HIMACHAL PRADESH NATIONAL LAW UNIVERSITY, SHIMLA

CENTRE FOR HUMAN RIGHTS AND DISABILITIES STUDIES, HPNLU

PRESENTS

AMULYA ADHIKAR MAGAZINE

DIGITAL JUSTICE: SAFEGUARDING HUMAN RIGHTS IN THE AGE OF ARTIFICIAL INTELLIGENCE

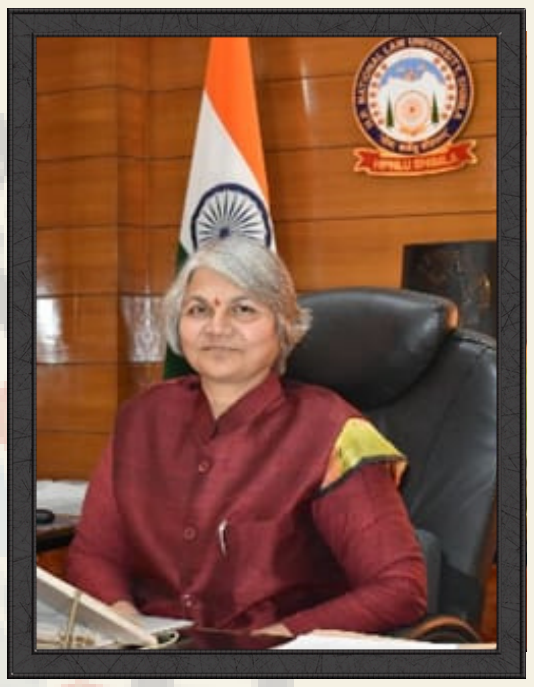
ABOUT THE CENTRE

Human diversity has been celebrated throughout history and recognized in transformative constitutionalism, emphasizing equality, autonomy, and dignity. Philosophical concepts like Immanuel Levinas' "One and other as Same" and Vedic teachings on selflessness highlight the value of pluralistic perspectives. However, stereotypes and binary narratives often perpetuate a gap between theoretical ideals and practical implementation. This gap results in discrimination against marginalized groups, particularly people with disabilities and Indigenous communities, who face systemic barriers to fundamental rights like healthcare and dignity.

The COVID-19 pandemic exposed these vulnerabilities further, prompting international directives from the UN and WHO to address such disparities. The Centre for Human Rights and Disability Studies at Himachal Pradesh National Law University, Shimla, aims to bridge this gap. It will amplify first-person voices, conduct action-based research, and engage with stakeholders to address healthcare, education, employment, and justice concerns. The Centre seeks to set a standard for integrating rights discourse into practice, focusing on dignity, autonomy, and equality for diverse communities.

MESSAGE FROM VICE CHANCELLOR

Prof. (Dr.) Priti Saxena, Vice Chancellor of HPNLU, Shimla, is a distinguished legal scholar with over 33 years of experience as an educator, researcher, and administrator. A gold medalist in LL.M. and Ph.D. holder, she has authored two books, edited four, and published over 85 research papers. She has been a key speaker at global institutions and has led numerous conferences, moot courts, and faculty development programs. Her expertise lies in Constitutional Law, Human Rights, and Governance, and she has mentored 18 Ph.D. scholars. She actively contributes to legal education, policy-making, and community outreach.



In an age of change spearheaded by digital innovations, Artificial Intelligence has gone a great way to become a strong engine for shaping industries, governance, and daily life. However, promising innovations come with the imperative duty of protecting the fundamental human rights that are put at stake by this new frontier. While AI provides unprecedented avenues for fostering equality, broadening access to information, and giving a voice to marginalized people, it comes with its fair share of profound concerns. Inherent biases, mass surveillance, and misinformation without any transparency, accountability, and ethical guardrails threaten to derail the very principles of dignity, equality, and autonomy, thereby eroding our trust in democratic institutions.

The 2024 UDHR theme, "Your Rights, Your Future, Right Now" serves as a rallying cry, reