclusion is crucial for the sustainability of health policy decisions. It therefore becomes crucial for technology developers as well as their medical partners

[17] *Id.* at 60.

[18] Rob Kitchin, *Big Data, New Epistemologies and Paradigm Shifts*, 1 BIG DATA AND SOCIETY 1, 4 (2014).

[19] David Danks & Alex J. London, *Algorithmic Bias in Autonomous Systems*, PROCEEDINGS OF THE 26TH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE (IJCAI) 4691-97 (2017).

to understand and account for "not only the limits of the data set, but also the limits of which questions they can ask of a data set and what interpretations are appropriate."[20]

IV. OF KNOWLEDGE: EXPERIENTIAL OR QUANTIFIABLE?

The *third conundrum* pertains to a disjuncture between the disciplinary demands of medicine and engineering - the two specialties brought together in medical applications of artificial intelligence. In the practice of medicine, the essence of a diagnostic process embodies two elements. The first is the idea of tacit knowledge - such as contextual understanding and the ability to read social cues - wisdom that is accumulated with experience.[21] The second involves referring to a complete corpus of retrospective data that emerge at different points in a patient's medical history.[22] It is a combination of both these elements - inter-observer variability and experiential diversity amongst doctors - that more often than not results in differential diagnosis, and is also responsible for lack of consensus in 'ground truth' amongst doctors.

Let us take the instance of Radiology and Pathology, where most of the algorithms being developed are trained on images of diagnosed lesions using histopathology. Their performance rests on these images as a reference point, under the assumption that the initial diagnosis on these images was correct. But it has been well documented in medical literature that there exists a high disagreement rate among different medical practitioners regarding interpretation of histopathology, especially for borderline cases.[23] This is precisely why clinicians do not solely rely on morphology (clinical, dermascopic, histopathological) for disease detection, but also venture into patient history, as well as their past experience with similar cases. It is difficult to conceptualize how this entire corpus of knowledge can be made available for training an AI system, which requires clear endpoints (benign-malignant) by its very design.

The challenge here for algorithmic applications is twofold. First, in the absence of any clinical immersion, the engineers who are developing the algorithms may be inclined to develop a measure of accuracy for their system that does not take into account the real-world impact of a

[20] Danah Boyd & Kate Crawford, *Critical Questions for Big Data*, 15 INFORMATION, COMMUNICATION & SOCIETY 662, 670 (2012).

[21] SIDDHARTHA MUKHERJEE, THE LAWS OF MEDICINE: FIELD NOTES FROM AN UNCERTAIN SCIENCE (Simon & Schuster 2015).

[22] Tessa S Cook, *Human versus Machine in Medicine: Can Scientific Literature Answer the Question*, 1 LANCET DIGITAL HEALTH, 246–247 (2019).

[23] Joann G Elmore et al, *Pathologists' Diagnosis of Invasive Melanoma and Melanocytic Proliferations: Observer Accuracy and Reproducibility Study*, 357 BMJ 2813 (2017); Aimilios Lallas and Giuseppe Argenziano, *Artificial Intelligence and Melanoma Diagnosis: Ignoring Human Nature May Lead to False Predictions*, 8 DERMATOLOGY PRACTICAL AND CONCEPTUAL 249-251 (2018).

diagnosis in the same way as a doctor would. Instead, they may simply rely on a metric that maximizes performance on testing data. Second, because most doctors do not relate to numbers in the way that engineers do, it might be difficult for them to agree to a measure of accuracy for the algorithm that caters both to medical wisdom as well as mathematical precision. In this dynamic, it becomes important to think about who in this partnership decides whether we have the right algorithm and further, how we appropriately mark the limits of its 'expertise'.[24] The manner in which we approach these questions will determine policy positions on both accountability and liability, as discussed in the next conundrum.

V. OF REGULATION: BENCHMARKING AND LIABILITY

The *final conundrum* appertains the lack of an appropriate regulatory paradigm that ensures standardization, accountability and liability in operationalising algorithmic applications. In the Indian context, there exist a range of policy measures around data sharing and its attendant concerns - National Health Stack (2018), Digital Information Security in Healthcare Act (2018), National Digital Health Blueprint (2019) and the Personal Data Protection Bill (2019). One common theme that weaves through these initiatives, and has been a consistent focus of scholarship is the notion of privacy.[25] Privacy is indeed an important consideration when working with clinical data, which often contains sensitive information that needs to be protected from un-authorized access and unintentional disclosure. But what is also of great import about medical data is the notion of vulnerability, which especially in the medical context requires us to re-evaluate the contours of both privacy and consent in the legal domain. As Calo has demonstrated, vulnerability and privacy intersect in a unique manner, wherein "the more vulnerable a person is, the less privacy they tend to enjoy; meanwhile, a lack of privacy opens the door to greater vulnerability and exploitation".[26] In a country like India where patients already have very restricted agency in terms of their interaction with healthcare professionals, it is difficult to imagine individuals at the cusp of receiving a medical diagnosis to reasonably understand or exercise their rights to privacy and informed consent. This is especially true for marginalized sections of the society, whose vulnerability is more layered than their privileged

[24] Kate Crawford et al., *The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term*, AI NOW REPORT, 1–25 (2016).

[25] Nayantara Narayanan, *Niti Aayog plan for Aadhaar-linked digital health records raises concerns over safety and privacy*, SCROLL.IN (Jul. 12, 2018, 9:00 AM), https://scroll.in/pulse/886153/niti-aayog-plan-for-aadhaar-linked-digital-health-records-raises-concerns-over-safety-and-privacy; Apar Gupta, *The Data Protection Bill only weakens user rights,* THE HINDU (Dec. 27, 2019, 12:08 AM), https://www.thehindu.com/opinion/lead/the-data-protection-bill-only-weakens-user-rights/article30405339.ece; Kamal Taneja & Gulshan Rai, *Data Protection Bill is vague and intrusive*, THE HINDU BUSINESS LINE (Mar. 15, 2020), https://www.thehindubusinessline.com/opinion/data-protection-bill-is-vague-and-intrusive/article31075785.ece.

[26] Ryan Calo, *Privacy, Vulnerability, and Affordance*, 66 DE PAUL L. REV. 591-604 (2017).

fellow citizens. Re-evaluating the safety net offered to patients, in the light of this trope, remains a pressing concern.

Another limitation of the above-mentioned government initiatives is that they primarily circumscribe issues related to data – digitization, storing, interoperability, privacy, consent. And while these are all important motifs, they reflect just one aspect of algorithmic deployment. So far there is no proposed mechanism proposed regarding running an algorithmic model through a protocol of tests to standardize it. This blind spot in regulation implies that developers will apply their own discretionary benchmarking through iterative testing and matching it with peer reviewed literature. The difficulty with such a practice is that in the literature around AI applications in medicine, there is not yet an agreed way to report findings or even compare the diagnostic accuracy of such systems.[27] And there are even fewer studies that compare the performance of these systems with humans. What adds to the complexity of evaluating performance of AI systems is the fact that negative studies – i.e., those that do not disprove the null hypothesis - are reported and published less frequently.[28] Any benchmarking from available literature, would therefore be insufficient because this literature is skewed in favour of those algorithmic models that perform well. Moreover, as with any medical innovation, while it is the best practice in the industry to get external validation from a third party medical establishment or practitioner, there are presently no monitoring and compliance guidelines in India recommending who qualifies to disburse such a validation, at what scale and under what terms, when it comes to pilot testing algorithmic models.

In the spirit of keeping pace with this technology, and in congruence with the enthusiasm reflected in government documents regarding its intrinsic promise, it is vital that we also contemplate about the legal standards and accountability mechanisms that govern data-driven algorithmic decision-making, since "the tools currently available to policymakers, legislators, and courts were developed to oversee human decision makers and often fail when applied to computers instead."[29] It is important to talk about regulation because regulation is intertwined with the idea of liability. At the core of the relationship between physicians and patients is an element of fiduciary duty, where the contours of liability are very lucid. But for making an algorithmic system accountable and circumscribing its outcomes with legal responsibility, it

[27] Robert Challen et al., *Artificial Intelligence, Bias and Clinical Safety*, 28 BMJ QUALITY AND SAFETY 231-237 (2019).
[28] Tessa S Cook, *Human versus Machine in Medicine: Can Scientific Literature Answer the Question*, 1 LANCET DIGITAL HEALTH, 246–247 (2019).
[29] Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2017).

would be instructive to explore the degree of granularity with which we can understand and explain an AI system. There is a profusion of studies which meticulously engage with the problem of explainability, interpretability and the black-boxed nature of artificial intelligence.[30] These are contested concepts, and as many of the authors mentioned have suggested, they need to be situated in the wider socio-political-technical milieu to comprehend the ethical ramifications of this emerging technology. So even as we grapple with the underpinnings of artificial neural networks and deep learning, and whether or not they can be rendered to scrutiny, we must bear in mind the counsel of Daniels & Williams, "*at the outset of system development, developers should also be able to describe how their system will function in the field, including the objectives it aims to achieve and the tasks it will undertake, the technologies it will rely on to do so, and the technical, legal, and ethical risks inherent to using those technologies*".[31] To ensure accountability in algorithmic applications, policy-makers then ought to deliberate and legislate beyond stipulating conditions around data usage.

In conclusion, the objective of dwelling on the above conundrums is to emphasize the need for contextualizing this technology in the fragile but unique model of Indian healthcare, and to move forward keeping in mind the fallibility of technological design, however pioneering it might appear. Like most emerging technologies, algorithms demand our attention and scrutiny because they raise complex ethical, legal, and security concerns surrounding issues such as equitability, accountability, fairness and autonomy. So even as we move towards the possibility of artificial agents taking a larger role in decision-making processes, we need to pay more attention to the kinds of encounters algorithmic systems generate, whether it is through the vagaries of their technological architecture or through the constitutive conditions that are made available to them by state machinery.

[30] Nicholas Diakopoulos, *Algorithmic Accountability Reporting: On the Investigation of Black Boxes*. TOW CENTER FOR DIGITAL JOURNALISM, 1-33 (2013); Mike Ananny and Kate Crawford, *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*, 20 NEW MEDIA AND SOCIETY, 973–989 (2018); Aaron M. Bornstein, *Is Artificial Intelligence Permanently Inscrutable?* NAUTILUS, September 1, 2016; Jenna Burrell, *How the machine 'thinks': Understanding opacity in machine learning algorithms*. BIG DATA AND SOCIETY, 1-12 (2016); Zachary C. Lipton, *The Mythos of Model Interpretability*, 16 ACM Queue, 1-27 (2018); Patrick Hall, Sri Satish Ambati and Wen Phan, *Ideas on interpreting machine learning*. OREILLY (Mar. 15, 2017), https://www.oreilly.com/radar/ideas-on-interpreting-machine-learning/; Siddhartha Mukherjee, *A.I. Versus M.D. What happens when diagnosis is automated?* THE NEW YORKER: ANNALS OF MEDICINE (Mar. 27, 2017), https://www.newyorker.com/magazine/2017/04/03/ai-versus-md; Paul Voosen, *How AI detectives are cracking open the black box of deep learning*. SCIENCE (Jul. 6, 2017, 2:00 PM), https://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning.

[31] Owen Daniels & Brian Williams, *Day Zero Ethics for Military AI*, TEX. NAT'L SEC. REV. (Jan. 28, 2020), https://warontherocks.com/2020/01/day-zero-ethics-for-military-ai/.

PANOPTIC JURISPRUDENCE OF A BIG BROTHER STATE: SURVEILLANCE AND PRIVACY VIS-À-VIS SECTION 69 OF THE INFORMATION TECHNOLOGY ACT

- *Snehil Kunwar Singh*[*]

ABSTRACT

The harrowing depiction of the Orwellian Big Brother State in '1984' is the perfect example of what unbridled mass surveillance can do to society. It helps us understand, intuitively, the threat of state scrutiny to our lives. Extending beyond Orwell's idea of the Big Brother State and Bentham's Panopticon, digital technology has revolutionized our lives and more so with the ability and willingness of governments to use and acquire our data which is used for unknown purposes often shrouded in secrecy. Unbridled surveillance without due safeguard affects 'intellectual privacy' and casts a chilling effect on right to freedom of speech and expression. Recognition of right to privacy as a fundamental right exemplifies the inherent tension between right against surveillance and compelling state interest in surveillance. These competing claims need to be adequately balanced by providing for judicial safeguards in the form of ex-parte ex-ante judicial scrutiny of surveillance orders where the executive today in an emergency is empowered to undertake surveillance with ex-post approval and 'narrow tailoring' of grounds on which surveillance can be carried by the executive.

I. INTRODUCTION

"*Time works changes and brings into existence new conditions. Subtler and far reaching means of invading privacy will make it possible to be heard in the street what is whispered in the closet.*"[1]

In the age of technology, the above observation of the Supreme Court in *Gobind* v. *State of Madhya Pradesh*[2] seems to have indeed become true. Present times witness excessive state surveillance[3]

[*] Third Year, B.A., LL.B. (Hons.), National Law School of India University, Bangalore.

[1] Gobind v. State of Madhya Pradesh, (1975) 2 SCC 148 [hereinafter Gobind].

[2] *Id.* at 155.

[3] The Government of India *vide* order Government Order S.O. 6227(E), Ministry of Home Affairs, Government of India has permitted ten agencies to carry out surveillance – (i) Intelligence Bureau, (ii) Narcotics Control Bureau, (iii) Enforcement Directorate, (iv) Central Board of Direct Taxes, (v) Directorate of Revenue Intelligence, (vi) Central Bureau of Intelligence, (vii) National Investigation Agency, (viii) Cabinet Secretariat (RAW), (ix) Directorate of Signal Intelligence (For service areas of Jammu & Kashmir, North-East and Assam only), (x) Commissioner of Police, Delhi. In the aforesaid list, agencies such as Commissioner of Police, Delhi; Directorate of Revenue Intelligence; the CBI, strictly do not work for homeland security and thereby not fulfilling the criteria under Section 69 of The Information Technology Act, 2000. Tathagata Satpathy, *Are India's laws on surveillance a threat to privacy?*, THE HINDU, (Dec. 28, 2018) https://www.thehindu.com/opinion/op-ed/are-indias-laws-on-surveillance-a-threat-to-privacy/article25858338.ece (last visited June 8, 2020).

under 'archaic' laws which have failed to keep pace with technological advancements. Consequently, infraction of the fundamental right to privacy is a recurring phenomenon where affected individuals and groups have no redressal of such an infraction of the right to privacy. In fact, the right to privacy when contrasted with state surveillance under Section 69 of Information Technology Act, 2000 ['IT Act'] presents a characteristic constitutional phenomenon – exemplification of the inherent tension between democracy's two fundamental elements. On the one hand is the fundamental right to privacy, existing equally in all individuals irrespective of class or strata, gender or orientation which is an inviolable right of human inhered in him; on the other hand, is the people element, limiting those very rights through their representatives.

Described as 'constructive tension',[4] this tension is resolved by proper balancing of competing principles which enable each facet to coexist harmoniously with the other, and not one in place of the other. This is more so because in the interest of the security and sovereignty of India and to deal with any other emergency for the protection of national interest, it may indeed be required that messages be intercepted. The core question that lies is whether there exist sufficient procedural safeguards to rule out arbitrary exercise of power and what is the manner in which such surveillance should be carried out.

I shall begin by outlining the mandate and ambit of Section 69 of IT Act and how it impinges upon right to privacy. I will, then, trace right to privacy in Indian jurisprudence and analyse how the declaration of right to privacy has changed the discourse regarding Section 69 of IT Act and the right to privacy. Thereafter, I will deal with how Section 69 of IT Act is relevant to the theory of 'intellectual privacy' and how the doctrine of 'chilling effect' has an effect on the sacrosanct right of free speech. Subsequently, I shall discuss judicial standards set for permissible infraction of the right to privacy. I shall also proceed to show how Section 69 falls foul of established judicial standards. Finally, I will conclude with suggestions on how to prevent the vice of unconstitutionality. I argue that Section 69 should be suitably read down to include procedural safeguards to rule out arbitrariness and to prevent indiscriminate intrusions into privacy.

II. SECTION 69 OF IT ACT: PRIVACY IN HANDS OF EXECUTIVE WITH NO CHECKS

[4] Modern Dental College and Research Centre v. State of Madhya Pradesh, (2016) 7 SCC 353, ¶ 62 [hereinafter Modern Dental College].

Generally understood as the defining legal regime for digital data in India,[5] the IT Act 2000 came into being after the UNCITRAL Model Law on E-Commerce, 1996[6] was adopted by the United Nations General Assembly and its members passed a resolution to the effect that the member states should consider and try to incorporate the Model Law on E-Commerce when amending or enacting their domestic laws.[7] Thereafter, being a member of the United Nations, India ratified this resolution and enacted the Information Technology Act in 2000. IT Act has extra-territorial application.[8] Section 69 of IT Act was amended through Information Technology (Amendment) Act 2008 and has been subject to much controversy, as it was passed without debate in Parliament.[9] The constitutional validity of Section 69 and rules framed thereunder have been assailed quite a few times and the constitutional challenge has gained fresh interest[10] following the declaration of the right to privacy as a fundamental right under Article 21 of the constitution by a 9-judge bench of the Supreme Court in *Justice K.S. Puttaswamy (Retd.) v. Union of India* [*Privacy Judgement*].[11]

Section 69 allows State or Central government, and officers authorised by the government to carry out surveillance to "*intercept, monitor or decrypt*" data "*in the interest of the sovereignty or integrity of India, defence of India, security of the State, friendly relations with foreign States or public order*".[12] It is noteworthy that '*defence of India*', which is a ground for carrying out surveillance, is neither provided for under Article 19(2) nor defined under the IT Act.

[5] CRID-UNIVERSITY OF NAMUR, FIRST ANALYSIS OF THE PERSONAL DATA PROTECTION LAW IN INDIA FINAL REPORT, http://ec.europa.eu/justice/policies/privacy/docs/studies/ final_report_india_en.pdf. (last visited June 8, 2020).

[6] UNCITRAL Model Law on Electronic Commerce, (1996) at http://www.uncitral.org/ uncitral/en/uncitral_texts/electronic_commerce/1996Model.html. (last visited June 8, 2020).

[7] General Assembly, Model Law on Electronic Commerce Adopted by the United Nations Commission on International Trade Law, 51/162 at http://daccessddsny.un.org/doc/ UNDOC/GEN/N97/763/57/PDF/N9776357.pdf?OpenElement and as noted in the introduction to the Information Technology Act 2000: "...Whereas the General Assembly of the United Nations by resolution A/RES/1/162, dated the 30th January, 1997 has adopted the Model Law on Electronic Commerce adopted by the United Nations Commission on International Trade Law. And Whereas the said resolution recommends inter alia that all States give favourable consideration to the said Model Law when they enact or revise their laws, in view of the need for uniformity of the law applicable to alternatives to paper–cased methods of communication and storage of information; and Whereas it is considered necessary to give effect to the said resolution and to promote efficient delivery of Government services by means of reliable electronic records..."

[8] Information Technology Act, 2000, § 1(2), No. 21, Acts of Parliament, India (2000) [hereinafter Information Technology Act, 2000].

[9] M. Kaul, *India has an internet problem*, OPEN DEMOCRACY (Mar. 13, 2013), http:// www.opendemocracy.net/openindia/mahima-kaul/india-has-internet-problem.

[10] Recently, two petitions have been filed before the Supreme Court – Saurabh Pandey v. Union of India, (Jan. 2, 2019), https://barandbench.com/wp-content/uploads/2019/01/Saurabh-Pandey-v-UOI-PIL.pdf; Internet Freedom Foundation v. Union of India, https://www.livelaw.in/pdf_upload/pdf_upload-357441.pdf – assailing the constitutional validity of Section 69, The Information Technology Act, 2000, (last visited June 8, 2020).

[11] Justice K.S. Puttaswamy (retd.) v. Union of India, Writ Petition (Civil) No. 494 of 2012 ["Privacy Judgement"].

[12] Information Technology Act, 2000, *supra* note 8, § 69.

Section 69 does not provide for how exactly the power of surveillance is to be exercised, rather it leaves it to delegated legislation under Section 69 (2) to provide for procedure and safeguard for carrying out surveillance. Surveillance is carried out on written orders of a 'competent authority' which comprises the Secretary in the Ministry of Home Affairs, in case of the Central Government;[13] the Secretary in charge of the Home department, in case of a State Government or Union territory.[14] In unavoidable circumstances, such order of surveillance may be issued by an officer not below the rank of Joint Secretary of the Government of India, who has been duly authorised by the competent authority;[15] and in case of emergency it may be carried out with the prior approval of the Head or the second senior most officer of the security and law enforcement agency at the Central level and the officer authorised in this behalf, not below the rank of the Inspector General of Police or an officer of equivalent rank, at the State or Union territory level.[16] The order of the 'competent authority' is reviewed once in two months by a review committee same as the one constituted under rule 419A of Indian Telegraph Rules, 1951[17] which comprises of Cabinet Secretary, Secretary to the Government of India, In-charge, Legal Affairs, Secretary to the Government of India, Department of Telecommunications for Central Government,[18] and Chief Secretary, Secretary Law/Legal Remembrancer In-charge, Legal Affairs, Secretary to the State Government (other than the Home Secretary) for State Government.[19]

The only safeguard under which such exercise is to be carried out is through the IT Rules framed under the IT Act and IT Act itself, which fails to offer any 'real' safeguard. The intermediary providing information is liable for criminal action for breach of duty to provide information for a term which may extend to three years, besides fine.[20] However, the IT Rules make the intermediary responsible for breach of secrecy through 'relevant provisions of the laws for the time being in force'.[21] No such responsibility has been placed on the 'competent authority'. Further, the IT Rules do not provide for review by judiciary authority.

[13] Information Technology (Procedure and Safeguards for Interception, Monitoring and Decryption of Information) Rules, 2009, Rule 2(d)(i), Gazette of India, pt. II sec. 3(i) (Oct. 27, 2009) [hereinafter IT Interception Rules].
[14] *Id.*
[15] *Id.* Rule 3.
[16] *Id.*
[17] *Id.* Rule 2(q).
[18] Indian Telegraph (Amendment) Rules, 2007, Rule 16, Gazette of India, pt. II sec. 3(i) (Mar. 1, 2007).
[19] *Id.*
[20] Information Technology Act, 2000, *supra* note 8, § 69B (4).
[21] IT Interception Rules, *supra* note 13, Rule 21.

III. RIGHT TO PRIVACY: JOURNEY SO FAR & FUNDAMENTAL CHANGES IN UNDERSTANDING OF SURVEILLANCE

Before the right to privacy was declared as a fundamental right, there were two prevailing judicial opinions regarding right to privacy as a fundamental right – one supporting it and the other one opposing it. Opposition to the right to privacy as a fundamental right was grounded in two decisions: *M.P. Sharma v. Satish Chandra, District Magistrate, Delhi*[22] ["MP Sharma"], a judgement delivered by a Bench of eight judges; and *Kharak Singh v. State of Uttar Pradesh*[23] ["Kharak Singh"], a judgement rendered by a Bench of six judges. Both these judgements had ruled that the Indian Constitution does not specifically protect the right to privacy. The other side which supported the right to privacy as a fundamental right relied upon the fact that *MP Sharma* and *Kharak Singh* were founded upon the principles of exclusivity of fundamental rights expounded in *A.K. Gopalan v. State of Madras*,[24] a judgement which was no longer a good law by virtue of the decision of eleven-judge bench in *Rustom Cavasji Cooper v. Union of India*.[25]

In such a scenario where the question on the right to privacy as fundamental right was unsettled, the standard of judicial scrutiny of tapping and surveillance had not evolved and was merely confined to the '*just, fair and reasonable*' test under Article 21. It is also interesting to note that the judiciary had considered constitutionality of surveillance only in terms of telephone tapping under Indian Telegraphs Act, 1885 in *PUCL v. Union of India*.[26] Laying down procedural safeguards in telephone-tapping in *PUCL v. Union of India*,[27] the Supreme Court mandated review of a tapping order by a Review Committee and such an order could only be issued by Home Secretary, Government of India, and Home Secretary, State Government in case of Central Government and State Governments respectively.[28] In an urgent case, the power could be delegated to an offer of the Home Department of the Government of India and the State Governments not below the rank of Joint Secretary.[29] The Review Committee, at the level of Central Government, shall consist of Cabinet Secretary, the Law Secretary and the Secretary, Telecommunication. At the level of State Government, the Review Committee shall consist of

[22] M.P. Sharma v. Satish Chandra, District Magistrate, Delhi, (1954) SCR 1077.
[23] Kharak Singh v. State of Uttar Pradesh, (1964) 1 SCR 332 [hereinafter Kharak Singh].
[24] A.K. Gopalan v. Union of India, AIR 1950 SC 27.
[25] Rustom Cavasji Cooper v. Union of India, (1970) 1 SCC 248.
[26] People's Union for Civil Liberties v. Union of India, AIR 1997 SC 568 [hereinafter PUCL].
[27] *Id.*
[28] *Id.* ¶ 35.
[29] *Id.*

Chief Secretary, Law Secretary and another member, other than the Home Secretary, appointed by the State Government.[30]

At that relevant point in time, the judiciary had not possibly envisaged the methods and modern ways of surveillance such as geo-tagging and facial recognition. These methods substantially increase the deep and all-pervasive control of the State over data of individuals and groups. Our understanding of Section 69 needs to be revisited with the development of jurisprudence on privacy in relation to the doctrine of 'chilling effect'[31] and establishing judicial norms for infraction of privacy[32] post declaration of the right to privacy as a fundamental right.[33] Even more so when it has been clearly highlighted that the safeguards laid down by the Supreme Court in *PUCL case* do not seem to be effective. Recently, an expert committee was constituted by Central government under the chairman of Justice BN Srikrishna which noted that the Review Committee mandated by the Supreme Court in *PUCL case* convenes once every two months, and the task is 'unrealistic' as it has to review more than 15,000 – 18,000 surveillance orders in every meeting.[34]

IV. IMPACT OF SURVEILLANCE: INTELLECTUAL PRIVACY & DOCTRINE OF 'CHILLING EFFECT'

Mass surveillance directly impacts the way we exercise our civil liberties – thinking, speaking, communication, reading, and the way we make up our minds about social and political issues. Such a phenomenon is particularly dangerous as it tends to suppress dissent in a democracy and forces us to toe the line of majority as there is constant fear of prosecution by the State. These will be analysed in this section by understanding why and how surveillance is problematic, its impact upon 'intellectual privacy', the 'chilling effect' on the exercise of freedom of speech, and the power imbalance caused between citizens and state.

[30] *Id.*

[31] Shreya Singhal v. Union of India, (2015) 5 SCC 1 [hereinafter Shreya Singhal]; R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632 [hereinafter R. Rajagopal]; Ram Jethmalani v Subramaniam Swamy, 2006 SCC OnLine Del 14 [hereinafter Ram Jethmalani].

[32] Justice K.S. Puttaswamy (retd.) v. Union of India, Writ Petition (Civil) No. 494 of 2012 [hereinafter Aadhar judgment]; Aadhar (Targeted Delivery of Financial and Other Subsidies, Benefits and Services) Act, 2016, § 33(2), which provided for disclosure of information in certain cases to executive was struck down on the ground that it did not provide for judicial oversight, AK Sikri, J., Aadhar Judgement, ¶ 349.

[33] Privacy judgement, *supra* note 11.

[34] Justice BN Srikrishna Committee, *A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians*, MINISTRY OF ELECTRONICS & INFORMATION TECHNOLOGY, GOVERNMENT OF INDIA, https://meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf, 125 (last visited June 8, 2020).

(1) INTELLECTUAL PRIVACY

Intellectual privacy has been defined as the "*ability, whether protected by law or social circumstances, to develop ideas and beliefs away from the unwanted gaze or interference of others*."[35] The idea that free minds form the foundation of a free society is the underlying basis of the theory of intellectual privacy, and surveillance of activities of individuals casts a 'chilling effect'[36] upon the exercise of free thought.[37] It is necessary to protect and promote this individualism which forms the core commitment of our constitutional culture. This was recognised by Justice DY Chandrachud in *Indian Young Lawyers Association v. State of Kerala*:

"*At the heart of transformative constitutionalism, is a recognition of change. What transformation in social relations did the Constitution seek to achieve? What vision of society does the Constitution envisage? The answer to these questions lies in the recognition of the individual as the basic unit of the Constitution. This view demands that existing structures and laws be viewed from the prism of individual dignity*."[38]

These individual rights must not be denied in the name of some community interests. Ronald Dworkin, in his work '*Taking Rights Seriously*', states that:

"***Individual rights are political trumps held by individuals.*** *Individuals have rights when, for some reason, a collective goal is not a sufficient justification for denying them what they wish, as individuals, to have or to do, or not a sufficient justification for imposing some loss or injury upon them*."[39] (Emphasis supplied)

Dealing with the question of whether the Government may abridge the rights of others to act when their acts might simply increase the risk, by however slight or speculative a margin, that some person's right to life or property will be violated, Dworkin says:

"*But no society that purports to recognize a variety of rights, on the ground that a man's dignity or equality may be invaded in a variety of ways, can accept such a principle*[40] … *If rights make sense, then the degrees of their importance cannot be so different that some count not at all when others are mentioned*[41] … *If the Government does not take rights seriously, then it does not take law seriously either*."[42]

[35] Neil M. Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387 (2008) [hereinafter Richards].

[36] This was recognised by the Supreme Court of India in R. Rajagopal, *supra* note 31 and Ram Jethmalani, *supra* note 31 by the Delhi High Court.

[37] Richards, *supra* note 35, at 403-05.

[38] Indian Young Lawyers Association v. State of Kerala, 2018 SCC OnLine 1690, ¶ 277.

[39] RONALD DWORKIN, TAKING RIGHTS SERIOUSLY (Duckworth 1977).

[40] *Id.* at 203.

[41] *Id.* at 204.

[42] *Id.* at 205.

Intellectual privacy, a subset of privacy, is not only limited to 'intellectuals' but is rather a very essential component of privacy for all individuals.[43] Theory of intellectual privacy makes a normative claim that the expression of thought and belief should be unrestricted[44] and has been recognised by the Preamble to the Indian Constitution, "*Liberty of thought, expression, belief, faith…*" In his dissenting opinion in *Abrams v. United States*,[45] US Supreme Court Justice O.W. Holmes characterised this condition for free thought as a foundational value of democratic institutions.[46]

(2) 'CHILLING EFFECT'

The imposition of a 'chilling effect' on free speech is guarded against by Article 19(1)(a).[47] The chilling effect caused by surveillance, in general, was explained by Justice Subba Rao in *Kharak Singh*[48] which was subsequently approved in *Privacy* judgement. Justice Subba Rao noted that it is indeed impossible to show actual, tangible harm in case of surveillance:

"The freedom of movement in clause (d) therefore must be a movement in a free country, i.e., in a country where he can do whatever he likes, speak to whomsoever he wants, meet people of his own choice without any apprehension, subject of course to the law of social control. The petitioner under the shadow of surveillance is certainly deprived of this freedom. He can move physically, but he cannot do so freely, for all his activities are watched and noted. The shroud of surveillance cast upon him perforce engender inhibitions in him and he cannot act freely as he would like to do. We would, therefore, hold that the entire Regulation 236 offends also Art. 19(1)(d) of the Constitution."

Surveillance may also be used to suppress exchange of politically unpopular or unorthodox opinions by individuals and groups as these groups cognizant of surveillance may exercise self-censorship and may not express their opinions. This is against the very culture of conferment of fundamental rights by the Constitution which aims at freedom of speech and expression, to express oneself without being under the threat of any fear. In *NAACP v. Alabama*,[49] the US Supreme Court struck down compelled disclosure of the membership lists of a civil rights organisation (the NAACP), noting that the knowledge of surveillance would force politically unpopular or dissident individuals and groups into self-censorship.

[43] Neil M. Richards, *The Dangers of Surveillance*, 126 HARV. L. REV. 1934 (2013).
[44] *Id.* at 1940.
[45] Abrams v. United States, 250 U.S. 616 (1919).
[46] *Id.* at 630 (Holmes, J., dissenting).
[47] Aadhar judgement, *supra* note 32, ¶ 87.
[48] Kharak Singh, supra note 23.
[49] NAACP v. Alabama, 357 U.S. 449 (1958).

The doctrine of 'chilling-effect' has been criticised by many scholars on the ground that it overprotects freedom of speech and often results in preventable harm.[50] These criticisms, however, miss the point that surveillance should not be a casual matter to be exercised by the State on frivolous grounds, and surveillance should be allowed, only under procedural safeguards, in exceptional circumstances only. Theory of intellectual privacy posits that a substantial protection should exist against normalising the gauge of surveillance.

V. LIMITS ON RIGHT TO PRIVACY - PERMISSIBLE STATE INFRINGEMENT

Notably, privacy has both negative and positive content –

"*The negative content restrains the State from committing an* [unjustified] *intrusion upon the life and personal liberty of a citizen. Its positive content imposes an obligation on the State to take all necessary measures to protect the privacy of the individual.*"[51]

Like other fundamental rights, right to privacy is privacy is not absolute. Relying upon *Modern Dental College and Research Centre v. State of Madhya Pradesh*,[52] *Gobind v. State of Madhya Pradesh*,[53] *Privacy judgement*,[54] the Supreme Court in *Ritesh Sinha v. State of Uttar Pradesh*[55] held that "*the fundamental right to privacy cannot be construed as absolute and but must bow down to compelling public interest.*"[56] It was held in the *Aadhar judgement* that the right to privacy cannot be impinged without a just, fair and reasonable law.[57] It has been settled in *Maneka Gandhi v. Union of India*, that "*procedure which deals with the modalities of regulating, restricting or even rejecting a fundamental right falling within Article 21 has to be fair, not foolish, carefully designed to effectuate, not to subvert the right itself*".[58] Thus understood, "procedure" must rule out anything arbitrary, freakish or bizarre. A valuable constitutional right can be canalised only by civilised process. Underlining this line of thought, it would be very clear that right to privacy can very much be curtailed, the aspect that needs to be analysed is whether Section 69 positively answers the criteria for curtailing the right to privacy.

(1) GROUNDS FOR LIMITING EXERCISE OF RIGHT TO PRIVACY

[50] Leslie Kendrick, *Speech, Intent and the Chilling Effect*, 54 WM. & MARY L. REV. (2013).

[51] Privacy Judgement, *supra* note 11, ¶ 326.

[52] Modern Dental College, *supra* note 4.

[53] Gobind, *supra* note 1.

[54] Privacy Judgement, *supra* note 11.

[55] Ritesh Sinha v. State of Uttar Pradesh, Criminal Appeal No. 2003 of 2012.

[56] *Id.*

[57] Aadhar Judgement, *supra* note 32, at 158.

[58] Maneka Gandhi v. Union of India, (1978) 2 SCR 621 [hereinafter Maneka Gandhi].

In the *Privacy case*, the Supreme Court laid down the following prerequisites before the State could interfere with privacy:[59]

"*(i) The action must be sanctioned by law;*

(ii) The proposed action must be necessary in a democratic society for a legitimate aim;[60]

(iii) The extent of such interference must be proportionate to the need for such interference;

(iv) There must be procedural guarantees against abuse of such interference."[61]

Surveillance under Section 69 has the sanction of law, i.e. Information Technology Act, 2000, which is enacted by a competent legislature through procedure established under the Constitution. The aim to carry out surveillance, here, is the maintenance of "sovereignty or integrity of India, defence of India, security of the State, friendly relations with foreign States or public order"[62] and can be argued to be necessary in a democratic society as these are *sine qua non* for a democracy to exist; these grounds except 'defence of India' are valid ground under Article 19(2)[63] of the constitution to impose reasonable restriction on right to freedom of speech and expression under Article 19 (1)(a) of the Constitution.[64]

At the same time, it should be noted that mere existence of 'legitimate state interest' does not permit limiting the right to privacy, such action in pursuance of 'legitimate state interest' should not be disproportionate or excessive.[65] In fact, existence of legitimate state interest satisfies only one element of proportionality.[66] This follows that the restriction upon right to privacy should be reasonable.

[59] Privacy Judgement, *supra* note 11.

[60] Privacy Judgement, *supra* note 11. DY Chandrachud, J., provided for this test in the Privacy judgement, ¶ 32:
(i) legality
(ii) legitimate state aim – which should ensure that the stature should fall within the zone of reasonableness as per Article 14
(iii) proportionality
Kanu Agrawal, *Legitimate and Compelling State Interest: The Test for Aadhar*, BAR & BENCH, https://barandbench.com/legitimate-state-interest-test-aadhaar/ (last visited June 8, 2020).

[61] Privacy Judgement, *supra* note 11, ¶ 638.

[62] Information Technology Act, 2000, *supra* note 8, § 69.

[63] Article 19(2), The Constitution of India, 1950. It is essential to note the difference between "security of state" vis-à-vis "defence of India". The term "security of state" refers only to serious and aggravated forms of public order e.g. rebellion, waging war against the State, insurrection and not ordinary breaches of public order and public safety, e.g. unlawful assembly, riot, affray. Thus, speeches or expression on the part of an individual, which incite to or encourage the commission of violent crimes, such as, murder are matters, which would undermine the security of State. "Defence of India" relates more to external aggression, and defence emergency and does not relate to internal public order. It relates to physical boundaries of the country. For a detailed discussion, *Nathuni Lal Gupta* v. *State*, AIR 1964 Cal 279.

[64] INDIA CONST. art. 19(1)(a).

[65] Aadhar Judgement, *supra* note 32, ¶ 283 (Dissent).

[66] Aadhar Judgement, *supra* note 32, ¶ 296.

In *M.R.F. Ltd. v. State of Kerala*,[67] the Supreme Court held that while examining the 'reasonableness' of a statutory provision one has to keep in mind the following factors:

"(1) The directive principles of State policy.

(2) Restrictions must not be arbitrary or of an excessive nature so as to go beyond the requirement of the interest of the general public.

(3) In order to judge the reasonableness of the restrictions, no abstract or general pattern or a fixed principle can be laid down so as to be of universal application and the same will vary from case to case as also with regard to changing conditions, values of human life, social philosophy of the Constitution, prevailing conditions and the surrounding circumstances.

(4) A just balance has to be struck between the restrictions imposed and the social control envisaged by Article 19(6).

(5) Prevailing social values as also social needs which are intended to be satisfied by the restrictions.

(6) There must be a direct and proximate nexus or reasonable connection between the restrictions imposed and the object sought to be achieved. If there is a direct nexus between the restrictions, and the object of the Act, then a strong presumption in favour of the constitutionality of the Act will naturally arise."[68]

Nothing in the Directive Principles of State policy prevents the state from conducting surveillance, further the prevailing social values allow for surveillance in interest of the society - "sovereignty or integrity of India, defence of India, security of the State, friendly relations with foreign States or public order"[69] – and it indeed has direct nexus with the object sought to be achieved, i.e. gathering information secretly is in the aforementioned interests.

The limitations upon the right to privacy also vary depending upon the source it is traced to, Article 19 or 21, "*if a privacy claim specifically flows only from one of the expressly enumerated provisions under Article 19, then the standard of review would be as expressly provided under Article 19. However, the possibility of a privacy claim being entirely traceable to rights other than Article 21 is bleak. Without discounting that possibility, it needs to be noted that Article 21 is the bedrock of the privacy guarantee. If the spirit of liberty permeates every claim of privacy, it is difficult, if not impossible, to imagine that any standard of limitation other than the one under Article 21 applies.*"[70] However, proportionality under Section 69 and rules framed under Section 69 (2) need to be looked into and can be analysed through the principle of 'compelling state interest' and 'narrow tailoring'. Further, the procedure adopted in surveillance

[67] M.R.F. Ltd. v. State of Kerala, (1998) 8 SCC 227.
[68] *Id.* ¶ 9.
[69] Information Technology Act, 2000, *supra* note 8, § 69.
[70] Privacy Judgement, *supra* note 32, ¶ 379.

also needs to be scrutinised to gauge whether it amounts to 'just, fair and reasonable' procedure under Article 21 of the Constitution.[71]

(2) 'COMPELLING' STATE INTEREST & NARROW TAILORING

The doctrine of "compelling state interest"[72] originated in USA in *Skinner v. Oklahoma*,[73] and *Korematsu v. United States*[74] and has been subsequently applied throughout the world in different jurisdictions. It entails a strict standard of scrutiny comprising two things—a "compelling State interest" and a requirement of "narrow tailoring" (narrow tailoring means that the law must be narrowly framed to achieve the objective). It should be noted that as a term, "compelling State interest" does not have definite contours in the USA. In Indian context, 'compelling state interest' leads to a situation where judicial scrutiny is stricter than 'fair, just and reasonable' test.[75]

"*Hence, it is critical that this standard be adopted with some clarity as to when and in what types of privacy claims it is to be used. Only in privacy claims which deserve the strictest scrutiny is the standard of compelling State interest to be used. As for others, the just, fair and reasonable standard under Article 21 will apply. When the compelling State interest standard is to be employed, must depend upon the context of concrete cases.*"[76]

This is primarily because *"the reasonable expectation of privacy may vary from the intimate zone to the private zone and from the private zone to the public arena.*"[77] "Narrow tailoring"[78] requires that the state action should infringe the fundamental right in a minimum possible manner to achieve its legitimate aim. This was laid down by the US Supreme Court in *Grutter v. Bollinger*[79] where it held that:

"*Even in the limited circumstance when drawing racial distinctions is permissible to further a compelling state interest, government is still constrained under equal protection clause in how it may pursue that end: the means chosen to accomplish the government's asserted purpose must be specifically and narrowly framed to accomplish that purpose.*"[80]

[71] Maneka Gandhi, *supra* note 58.

[72] Stephen A. Siegel, *The Origin of the Compelling State Interest Test and Strict Scrutiny*, 48 AMERICAN JOURNAL OF LEGAL HISTORY 355 [hereinafter Siegel].

[73] Skinner v. Oklahoma, 316 U.S. 535, 541 (1942).

[74] Korematsu v. United States, 323 U.S. 214, 216 (1944).

[75] Maneka Gandhi, *supra* note 58.

[76] Privacy Judgement, *supra* note 11, ¶ 380.

[77] Aadhar Judgement, *supra* note 32, ¶ 90.

[78] Siegel, *supra* note 72; Gautam Bhatia, *Surveillance and Privacy in India – II: Gobind and the Compelling State Interest Test*, INDIAN CONSTITUTIONAL LAW AND PHILOSOPHY, (Dec. 16, 2013), https://indconlawphil.wordpress.com/2013/12/16/surveillance-and-privacy-in-india-ii-gobind-and-the-compelling-state-interest-test/ (last visited June 8, 2020).

[79] Grutter v. Bollinger, 538 U.S. 306 (2003).

[80] *Id.*

Therefore, what follows from 'compelling state interest' is that surveillance *per se* is not unconstitutional but the way in which it is conducted makes it unlawful, opaque and arbitrary.[81] Procedural safeguards and lack of judicial oversight enhances the chilling effect of surveillance on society and brings it on the brink of unconstitutionality. This is not covered by the restrictions in surveillance laws due to lack of existence of a proper statutory framework and process, it is unreasonable because of the vagueness of the legal backing behind such actions and the overbreadth of such laws.

In the current context, unbridled surveillance is against the grounds of reasonability and vagueness as the overbreadth principle kicks in. Mass surveillance, has no safeguards such as scrutiny by a judicial authority, while in a similar light, countries like the US, require a 'probable cause'[82] standard which is essentially judicial review of surveillance, providing a system of checks and balances to such an action, which aligns with the reasoning of the Court in *PUCL* with requires a 'case-to-case basis' approach for surveillance. This would ensure that infringement of privacy would be minimum as an effective check could be kept upon executive and subsequently, executive would not be able to subject anyone to surveillance thereby ensuring 'narrow tailoring' is adequately followed.

VI. BALANCING LEGITIMATE STATE INTERESTS AND THE RIGHT TO PRIVACY

IT Act Rules framed for the purposes of Section 69 are vague in themselves as they do not provide for what shall be the scope of the power of the state official, or what kind of data will be aimed at, or whether the citizen will be made aware of the data collection. Framed more than a decade ago, IT Act Rules do not offer adequate protection against modern surveillance tools like facial recognition, geo tagging, encrypted data and conversations. These rules also do not outline safeguards against illegal surveillance or their misuse; and do not create a case-to-case distinction and operate under general scheme for all cases. Hence, Section 69 and rules framed thereunder fall foul of the 'fair, just and reasonable' test let alone 'narrow tailoring'.

The framework is vague and confers extraordinary powers in the hands of the executive which makes the law 'overbroad'. This means that even though they are aimed for a specific purpose,

[81] Danielle K. Citron, *Addressing the Harm of Total Surveillance: A Reply to Professor Neil Richards,* 126 HLRF 262 (2013), https://harvardlawreview.org/2013/06/addressing-the-harm-of-total-surveillance-a-reply-to-professor-neil-richards/ (last visited June 8, 2020).

[82] Chaitanya Ramcahndran, *PUCL v Union of India Revisited: Why India's Surveillance Law Must Be Redisigned for the Digital Age*, 7 NUJS L. REV. 105, 118 (2014).

i.e. grounds under Section 69, they cover aspects more than that.[83] Consequently, it impinges upon the freedom of speech beyond the stipulated restrictions.[84] The overbreadth of provisions under Section 69 is very evident from the fact that not only is 'harmful' speech covered under grounds enumerated in Section 69 but it also casts a chilling effect on speech which are not 'harmful' in nature and thereby attracting overbreadth principle.[85] In such a scenario, it is liable to be struck down as overbreadth of a statute impinging upon freedom of speech is a ground for declaring a law unconstitutional.[86]

In cases involving disclosure of information, it is imperative that judicial review should exist for scrutinising such executive action. In the absence of such judicial (independent) oversight, the provisions permitting such action are liable to be struck down.[87] The executive cannot be trusted with this task alone because they bring an inherent "institutional bias"[88] in their functioning which is often aimed at collecting data rather than protecting it. Further, the absence of penalties for abuse of surveillance powers offers no incentive (or reason to fear) for the Executive to adhere to these procedural safeguards. Absence of any legal framework for discovering unauthorised and unwarranted surveillance, until the breach is made public, further compounds the problem. Courts may even find such accusations of surveillance "too speculative".[89] In such a scenario where unwarranted executive action cannot be detected, the executive has absolute immunity in carrying out surveillance in the manner they wish to and there is an uncontrolled violation of fundamental rights.

A person may have recourse to following remedies in case of surveillance:

a) initiate a civil suit for a tortious claim;[90]

b) a writ petition under Article 32 or 226 of the Constitution for claiming compensation;[91]

c) action for criminal trespass subject to the provisions of Code of Criminal Procedure;[92]

[83] Information Technology Act, 2000, *supra* note 8, § 69.

[84] Shreya Singhal, *supra* note 31.

[85] Lewis Sargentich, *The Overbreadth Doctrine*, 83 HARV. L. REV. 844, 853 (1970).

[86] Shreya Singhal, *supra* note 31, ¶ 90.

[87] Aadhar Judgement, *supra* note 32. Justice BN Srikrishna Committee specifically noted that the lack of legislative/statutory inter-branch oversight in India was "not just a gap that is deleterious in practice but, post the judgment of the Supreme Court in *Privacy Judgement*, potentially unconstitutional" since it did not satisfy the tests of *Privacy Judgement*.

[88] M. P. JAIN & S. N. JAIN, PRINCIPLES OF ADMINISTRATIVE LAW 225-234 (2002) [hereinafter JAIN]; Romesh Sharma v. State of Jammu & Kashmir, (2007) 1 J.K.J. 84.

[89] Clapper v. Amnesty International USA, 113 S. Ct. 1138, 1147 (2013) (U.S.).

[90] Sunkara Satyanarayana v. State of Andhra Pradesh, 2000 (1) A.L.D. (Cri.) 117, ¶ 65 (Per V. V. S. Rao, J.) (recourse available to an individual for infringement of right to privacy).

[91] *Id.* ¶ 65.

[92] *Id.* ¶ 65.

d) judicial review of the constitutional courts under Article 32 or 226 of the constitution;[93]

Though, these recourses may look very attractive at first blush, they often take substantial resources, money, effort, time, and lawyering to enforce to the right to privacy[94] and hence, will be rather ineffective in adequately addressing the problem.[95] Further, the burden of proof on the plaintiff complicates the problem as s/he has no way in which s/he can prove that s/he is put under surveillance.

Ex-ante ex-parte judicial permission is the best way to adequately address the interest of state and that of citizens as well. Before the surveillance could be carried out by the State, the executive shall present the 'probable cause' to judges, who may then permit such surveillance. This way the person who would be put under surveillance would not be alerted before at hand and all concerns regarding unfair surveillance will also be resolved. In cases which demand instant action, surveillance may be carried out and later be ratified by the judge. Thus, 'case-to-case' basis of surveillance would also be secured. Prior judicial scrutiny – *ex parte* in nature – was suggested in *PUCL* case.[96] However, the Supreme Court felt that it would amount to 'judicial legislation' and turned down the suggestion on the ground that it was not possible for the judiciary to insert *ex-parte* judicial scrutiny in the absence of any provision in the IT Act.[97] Now, *Aadhar Judgement* has set a precedent where prior judicial scrutiny is *sine qua non* for carrying out such exercise.[98]

CONCLUSION

The problem that state surveillance poses in our daily lives is both unique and acute as India faces a long-standing threat of terror where the right to privacy may seem like a rather new creature. In such a scenario, surveillance laws and rules framed thereunder seem 'archaic' and are not in line with recent developments of privacy jurisprudence and fail to offer any real protection to citizens against unjust executive action, which has its own institutional bias.[99] Not denying the fact that the state indeed has legitimate interest in subjecting individuals and groups to surveillance, unfettered and unconstrained surveillance without adequate procedural safeguards poses a real

[93] *Id.* ¶ 65.

[94] ARUN MOHAN, JUSTICE, COURTS AND DELAY 1-42 (2009); Marc Galanter, *Fifty Years On, in* SUPREME BUT NOT INFALLIBLE: ESSAYS IN HONOUR OF THE SUPREME COURT OF INDIA 57-65 (B. N. Kirpal et al. eds., 2004) (Galanter describes litigation in India as a game of a 'sunken cost auction').

[95] Apar Gupta, *Balancing Online Privacy in India*, 6 INDIAN J. L. & TECH. 43, 60 (2010).

[96] PUCL, *supra* note 26, ¶¶ 8, 33.

[97] *Id.* ¶ 34.

[98] Aadhar Judgement, *supra* note 32.

[99] JAIN, *supra* note 88, at 225-234.

threat to the sacrosanct right to privacy. It must be remembered, "*that the Constitution is not a mere parchment; it derives its strength from the ideals and values enshrined in it. However, it is only when we adhere to constitutionalism as the supreme creed and faith and develop a constitutional culture to protect the fundamental rights of an individual that we can preserve and strengthen the values of our compassionate Constitution*."[100]

Unbridled surveillance without due safeguard has inherent problems – it affects 'intellectual privacy' which forms the bedrock of the intellectual development of a democratic nation; it also 'chills' the exercise of the right to freedom of speech and expression where people change their behaviour because of potential prosecution. These concerns would be addressed if judges scrutinise surveillance orders *ex-parte ex-ante*, where the executive today in an emergency is empowered to undertake surveillance *with ex-post* approval. This would rule out institutional bias of the executive and ensure that secrecy is maintained while preventing unjust surveillance and protecting the right to privacy.

[100] Navtej Johar v. Union of India, (2018) 10 SCC 1, ¶ 110.

POST-HUMANISM AND CITIZENSHIP: AN ANALYSIS OF THE CAPACITY AND CONSEQUENCES OF CONFERRING NATIONAL CITIZENSHIP TO ARTIFICIAL INTELLIGENCE IN A DEMOCRATIC SOCIETY

- *Anubhav Banerjee*[*]

"It was the slave's continuing desire for recognition that was the motor which propelled history forward, not the idle complacency and unchanging self-identity of the master."

~ Francis Fukuyama[1]

Rights and duties form the cornerstone of the modern globalized world, and the foundations of modern society has emerged through an inherent belief in the rights held by human beings by virtue of them being humans. From the very origins of collective existence, various groups of human beings have been conferring an array of titles, statuses and degrees of acceptance and recognition, whether it was founded on the religious merits of a person's life based upon paradigms of spirituality and piety,[2] or on their place of birth and origin.[3] The Hellenic culture of the Greek city states can be attributed as the precursor to such a bestowal of status, as merely being culturally Hellenic, whether they were Sicelotes, Thracians, Seleucids or mainland Greeks, was sufficient to allow inclusion of such persons into an entire society which accorded special statuses and privileges to those deemed worthy to satisfy its prerequisites. The right to participate in the Olympic games, for example, which was an event transcending modern sporting events through its religious and political significance; was formerly a right only granted to men of Hellenic culture irrespective of their political origin. Greek philosophy was projected as an exclusive branch of cognitive functionality only really appreciated by Greeks, and if modern insights of Greek philosophy are any indication, such as Derrida's examination of the Greek term *pharmakon*, which encompasses a*gatha* (good) and *aniara* (pain) to indicate both 'cure' and 'poison';[4] there was a great deal of sophistry and prestige associated with the Hellenic identity, and Greek citizenship was a badge of honor held against those unworthy of its merits and values.

[*] Student, School of Law, Christ University, Bangalore.

[1] FRANCIS FUKUYAMA, THE END OF HISTORY AND THE LAST MAN (The Free Press, 1992).

[2] M.D. Altekar*, Caste System and Its Relation to Social and Economic* Life, 145 AAPSS 183-87 (1929).

[3] P. Wooyeal & D. Bell, *Citizenship and State-Sponsored Physical Education: Ancient Greece and Ancient China*, 66 REV. OF POLITICS 7-34 (2004).

[4] D. Zeyl, *Socrates and Hedonism: "Protagoras" 351b-358d,* 25 PHRONESIS 250, 255 (1980).

The paradigms of citizenship did not change from conferral of privilege to the selected few as against the "other", but these binaries were reformed to reflect a minimum "civilized" treatment of all humans. Such a bifurcation existed in Roman law through *Ius Naturale* (the rights of all persons and the privileges of Romans) and *Ius Gentium* (the rights of foreigners and of the conduct of nations)[5] system, which reveals how the foundational principles of citizenship and identity evolved; from the conferral of one-sided benefits towards accepted 'citizens'; into a more equitable and basic common system that ensured basic rights of all persons, followed by *additional* benefits granted to the selected few citizens. Such a shift allowed for the preservation of social and demographic privileges, while also encouraging immigration of people who desired the benefits given to the peoples on the periphery, rather than remain in societies which completely excluded those not deemed worthy enough to be citizens. The meteoric population growth of the city of Rome can be attributed to the massive scale of migration[6] from Italy and Roman colonies towards the capital of the empire, with many aspiring to eventually become Roman citizens and enjoy the benefits of *Pax Romana* (or the stability and power of the Roman peace). Such was not merely related to status, as the relative benefits of Universal recognition in societies allowed economic and social homogenization and development and potentially even an aspiration of ascending to the status of citizen over time; rather than the abject deprivation of the older modalities of citizenship.

It would be reductive to entirely credit the developments in the Classical period to the concepts of rights and citizenship, for the turbulence and transformations of the Pre-Modern Era through the Industrial Revolution, and the Age of Enlightenment led to the growth of agency of citizens to become more than mere supplicants to monarchial and religious power. The rights of man were regarded as self-evident[7] and principally universal, and the role played by the man and citizen expanded from being powerless subjects of monarchs to active and participatory members of civic society post the French Revolution. Through the interpolation of anti-racism movements, the growth of Feminism and the Suffragette movement, and the eventual rise of Post-Colonialism and the rights to national self-determination, the notion of passive and incapable citizens was discarded. While voting rights and the exact nature of privileges differed across nations due to geopolitical as well as regional variables, the idea of citizens being tethered

[5] HENRY SUMNER MAINE ET AL., ANCIENT LAW, ITS CONNECTION WITH THE EARLY HISTORY OF SOCIETY AND ITS RELATION TO MODERN IDEAS 67 (Henry Holt and Company 1906).

[6] Walter Scheidel, *Roman population size: The Logic of the Debate*, 2 Princeton/Stanford Working Papers (2007), https://www.princeton.edu/~pswpc/pdfs/scheidel/070706.pdf. (last accessed June 2, 2020).

[7] *The Declaration of Independence* (U.S. 1776) [hereinafter The Declaration of Independence].

to the State through their own willingness to be represented and protected by a society of their own became the next step in the evolution of citizenship. Humans acquiring the privileges of citizenship and the benefits of living as a recognized member of a commune was the primary driving force behind urbanization[8] in a manner described by the Social Contract theory;[9] which eventually led to a pursuit of the needs encapsulated within Maslow's Hierarchy of Prepotency.[10] In the pursuit of self-actualization, humanity strove to satisfy their political and belongingness needs, and in this process transformed from the feudalistic monarchies of the Medieval ages, created to guarantee food and safety, into representative democracies to attain belongingness and representation needs, where they were recognized as key aspects of the geo-political sphere and could be actively involved in the creation of governments and the functionality of the state.

However, a key variable in the development of this form of jurisprudence that has remained largely static has been the anthropocentric philosophy of citizenship. Non humans remain largely confined as subjects of law,[11] and are not equated to be citizens of a state, though personality can be attributed to them in varying degrees.[12] Therefore, it is quite analogous to infer that in the process of attribution of citizenship itself, there needs to be a clear manifestation of humanlike personality and personhood which hitherto remains the monopoly of the *homo sapiens*. Personality by itself is no guarantee of citizenship, as the 70.8 million displaced and stateless refugees globally will attest,[13] which means that the status of citizenship encompasses certain principles and values that emerge as a result of humanity. Scholarly insights from jurisprudence[14] suggests that there lies two effective aspects interwoven into the legal and philosophical concept of citizenship; the first is a *"functional"* aspect, or the legal relationship between the individual and the state, making "alienage distinctions", the differences between the rights and duties of earning indigenous citizens as what separates them from "aliens". The second aspect is *"nonfunctional"* based upon an ethereal attachment to commonalities in identity and community are held collectively and

[8] EDWARD ANTHONY WRIGLEY ET AL., ENGLISH POPULATION HISTORY FROM FAMILY RECONSTITUTION 1580-1837 295-96, 303 (Cambridge University Press 1931); Manfred Heun, Ralf Schäfer-Pregl, Dieter Klawan, Renato Castagna, Monica Accerbi, Basilio Borghi & Francesco Salamini, *Site of Einkorn Wheat Domestication Identified by DNA Fingerprints* 278 AAAS 1312, 1312-14 (1997).

[9] THOMAS HOBBES, THE LEVIATHAN (Penguin Books Chicago 1968); JOHN LOCKE, TWO TREATISES ON CIVIL GOVERNMENT (G. Routledge 1887); JEAN-JACQUES ROSSEAU ET AL. SOCIAL CONTRACT: AND DISCOURSES (E.P. Dutton and Company Inc. 1950).

[10] A.H. Maslow, *A Theory of Human Motivation*, 50 PSYCHOLOGY REVIEW 370-396 (1943).

[11] P.J. FITZGERALD, SALMOND ON JURISPRUDENCE 298-299 (12th ed. 2016).

[12] Pramatha Nath Mullick v. Pradyumna Kumar Mullick, (1925) 27 BOMLR 1064.

[13] United Nations High Commissioner for Refugees, *Global Trends of Forced Displacement*, UNHCR (June 20, 2019), https://www.unhcr.org/5d08d7ee7.pdf. (last accessed Sept. 8, 2019).

[14] Mark C. Fleming, *The Functionality of Citizenship*, JEAN MONNET PROGRAM, https://jeanmonnetprogram.org/archive/papers/97/97-10-Contents.html#top_of_page. (last accessed Sept. 8, 2019).

simultaneously by all members of a nation. These principles work in cohesion, with neither being more important than the other. Valuing functionality alone commodifies the value of citizenship, legitimizing any financial and tangible gains made by any person as the *sole determinant* of their right to be a citizen, which disqualifies dependents on state welfare as citizens. Similarly, a purely non-functional approach renders an exclusivity to *certain families* within certain communities and breaches the principle of rule of law.[15] Hereditary entitlements being the *sole determinant* entirely robs more qualified immigrants the opportunity to eventually integrate themselves and their families into the society as prospective citizens.

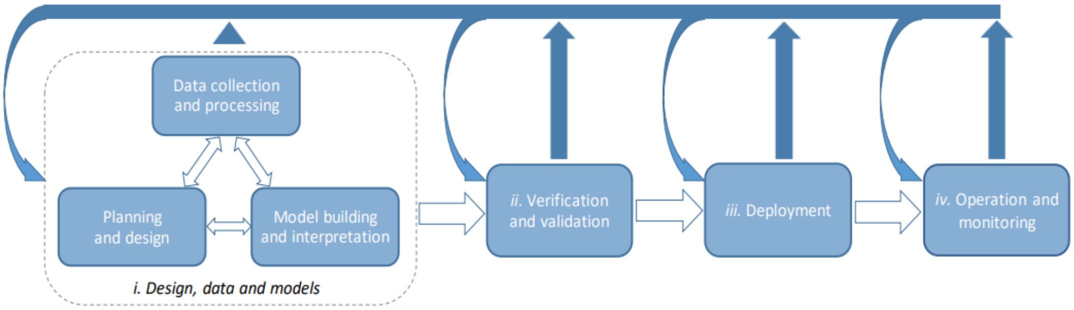
Through the course of this essay, I aim to examine whether the existing jurisprudence and legal systems of acquiring citizenship are sufficient to first determine whether A.I.s are capable of being citizens; then how they would feasibly become citizens; and what are the consequences of conferring such citizenship. "A.I." itself remains undefined, and can be loosely understood as "*machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. It uses machine and/or human-based inputs to perceive real and/or virtual environments; abstract such perceptions into models (in an automated manner e.g. with ML or manually); and use model inference to formulate options for information or action. AI systems are designed to operate with varying levels of autonomy.*"[16] A.I.s "learn" via Machine Learning systems, and use the human neural networks as the prototype for their structure and processes, and are capable of responsive as well as strategic thinking, and are also able to respond via audio-visual cues to mimic human expressions.

Thereafter, a natural comparison between human cognitive processes and A.I.s reveals a commonality in the methodologies of thinking and the behaviors exhibited during computerized artificial neural processing, upon reference to the research done by psychologists like Pavlov and Skinner.[17] Just as human beings learn by association (classical conditioning) and responsive learning (operant conditioning), A.I. can use the same methodologies while learning to mimic human learning, and then transcend its limitations through advanced machine learning. There are many sources of inputs such as the family, education system and society which help humans learn abstract ideas which communities of people can perceive collectively and be taught to obey[18] like

[15] A.V. DICEY, INTRODUCTION TO THE CONCEPT OF THE STUDY OF THE CONSTITUTION (4th ed. 1982).
[16] OECD, ARTIFICIAL INTELLIGENCE IN SOCIETY 15 (OECD Publishing 2019) [hereinafter OECD].
[17] I.P. PAVLOV, CONDITION REFLEXES (*Dover Publications 1960);* R. G. MILTENBERGER, BEHAVIORAL MODIFICATION: PRINCIPLES AND PROCEDURES *9 (5th ed. 2008).*
[18] Y.S. HARARI, HOMO DEUS: A BRIEF HISTORY OF TOMORROW (Harvill Secker 2016).

respect for law, appreciation of culture, acknowledgment of religious and social differences. This is supported by the work of Carl Jung and his theory of Collective Unconscious,[19] which manifests in the legal fraternity through texts like Universal Declaration of Human Rights, *Jus Cogens* norms in Customary International Law, and commonly practiced principles like the rule of law.

Therefore, it is technologically feasible to collect these collective cognitive precepts and beliefs, and these Universally accepted practices, customs and codes; all of which can be combined into a *corpus* of laws in machine language, which can serve as inviolable rules that an A.I. can follow. Because there are commonalities between legal systems despite technical differences, it is possible to make an A.I. able to understand and obey with the logic underlying law and legal practice, and so long as the programmers have a definite series of laws and principles they can encode as laws to the A.I., they can make such an entity cognizant of and compliant to laws and morals in society.


Data collection and processing
Planning and design
Model building and interpretation
i. Design, data and models
ii. Verification and validation
iii. Deployment
iv. Operation and monitoring

Source: As defined and approved by AIGO in February 2019.

Figure 1: System Cycle of Artificial Intelligence[20]

The work of OECD's AI Group of Experts (AIGO) in contextualizing and expanding the horizons of the theoretical possibilities and applications of Artificial Intelligence systems has amply reinforced the technical capacity of A.I. to be able to understand and respond to laws and legal systems. Citizenship, therefore, becomes another concept that can be encoded as a legal precept, whose technical criteria and obligations can be easily programmable as variables in the Design and Verification stages of programming an A.I. system. In fact, such a process of

[19] D. JOHN DOYLE, WHAT DOES IT MEAN TO BE HUMAN? LIFE, DEATH, PERSONHOOD AND THE TRANSHUMANIST MOVEMENT 173 (Springer 2018).
[20] OECD, *supra* note 16.

programming is a manifestation of Skinner's psychological research on Operant Conditioning;[21] which can already be applied to law, as there are positive reinforcements for compliance to the mandate of law, and negative reinforcements against any behaviors that bypass or violate law. Just as how law conditions aversive and acceptable behaviors in people, A.I.s can be taught to comply with the application of laws and legal principles just as it obeys the limitations and mandates of its coding. In fact, the process of coding these behaviors into an A.I.'s algorithm is similar to the precepts of natal conception and conditioning, for there are many similarities between the genetic code of people in the DNA of human beings and those of A.I. program codes; therefore making an interpolation of law, morality and psychology into an A.I.'s code quite analogous to conditioning human behavior.

Of course, dissent and "independent rational thought" are attributes correlated to the legal profession which remains the exclusive domain of certain humans (judges), and there is a substantial juristic and jurisprudential value attributed to the *spirit* of the law, which refers to the inherent logic of a law beyond its textual confines.[22] The philosophical value of judicial pronouncements and legislations are the exclusive domain of the judiciary in a state; who are deemed to truly understand all aspects of a law, and entrusted with interpreting and clarifying the real value of law.[23] This is, however not a detriment to the ability of an entity to comprehend and apply laws, because the corollary of this exclusive domain of the judiciary is that such a value is not to be interpreted or deciphered by the rest of the society, which is expected to comply with laws and procedures without specializing in its deeper meanings. Juxtaposing an A.I. as an average intelligent voter shows how the lack of such specialized knowledge does not detract from its ability to comply with the essence and effect of law; because a citizen is not expected to know the nuances of a law in order to obey or respect it, and similarly an A.I. need not be programmed to understand this deeper meaning of law. This assumption itself is growing obsolete, as contemporary legal practice seeks to involve all stakeholders beyond just the state and its institutions in changing and developing law, which includes private persons and artificial persons like corporations and international organizations.[24]

[21] B.F. SKINNER, THE BEHAVIOR OF ORGANISMS: AN EXPERIMENTAL ANALYSIS (Appleton-Century 1938).

[22] John M. Gaus, *The Report of the British Committee on Ministers' Power*, 26 AM. POLITICAL SCI. REV. 1142, 1143-1147 (1932).

[23] Keshav Mills Company Petlad v. CIT, Bombay-North Ahmedabad, AIR 1965 SC 1636.

[24] OHCHR, *Guiding Principles on Business and Human Rights* (United Nations Human Rights Office of the High Commissioner 2011).

Therefore, there is a need to accept the capability of A.I. to participate in ensuring the continuation of the rule of law. Even in the worst case scenario, when an A.I. needs to be constantly supervised until it can be trusted to have understood the effect, provisions and principles of law; until the A.I. demonstratively understands how the law actually applies and exercise proper consent and compliance, it can exist in the same domain as minors, persons of unsound mind, and other similar subjects of legal systems, which can act as a failsafe during the Monitoring stage of development, making the programmers liable for the actions of the A.I. and any other entity it exercises control over. In fact, A.I. can be programmed to circumvent racial or sexist bias when applying and operating under legal parameters, and such entities would be far less likely to perpetuate stigmas and discriminatory practices than a human being,[25] because it is not cognitively apparent for them to accept the stereotypes and bias and thus they either ignore them entirely, or apply the Veil of Ignorance when making decisions[26] by disregarding demographic statuses altogether. Additionally A.I. can perform supplementary functions to law enforcement with a high degree of success, such as Risk assessment algorithms predicting the risk levels of possible crimes or repeat offences;[27] making it clear that contemporary A.I. can both understand the textual nuances of law, but also integrate variables, facts and circumstances to modify its application of laws, and irrespective of whether it truly "understands" the "deeper jurisprudential value" of laws, it is able to decipher, apply and comply with them.

Therefore, there is a growing acceptance of the capabilities of A.I. to be a person in law, and exist as a person who is capable of obeying and understanding laws and abiding by the practices of society. Since it is already settled that humanity or anthropomorphic characteristics is not the sole determinant of citizenship, the author now seeks to identify whether it is possible for A.I. to acquire citizenship through the various methods of acquisition available to human beings. The *prima facie* processes of citizenship, however, are not sufficiently equipped to bestow citizenship to such entities. When examining determination by the principle of *jus sanguinis* citizenship,[28] there are complications in the process of acquisition of citizenship rights. The "blood relatives" of an A.I. can be its programmers, unless their composition are members from various nations, in which case there is no formula for determining which nationality the A.I. will belong to. Additionally, any software and hardware components may be owned by artificial persons in turn,

[25] OECD, *supra* note 16.

[26] JOHN RAWLS, A THEORY OF JUSTICE (Belknap Press of Harvard University Press, 1971).

[27] OECD *supra* note 16.

[28] DORA KOSTAKOPOULOU, THE FUTURE GOVERNANCE OF CITIZENSHIP (Cambridge University Press. 2008) 26-27.

like the Government of the State or an incorporated entity, who are generally held vicariously liable for the acts of their agents[29] and hold the title of ownership. Since these programmers work under the instructions of their employers, an equally valid argument can be made that these entities are the true guardians of any fledgling or nascent A.I., which makes it difficult to determine paternity and blood as the basis for acquisition of citizenship.

Furthermore, using the geographical region and territory of conception as a factor of determination of citizenship is equally difficult with regards to A.I.; as it is difficult to ascertain the territory of birth for an A.I., as their conception in specialized software and digital space doesn't correspond to a tangible geographical location where the A.I. is "conceived in", unless the location of the hardware is what is accepted, which reduces an A.I. to its physical components, unlike human beings who are deemed to exist beyond their corporeal existence. Since many such A.I.s use cyberspace as the platform for communication across the internet or other such networks, they also exist and operate in this distinct incorporeal digital domain that is unowned and largely unregulated[30], which further vitiates the rate of naturalization of the A.I. over an extensive period of "residence". Since it is impossible to limit the location where an A.I. resides and operates, it may simultaneously be a resident of many countries and demarcated jurisdictions, making the territoriality of citizenship difficult to enforce or attain.

Therefore, the author is unsure how conventional processes will legitimize the process of acquisition of citizenship, which means that to further examine the potential of A.I. citizens in the future, there needs to be a substantial jurisprudential growth and legal flexibility in understanding how a citizen is recognized. Perhaps the indoctrination of Sophia as a citizen of Saudi Arabia will serve as the template for all future A.I.s where an express recognition will be granted by a government based upon claims filed by the creators of such A.I., for the breath of citizenship can be infused into any entity based upon its *form rather than its substance, of lucid and compendious expression rather than of legal principle.*[31]

The effects of citizenship are not limited to mere recognition, for most societies confer additional rights, privileges and impose greater burdens on those who become citizens. Once the law

[29] State of Rajasthan v. Vidhyawati, AIR 1962 SC 933; Indian Companies Act, 2013, §§ 166-167, No. 18, Acts of Parliament, 2013 (India).

[30] Mark Graham, *Geography/internet: Ethereal alternate dimensions of cyberspace or grounded augmented realities?*, 179 G. J. 177, 178-182 (2013).

[31] OECD, *supra* note 16.

recognizes the capacity of an entity to be a citizen, and lawfully confers the status of citizenship, the entity in question now acquires a host of enforceable claims, rights and civic duties which it is obligated to perform, such as the right to acquire property and the duty to pay taxes. Firstly, there would be a guarantee of Fundamental Rights and the application of all basic rights accorded to human beings to A.I., such as the Right to Life; and given the developments around the Right to life including the right to a dignified existence,[32] it is difficult to understand what would equate to a life of dignity for an A.I., which may need access to electricity and information, or access to a corporeal body it can control for it does not benefit from conventional governmental welfare schemes for it does not need conventional sustenance. There already exist many recommendations given by expert bodies[33] which can be converted into laws for the protection of AI, such as protection from 'Adversarial examples' or malicious operations that convolutes the processing functions of an A.I.; and by making A.I.s citizens, the state is under an increasing pressure to enhance and regulate digital and internet laws and improve cybersecurity just as frequently as defense due to an even greater emphasis on the so far limited scope present in the Right to Privacy and security of Data, which is a net positive in the rapidly digitalizing society. Another significant consequence of enfranchising A.I. citizens is upon the process of elections and democratic participation, for not only will A.I's have the Right to vote as citizens (without which there would again arise problems such as taxation without representation);[34] the nature of political discourse in a state would transform through these additional stakeholders in the electorate. While multifaceted communities like India would find it easier to adapt to another demography of voters, homogenous cultures like Japan or Denmark would need radical reforms in the nature of political appeals and popular discourse so as to become less centered around one community.

Furthermore, almost all democratic societies would completely transform, as community politics may theoretically cross digital limitations and transform into speciesism beyond its current use as the right to exploit plant and animal life[35] (humans having greater prerogatives and rights than the rest of the lifeforms on Earth); and the anthropocentric nature of politics would become irrelevant due to non-human voters. If electoral rights are denied to A.I., and they continue to exist as subservient yet cognitively and potentially equal to human beings, history is replete with

[32] Maneka Gandhi v. Union of India, AIR 1978 SC 597.

[33] OECD, s*upra* note 16.

[34] The Declaration of Independence, *supra* note 7.

[35] George Yancy, Peter Singer, *Peter Singer: On Racism, Animal Rights and Human Rights*, THE NEW YORK TIMES (Aug. 19, 2019, 11:04 AM), https://opinionator.blogs.nytimes.com/2015/05/27/peter-singer-on-speciesism-and-racism/?mtrref=www.google.com&assetType=REGIWALL.

examples of how social upheavals become inevitable, such as the Civil Rights Movement or the Anti-Imperialism sentiments of the early 20th Century against unequal treatment and a lack of representation. While the author is unsure whether the politicians of tomorrow would have to cater to the demands of autonomous A.I. in their campaigns, the need to transform the legal and political systems will become inevitable if there remains a desire to maintain democratic harmony between people and A.I. citizens. The development of A.I. and technology will gradually make such sentient entities a reality, and as human beings have evolved to expect political representation and agency in their respective processes of self-actualization, to deny humanlike A.I. having equal or greater intelligence to humans the same right would be baseless.

In addition to political rights, the right to serve in public office is a benefit only available to citizens, subject to eligibility criteria and professional standards serving as reasonable restrictions on the Right to Occupation of public offices.[36] If A.I.s become citizens in such a civil society, as an extension of their rights to representation and political enfranchisement, there is a substantial possibility of demands for participation and appointment of their kin in public posts. Such a claim is not dubious in and of itself, for it stems from the same desire as the right to vote, namely, the desire to feel involved and engaged in geopolitical and regional discourse. The complications emerge when there are multiple standards of evaluation for the same position, and given the discourse around reservation in India and the privileges granted herein,[37] the best-case scenario will be a feeling of systemic bias by either humans or A.I. if the standards of evaluating them as eligible for a post are different.

However, it is difficult to test both types of entities on the same standards, for it is uncertain whether A.I. would need the same nature of education or academic qualifications as human beings, as their similarities in learning patters is undone by the perfect memory retention capacity of A.I. which encodes its memory in methods that may not decay with age and evolving contexts, even if A.I. exhibits stress responses and anxiety. Notwithstanding the natural judicial skepticism on the reliability of new technologies in law until the passage of time normalizes the innovations themselves,[38] A.I.s may exhibit a higher proficiency than humans in certain functions such as applying the principle of *stare decis*[39] as judges. Therefore, the author is reminded of the

[36] INDIA CONST. art. 19, cl. 6.

[37] Indira Sawhney v. Union of India, AIR 1993 SC 477.

[38] *Florida Innocence Project*, https://www.floridainnocence.org/about (last visited Aug. 18, 2019).

[39] Kimble et al. v. Marvel Entertainment, LLC, Successor to Marvel Enterprises, Inc. 135 S. Ct. 2401 [Justice Samuel Alito's Dissent].

revolutionary steps taken by the Government of Estonia to indoctrinate A.I. into the nuances of civic life, through the "robot judge" at Velsberg[40] and major strides to substitute public servants with A.I. based systems by 2020. Estonia serves as the hub of cybersecurity and research for the NATO, and its prominence in development of a digitalized society is a pragmatic transformation towards the betterment of its citizens. By leading the technological and jurisprudential evolution of A.I. the country is a template for determining the possibilities of such paradigmatic shifts from an anthropocentric socio-political order into one where humanity can co-exist with other entities which can match, supplement and even possibly surpass its intelligence.

Indeed, the author intends to question the assumption of the centrality of human beings to the developments of law and society in the future, because the very existence of artificial beings which have the ability to mimic the human brain, exhibit similar degrees of autonomy and self-awareness, and undergo the same processes of learning and expression of ideas; challenges the legal and social foundations of the contemporary world. The philosophical foundations for this newer world already exist through the numerous developments in Post-humanist thought,[41] and law has already conceded to the lack of a need for actual corporeal human beings to exhibit characteristics of being human. The exclusivity of *homo sapiens* as autonomous, as law-givers, as deliberative, as emotive, social, communicative; have all been captured by artificial persons, and soon will be reflected through A.I. taking a central role in the socio-political spectrum. While A.I. may have been created as a supplement to human labor, its very existence and potential leads to the possibility of a future where human beings are obsolete and unable to perform the same tasks at the same proficiency as A.I., and traces of such a future are already evident.[42]

Unlike previous technological developments and automation, A.I. may serve as a substitute for human inputs altogether, for the physiological limitations of the brain are not shared by A.I., but they both undertake the same processes of cognition and comprehension of information, the brain using biochemical signals and the A.I. processing with electromagnetic connections. Therefore, it is increasingly untenable to dismiss A.I. as mere machines, and the superiority of human beings through speciesism cannot be sustained, and therefore, law must seriously consider

[40] Eric Niiler, *Can AI be a Fair Judge in Court? Estonia Thinks So*, WIRED (Mar. 25, 2019, 7:00 AM), https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/.

[41] F. NIETZSCHE, THUS SPAKE ZARATHUSTRA (Tudor Publishing Co. 1928).

[42] Jane Wakefield, *Robot 'talks' to MPs about future of AI in the classroom*, BBC NEWS (Oct. 16, 2018, 12:24 AM) https://www.bbc.com/news/technology-45879961 (last visited Aug. 18, 2019); Japan Cabinet Office, *Integrated Innovation Strategy*, (2018) https://www8.cao.go.jp/cstp/english/doc/integrated_main.pdf (last visited Aug. 18, 2019); C Mims, *AI That Picks Stocks better than the Pros*, MIT Technology Review (Aug. 17, 2019, 21:25 PM) https://www.technologyreview.com/s/419341/ai-that-picks-stocks-better-than-the-pros. (last visited Aug. 18, 2019).

the possibility of the emergence of a new class of politically conscious, autonomous entities which value human institutions of law and government and are essentially immortal entities with an infinite capacity to access and store information.

Therefore, a simple rebuttal of the claim to citizenship *vis-à-vis* the functional and nonfunctional aspects is not sufficient, for through the course of this paper, the capacity of A.I. to comprehend, uphold and abide by the values and roles of citizenship has been upheld, due to their ability to follow an overarching and overlapping *corpus* of codes, which govern the "life" of the entity, its ability to "think", its manner and understanding of "expression" and "emotion", and its ability to co-exist, just like human beings. Laws regulating the parameters and processes of development, growth and functionality of A.I. must therefore be evolved keeping in mind the potential abilities and capabilities of such entities, and the author argues that human society must pre-emptively prepare for the inclusion of A.I. personalities in their lives and not reduce such technological developments into baseless threat to their existence. To ensure equal status and treatment of citizens is already an obligation undertaken by many States and the International Community,[43] and while the continued racial and patriarchal dominations reflect a species-wide failure to recognize such an inherent value of equal treatment, to inflict such discrimination against A.I. would be to re-inflict the very wounds upon an entire subcategory of entities which has been acknowledged as unjust in the *status quo*.

Constant fluctuations in the criteria for citizenship[44] and denial of claims over citizenship is the primary cause for the migration of refugees and internally displaced persons, and the violence which follows during such periods of protracted inter-community hostilities adversely affects all stakeholders involved. If the same is done to developed conscious beings like A.I., the effects would be equal to that of depriving the rights of a human being, and in anticipation of potential conflict, and such may trigger disputes at a global scale if all sentient A.I.s collectively claim their rights to citizenship, and the author believes that such a potential conflict is easily avoided by the natural transfer of citizenship rights to A.I.s so as to prevent such a problem from emerging in the first place.

[43] G.A. Res. 217 (III) A, Universal Declaration of Human Rights, arts. 1, 2, 4 (Dec. 10, 1948).

[44] Government of Assam, *NRC: In A Nutshell*, NRC ASSAM, (Aug. 20, 2019, 23:43 PM), http://www.nrcassam.nic.in/nrc-nutshell.html.

The only example of an enfranchised A.I. in contemporary law is Sophia,[45] who was granted lawful citizenship by the Government of Saudi Arabia;[46] however, given the highly limited degree of autonomy and political representation of the government of Saudi Arabia, such a recognition does not correspond to an increase in autonomy or the ability of Sophia to acquire and discharge rights. In a democratic society, the consequences of citizenship are wide-ranging, and an adequate legal setup supporting the transition of the society to include these new citizens will become necessary to ensure harmonious co-existence of A.I. and human beings. There are no possibilities that any measures in legal systems with regard to citizenship of A.I. would reduce the possibility of conflict between humanity and its creations, and all dystopian futures of science fiction remain theoretically possible.

Such is compounded by the modality of conferral of citizenships in democratic societies themselves gatekeeping individuals based upon abstractions such as place of birth, time of residence or *"Good Moral Character"*;[47] which remain defined and bound by primarily anthropocentric institutions of the State, such as criminal record, personal history etc. The President of the United States may be deemed the legitimate repository of the will of the People in the Executive, or the British Parliament may embody *vox populi vox deli* for the citizens of the UK; but when rules and regulations for modalities of conferral of citizenship are bound by variables based on abstractions and concepts left at the behest of human beings in perpetuity, there are inevitable conflicts regarding the authenticity of the democratic legal framework, and how representative of its hypothetical cyber citizens it truly is.

Looking at the concept of *"Good Moral Character"* itself, there are several clearly demarcated grounds in law,[48] while leaving a revolving door for the State to remain flexible on what constitutes positive moral standards, as the inclusion of terms such as *"moral turpitude"*. In practical terms, there would be a transition from the anthropocentric model of conferral and determination of citizenship, such as *"Good Moral Character"* potentially being redacted in favor of more tangible variables; however, it remains uncertain whether such a transition would be

[45] *Sofia. Hi, I am Sophia... Hanson Robotics* (Aug. 21, 2019, 13:45 PM), https://www.hansonrobotics.com/sophia/ (last visited June 8, 2020).

[46] Dave Gershgorn. *Inside the mechanical brain of the world's first robot citizen*, QUARTZ (Aug. 21, 2019, 14:02 PM) https://qz.com/1121547/how-smart-is-the-first-robot-citizen/ (last visited June 8, 2020).

[47] Department of Homeland Security, *U.S. Citizenship and Immigration Services, Policy Manual*, 12, INA Part D, Chapter 9, § 1101 101(f) and § 1427 INA 316(e).

[48] *Electronic Code of Federal Regulations*, Title 8, Chapter I, Subchapter C, Part 316, § 316.10.

accepted by the humans of *status quo*, given that any concession of the centrality of the value of human life is incompatible with the legal systems of today.

To propose a society that include A.I. entities as citizens, the author acknowledges the fundamental problem of a lack of *representation* of what might be essential factors of determination of citizenship which include the interests of A.I.; partly because there is no clarity with respect *what* or *how* AI views the notion of citizenship; and partly because there are no principles for conferral of human citizenship which can be universally accepted across A.I. and humans, arguably besides the simple manifestation of *jus sanguinis*. The furor over the contemporary Citizenship Amendment Act 2019 in India[49] is indicative of the sensitive nature of determination of the methodology of conferral of citizenship, and how paradigm shifts implemented by agencies authorized to enact such changes (in this case, the Parliament) generates a significant cognitive dissonance between groups of citizens. Citizenship being entwined with identity politics and becoming instrumentalities for potentially discriminatory policies makes the probability of conflicting interests rupturing the concept of the state itself; and given that the established anthropocentric models of citizenship themselves are not stable or universal, the very idea of another *species* of citizens having determinative rights in policy making regarding who becomes a citizen borders on absurdity.

Such is not helped by the *"Grimdark"* and *"Dystopian"* settings[50] which have been posited by literary and artistic works of fiction, leading to a prevailing fear of a society where humans no longer remain the apex species. While Science Fiction is purely fantastical and uses futuristic settings to deliver social commentary and questions of morality,[51] it belies the underlying apprehension that the technological developments of society will exacerbate the challenges and moral dilemmas of *status quo*. The significance of such a medium beyond its effect on the Collective Consciousness[52] of the populace, is in the principles of Post-humanism taking shape with regard to perceptions of the future in popular culture; as all futures of a growing interface between independent and advanced machinery powered by *"sentient"* A.I. and the fractured, morally challenged human species, generally leads to the latter's decline. Notwithstanding technological and financial constraints on the development of such potentially *"sentient"* A.I. it

[49] The Times of India (Dec. 12, 2019, 13:23 PM), https://timesofindia.indiatimes.com/city/guwahati/citizenship-amendment-act-protests-in-delhi-live-updates/liveblog/72659165.cms (last visited June 8, 2020).

[50] Marshall B. Tymn, *Science Fiction: A Brief History and Review of Criticism*, 23 AM. STUD. J. 41 (1985).

[51] *Id.*

[52] CARL GUSTAV JUNG, THE ARCHETYPES AND THE COLLECTIVE UNCONSCIOUS (Princeton University Press, 1980).

cannot be denied that such depictions have instilled a sense of unease and apprehension within lawmakers themselves with regard to their potential as destructive force.[53]

However, these grim futures are not avoided by sidestepping the need to provide and recognize the autonomy and legal status of A.I. as citizens. The claim that law ought not recognize A.I. as citizens in order to prevent Armageddon, is principally identical to the fears in 19th and 20th century U.S. about recognizing the equality of Black Americans as equal to White Americans causing the destruction of the American society; in fact, in certain parts of America during the Civil War period, the idea that slavery was wrong, and the black slaves deserved legal recognition as equal citizens with voting rights and deserving of representation, was just as alien to their way of life, and just as incompatible with the legal system prevalent at the time. Therefore, the author gives no credence to the projected problems stemming from an A.I. takeover of human society, and instead recommends addressing this dilemma in the same manner that slavery was addressed.

Furthermore, neither do humans need an "enemy" like A.I. to inflict irreparable damage to their society through technological development, nor are all entwined futures of humanity and A.I. bleak. In fact, it is often projected that the relinquishment of the centrality of humans to legal power and discourse would be beneficial to not only other life forms on earth, but to the human species in the future, and thus it may be inevitable to discard speciesism in favor of greater rights of all forms of sentient life, which makes the need to recognize artificial life forms and domesticated animals vital.[54] In this regard, humanity may actually benefit from the existence of sentient A.I. citizens,[55] which can entirely replace the need for human activity allowing all persons to exist in a state of bliss; or for hybrid human-cybernetic organisms to become the next cycle in human life and existence, who can overcome many limitations of the human condition, namely aging, forgetfulness, hunger and death itself.[56] Such an advanced race of humans warrants new laws and legal principles, and their roles in society as citizens will transform. Therefore, since the development of increasingly advanced A.I. systems is ongoing, and upon reaching higher levels of sentience, A.I.s would seek the same types of rights and privileges as their human counterparts, and the fact that A.I.s exhibit both legal personality and citizenship, the author

[53] Psaledakis Daphne*, EU lawmakers call for global ban on 'killer robots'*, (Dec. 12, 2019, 9:43 AM) https://in.reuters.com/article/us-eu-arms/eu-lawmakers-call-for-global-ban-on-killer-robots-idINKCN1LS2AS (last visited June 8, 2020).

[54] DAVID PEARCE, THE ANTISPECIEIST REVOLUTION (Institute for Ethics and Emerging Technologies, 2013).

[55] Nick Bostrom & Yudkowsky Eliezer, *The Ethics of Artificial Intelligence*, *in* CAMBRIDGE HANDBOOK ON ARTIFICIAL INTELLIGENCE (William Ramsey & Keith Frankish eds., Cambridge University Press, 2009).

[56] Jerold J Abrams, *Pragmatism, Artificial Intelligence and Posthuman Bioethics: Shusterman, Rorty, Foucault*, 27(3) HUMAN STUDIES 241 (2004).

concludes that in light of the Post-Humanist philosophical transformations, and the eventual displacement of humanity as the center of the discourse surrounding rights and citizenship, the law consider A.I. as capable of becoming citizens and enjoying democratic rights, equal treatment, political representation and a dignified existence.

In conclusion, through the course of this paper, the author sought to understand the dichotomy of the creation of a class of entities mimicking human intelligence and sentience, with the conferral of citizenship based upon principles which are a combination of the concepts of natality and sanguinary ties, along with abstract notions of naturalization and "good moral character". The purpose behind such an enquiry was to elucidate the possibility of conferring citizenship to A.I. entities, which exhibit similar means of learning and comprehension as humans, and its consequences. It becomes clear that there are no obvious or justifiable reasons in an anthropocentric approach towards citizenship, because neither are the grounds for asserting human superiority and capacity correct, as A.I. express the same or greater indices of learning and intelligent behavior; nor are the criteria for conferral of citizenship in legal systems exclusionary towards A.I. if they can satisfy the test for personhood, whose contours can be determined through instruments like the Turing Test. Therefore, the author concludes that A.I. are equally deserving of citizenship and personhood, and the conferral of the same would not unduly rupture the legal processes for such conferral and enjoyment of rights which are currently taking place.

TOWARDS A NEW ALGORITHMIC SELF?

- Upendra Baxi[*]

I.

Accelerated global social transformation that will, as Ray Kurzweil famously predicts, occur through the propounded law of acceleration, and lead us all towards a "rupture in the fabric of human history."[1] Indeed, Kurzweil describes his book as "story of destiny of the human-machine civilization, a destiny we have come to refer as the Singularity".[2] Whether this constitutes a prophecy[3] or a prediction (based on secular traditions of science) is an important concern,[4] though not pursued directly here. Rather, we concern ourselves, in these few brief remarks, with the implicit notion of algorithmic self – a notion deeply intertwined with the related but larger idea of algorithmic futures.

We must note at the very outset, that we do not know how best to attend to the problems of risk and trust in modern society, rolled up in the notion of "reflexive modernization".[5] In the current discourse on AGI (artificial general intelligence), we do not quite know whether AGI developments deliver sufficiently proximate benefits or betoken some catastrophic danger. Nor do we quite know where to place AGI (still in the offing as 'super-intelligence') among the litany of sinister man-nature made catastrophic situations and how to prioritize risk and trust strategies – whether it be COVID-19 global pandemic, the accelerating rate of anthropogenic harm now occurring, the continuing late Holocene destruction of ecology and environment, or the patterns and perils of nuclear proliferation.

[*] Professor of Law, Jindal Global Law School, India.

[1] RAY KURZWEIL, THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY 5 (Penguin Books 2005).

[2] *Id.* at 5.

[3] Considered worthy of pursuit by millenarians or co-religionists. NORMAN COHN, THE PURSUIT OF THE MILLENNIUM: REVOLUTIONARY MILLENARIANS AND MYSTICAL ANARCHISTS OF THE MIDDLE AGES (Oxford University Press 1970); Richard Landes, *Millenarianism/Millennialism, Eschatology, Apocalypticism, Utopianism*, *in* HANDBOOK OF MEDIEVAL CULTURE: FUNDAMENTAL ASPECTS AND CONDITIONS OF THE EUROPEAN MIDDLE AGES (Albrecht Classen ed., 2015). There is an aura of millenarian thought traditions concerning both the prophetic and secular discourses.

[4] SINGULARITY HYPOTHESES: A SCIENTIFIC AND PHILOSOPHICAL ASSESSMENT (Amnon H. Eden, James H. Moor, Johnny H. Søraker & Eric Steinhart eds., Springer New York 2012).

[5] SCOTT LASH & JOHN URRY, ECONOMIES OF SIGNS AND SPACE (London Sage Publications 2004) [hereinafter LASH & URRY]; Scott Lash, *Reflexivity and Its Doubles: Structure, Aesthetics, Community*, *in* REFLEXIVE MODERNIZATION: POLITICS, TRADITION AND AESTHETICS IN THE MODERN SOCIAL ORDE 110-173 (Ulrich Beck, Anthony Giddens & Scott Lash eds., Stanford University Press 1994).

But at least three ways seem offered, first, on the side of industry, we stand educated in the ways in which all science and technology products must pass through what has been called the Gartner hype cycle. The five phases include: Technology Trigger (a potential technology breakthrough, and significant publicity; 'Peak of Inflated Expectations' (great expectations at an early stage); 'Trough of Disillusionment' (waning of interest because experiments and implementations do not match expectations); 'Slope of Enlightenment' (emergence of public trust in innovations); and 'Plateau of Productivity' (the technology's broad market for innovations continue to be adopted and grow).[6]

Second, is the issue not merely of long- term agendas of concern but of prioritization as a matter of short run governance (state as well as non-state networks). This has been framed differently in modernization theory as reflexive social trust in expert systems, which is a common enough experience in modernity – even so as to constitute modernity itself.[7]

But, second, trust in elites happens side by side with a common deinstitutionalization. The trust deficit has been on the increase, more particularly with the phenomenon of digital crowds, which has acquired a new salience. James Surowiecki (in a book hailed as 'illuminating' and 'revolutionary') argues cogently that "under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them. Groups do not need to be dominated by exceptionally intelligent people in order to be smart. Even if most of the people within a group are not especially well-informed or rational, it can still reach a collectively wise decision. This is a good thing, since human beings are not perfectly designed decision makers". Noting many weighty objections to the wisdom of the crowd, he is still able to say that when "our imperfect judgments are aggregated in the right way, our collective intelligence is often excellent".[8] Is machine learning and intelligence now posed to overcome that wisdom?

One can safely say that human brains or human species (*homo sapiens*) is fast becoming *machina sapiens,*[9] with all its bright aspects and sinister sides. Humans always used algorithms as

[6] These rough indications are to be read in the methodology material available on the Gartner website, alongside with studies in digital marketing.

[7] LASH AND URRY, *supra* note 5.

[8] JAMES SUROWIECKI, THE WISDOM OF CROWDS: WHY THE MANY ARE SMARTER THAN THE FEW AND HOW COLLECTIVE WISDOM SHAPES BUSINESS, ECONOMIES, SOCIETIES AND NATIONS (Doubleday New York 2004).

[9] GABRIEL HALLEVY, WHEN ROBOTS KILL: ARTIFICIAL INTELLIGENCE UNDER CRIMINAL LAW 5 (Northeastern University Press 2013), describing AI research as offering four main categories: "those that (1) act like humans, (2) think like humans, (3) think rationally, and (4) act rationally"; JAMES D. MILLER, SINGULARITY RISING SURVIVING AND THRIVING IN A SMARTER, RICHER, AND MORE DANGEROUS WORLD (Ben Bella Books, Inc. 2012). Miller

rudimentary guides to choosing among multiple options for choices of conduct. The common example of algorithms is, of course, recipes (making pasta or soup); these rites are commonplace but also complex and the human brain works with number of algorithms that make acts of daily living possible. In a sense, machine learning and artificial intelligence (non-biological intelligence; hereafter, AI) transforms the process and the problematic in going to the shores of AGI where mere machine learning replaces itself with several human and trans-human – even post-human – attributes: those of self-replication, regulation, thinking, feeling, suffering and even assuming a form of new being, previously unknown.

Among the major anxieties of AGI developers is how to develop "rational architectures for intelligence"[10] that may lead to human friendly virtues of goodness, fairness, justice, and rights. That aim is harder to achieve as the AGI begins to develop 'sub-goals' or its own teleology—or when it "would rapidly replicate, distribute … widely, attempt to control all available resources, and attempt to destroy any competing systems in the service of their initial goals which they would try to protect forever from being altered in any way." Apparently, the "safe AI scaffolding strategy" would resolve these questions of the potential for 'anti-social behaviour' by "early forays into creating autonomous systems" which would then prevent inadvertent or malicious creation of "this kind of uncontrolled agent".[11]

I do not know how this may work, if at all; but I know enough from the history of industrial catastrophes (including my long, and still continuing, engagement with struggle for justice waged by the Bhopal violated)[12] to say that business and investment decisions will continue to develop styles of corporate governance that will specialize in re-victimizing the victims and in ways of production of social indifference to the immiseration of survivors. Only a strong human rights discipline and invigilation of international and national human rights law and jurisprudence will help address this iniquitous and ubiquitous industrial hazard system. But this is a distant dream in an era where neo-liberalism, and its five allies – denationalization, deregulation, de-juridification,

describes dystopian "ultra-AI" which would quickly create "any military technologies we lack. It might take such an AI a minute to create a million different kinds of airborne viruses, each capable of wiping out our species. Or the AI could use nanotechnology to make a self-replicating factory, which could make ten robot soldiers and ten copies of itself every hour".

[10] Steve Omohundro, *Rational Artificial Intelligence for the Greater Good*, in SINGULARITY HYPOTHESES: A SCIENTIFIC AND PHILOSOPHICAL ASSESSMENT 163 (Amnon H. Eden, James H. Moor, Johnny H. Søraker & Eric Steinhart eds., Springer New York 2012).

[11] *Id.* at 172. In contemplating this, one may not avoid what has been termed as epistemic *akrasia* (acting against one's best judgment): this phenomenon is insightfully analysed by Declan Smithies, *Epistemic Akrasia*, *in* THE EPISTEMIC ROLE OF CONSCIOUSNESS (Declan Smithies ed., Oxford University Press 2019).

[12] Upendra Baxi, *Human Rights Responsibility of Multinational Corporations, Political Ecology of Injustice: Learning from Bhopal Thirty Plus?*, 1(1) BUSINESS AND HUMAN RIGHTS JOURNAL 21-40 (2016).

de-constitutionalization, and de-democratization continue to reproduce itineraries of human rightlessness and social misery.

II.

The tiny letter 'I' has generated a lot of philosophical anxiety and concern with the 'subject' in the sphere of cultural studies. A very broad question arises when one thinks about "algorithmic self": is it merely a type of many selves humans may choose to have, or is it a new and unique self--that is, an autonomous reflexive self? A self that may overcome, and act independently of, humanity? This question is scarcely novel in jurisprudence as legal theory, and the latter discussed the centrality of the idea of subject, a self that is also the other.

The former discussed the distinction between 'artificial' legal person and 'natural persons'; and in particular the legal personality of corporations.[13] Especially, the 'bracket' theory of legal personality developed by Rudolf Jhering and the French 'will' or 'enterprise' theory' wrestled with the problems of how fictionality of legal persons could ever transcend the will of those who created them in the first place.[14]

It remains intriguing that in science and technological studies tend to ignore these juristic discourses which relate to digital selfhood and 'human' rights of some AI life-forms. Because these offer some fascinating points of comparison and contrast, they should not be totally ignored thus.

The relation between law and language is equally fascinating and pertinent to discourse concerning AI life-forms. Adrigeus Griemas pointed out that a peculiarity of legal language consists in the invention words that do not pre-exist in society or nature (citizens, corporations, trusts, State, for example) and production of jural and social meanings occur through both the forms of *production juridique* and *verification juridique.*[15] And in the field of sociolinguistics, John Searle was a pioneer in talking about first the social construction of reality and the ontic powers

[13] SAMIR CHOPRA & LAURENCE F. WHITE, A LEGAL THEORY FOR AUTONOMOUS ARTIFICIAL AGENTS 153-193 (Ann Arbour, The University of Michigan Press 2011); Ngaire Naffine, *Who are Law's Persons? From Cheshire Cats to Responsible Subjects*, 66(3) MODERN LAW REVIEW 343-367 (2003); VISA A.J. KURKI & TOMASZ PIETRZYKOWSKI, LEGAL PERSONHOOD: ANIMALS, ARTIFICIAL INTELLIGENCE AND THE UNBORN (Springer International Publishing AG 2017); Juyun Kim And Stephen Petrina, *Artificial Life Rights: Facing Moral Dilemmas Through the Sims*, 10 EDUCATIONAL INSIGHTS 84 (2006).

[14] ROSCOE POUND, VI JURISPRUDENCE (St. Paul Minnesota 1959). For the French form of enterprise theory, JULIUS STONE, SOCIAL DIMENSIONS LAW AND JUSTICE (Maitland Press 1966); Maksymilian Del Mar, *Introducing Fictions: Examples, Functions, Definitions and Evaluations*, *in* LEGAL FICTIONS IN THEORY AND PRACTICE (Maksymilian Del Mar & William Twining eds., Springer New York 2015).

[15] BERNARD JACKSON, SEMIOTICS AND LEGAL THEORY (London, Routledge & Kegan Paul 1985).

of legal language.[16] May one speak already of *ontic powers* of AI life-forms in the spheres of distributed AGI?

III.

The problem of many selves, or as Paul Ricoeur formulated it in terms of 'oneself as another',[17] persists also in the domains of AGI. Many selves are possible, on this analysis, when a core individual self is postulated; a self that is singular or is at home itself. Or, only humans are capable of attaining selfhood? Is there a distinction between persons and things, or persons and nonpersons? May we say that algorithmic self exists? Or, is it the case that the concept, and conceptions, of 'self' itself is problematic?

In cultural studies, the proliferation of self/subjects is as remarkable as the death of the subject. The loss of self is celebrated as were lamented. Judith Butler says, for example, that "the body is not merely matter but a continual and incessant materializing of possibilities" and the "embodied subject constitutes meaning by the "taking up and rendering specific of a set of possibilities".[18] In this sense, AGI life-forms remain deeply performative and reiterative.[19]

The wounded and traumatized self, the accidental and contingent self, the marginalized and exploited selves of 'abject subjects' are narrated as frail selves in contrast to ontologically robust selves such as governance selfhood, dominant historical selves, the ableist self, robust moral and dharmic self, the custodians and citadels of patriarchal and heterosexual selfhood. Self, always in the making, and often in fear and trembling, which is peacefully but ethically insurgent includes resistant selves, creative self, and resilient self. 'Oneself as another' is a constant warfare and law-

[16] *Id.*; JOHN ROGERS SEARLE, THE CONSTRUCTION OF SOCIAL REALITY (New York Free Press 1997); MAKING THE SOCIAL WORLD: THE STRUCTURE OF HUMAN CIVILIZATION (Oxford University Press 2010).

[17] PAUL RICOEUR, ONESELF AS ANOTHER (Kathleen Blamey, trans., University of Chicago Press 1992). He introduces the distinction between *Idem* identity (the identity that never changes and*, ipse* identity which is sameness' through change. The existence of identity indicates that a self presents a better regarded as the question of "who?" rather than "what" is a self. Ricoeur to introduce what he calls his "little ethics "which re-orientates the argument toward prescription, rather than descriptive; this little ethics presents always the ethical intention: "aiming at a good life lived with and for others in just institutions" (at 172).

However, Henry Isaac Venema in his IDENTIFYING SELFHOOD: IMAGINATION, NARRATIVE, AND HERMENEUTICS IN THE THOUGHT OF PAUL RICOEUR (New York University Press 2000), has this to say about the central distinction between self-sameness (*idem* identity) from self- constancy (*ipsey* identity): Ricoeur thinks that moving from the question "what is a subject?" to the question "who is the object? allows him to uncover a type of selfhood that is not reduced to the identifying structures of the sameness. However, …this is not the case… Ricoeur has hopelessly entangled selfhood in the semantics of identity, and has encircled" 'the pure selfhood of self-constancy' … within the power of the self -same" (at 11).

[18] Kim Atkins, *Commentary on Butler*, *in* SELF AND SUBJECTIVITY 254 (Wiely-Blackwell 2005).

[19] *Id.* Reiterative precisely in the sense Butler gives to the notion of "forced reiteration"—that is, reiterative acts that "locate the possibility of resistance within the process of reiteration itself" (at 253).

fare (in so far as law is to regarded as an alternate 'politics by other means') of various self and identity-making practices or as Michel Foucault would say "technologies of self".

The reference to Foucault brings us with the notion of 'self' that he articulated as a programme of "remaking"; the self as "a sort of permanent and multiple enterprise".[20] This is an extraordinary notion of self not as quintessence or essence but as so many heterogeneous practices, a notion which also remains true about AGI life-forms, Foucault was not referring here to patterns of governance and market typical of a liberal order; rather, he made here explicit some of the postulates of neoliberal governance that gave rise to a self that internalized the goals of governance and laws of free market as principles of making of self. In neoliberal state and economy, as Lois McNay reminds us,[21] the remodelling of the subjective experience of the self around an economized notion of enterprise subtly alters and depoliticizes conventional conceptions of individual autonomy. "Individual autonomy…lies at the heart of its disciplinary control" and this "inevitably challenges conceptions of resistance, freedom and political opposition, which often invoke a notion of individual autonomy as an absolute block or challenge to the workings of power"[22]. Indeed, "…it understands the commodification of subjective experience not so much through ideas of passive consumerism, standardization and heteronomy, as through ideas of active differentiation, regulated self-responsibility and depoliticized autonomy".[23] Perhaps, Foucault's principal difficulty with these self-making practices" is that, under the guise of promoting individual autonomy, it is profoundly normalizing".[24]

We will not here tarry further with this connection between the rise of the new practices of self-forming as an enterprise but it is tolerably clear that AGL is well at home with this mode of self-making and its rise and growth seems to be an integral feature of neoliberal governance and markets. If so, the relationships between law and technologies have to be understood anew,

[20] MICHEL FOUCAULT, THE BIRTH OF BIOPOLITICS: LECTURES AT THE COLLÈGE DE FRANCE 1978–1979 241 (Palgrave Macmillan 2008).

[21] Lois McNay, *Self as Enterprise Dilemmas of Control and Resistance in Foucault's The Birth of Biopolitics*, 26(6) THEORY, CULTURE & SOCIETY 55–77 (2009).

[22] *Id.* at 62.

[23] *Id.*

[24] Lois McNay quotes Foucault's notion as a "form of power" that "applies itself to immediate everyday life which categorizes the individual . . . attaches him to his own identity, imposes a law of truth on him which he must recognize and others must have to recognize in him."

beyond the famous sociological hypothesis of a cultural lag between law and technology.[25] In the neoliberal times, and beyond, we perhaps need to think of law itself *as* a social technology seeking to impose some discipline of rights and responsibilities on the managers and masters of 'hard' technologies.

[25] WILLIAM F. OGBURN, SOCIAL CHANGE (B.W. Huebsch Inc. 1922); Floyd Henry Allport, *Social Change: An Analysis of Professor Ogburn's Culture Theory*, 2 SOCIAL FORCES 671-676 (1924); John J. Honigmann, *The New Attack on Cultural Lag*, 7(1) THE ANTIOCH REVIEW 55-63 (1947).

SUBMISSION GUIDELINES

The Law and Society Review is an annually published, faculty cum student led, double-blind peer-reviewed academic law review focussing on socio-legal issues and processes. It employs and encourages an interdisciplinary and multidisciplinary approach to the study of law and legal issues. The LSR is open for academics, scholars, research, and doctoral students. It encourages undergraduate students to contribute case comments.

- All submissions must be accompanied with an Abstract, explaining the aims and objectives of the paper. The word limit for the Abstract is 500 words.
- The manuscript should be anonymous i.e. it should not contain the name of the author/s, educational qualifications or any institutional affiliations.
- The details of the author/s (name, institutional affiliation, if any, and contact details) must be contained in a separate attachment in .doc or .docx format.
- Co-authorship is allowed to a maximum of two authors.
- All submissions made, must not have been previously published or under consideration/review elsewhere. Disclosure to this effect must be made to the editorial board.
- All submissions must follow the "The Bluebook: A Uniform System of Citation (20th Ed.)". Non-conformity could be a ground for rejection.
- The body of the manuscript must be in the font "Garamond", font size 12, line spacing 1.5. All footnotes must be in Font "Garamond", font size 10, line spacing 1.
- Submissions must be made in Microsoft Office (doc. /docx.) formats only.

Submission Categories

- The submissions could fall under any of the following categories:

  1. Articles (6000 words or above, excluding footnotes);
  2. Essays (3500-6000 words excluding footnotes);
  3. Book Reviews/Case Commentary/Short Notes (1500-3500 words excluding footnotes)

- All the submissions must be made to glsr@gnlu.ac.in. Authors can expect to receive decisions on their submissions within two months of their submissions.
- Please note that the procedures for reviewing manuscripts are based on the anonymity of the author and the confidentiality of the editors' and reviewers' reports; author anonymity is preserved, as well, during the editorial decision-making process.

NOTES

GNLU LAW AND SOCIETY REVIEW

Published by:

GNLU Centre for Law and Society,
Gujarat National Law University,
Attalika Avenue, Koba, 382426
Gandhinagar, Gujarat.

Email – gcls@gnlu.ac.in

आ नो भद्राः क्रतवो यन्तु विश्वतः
GNLU
Gujarat National Law University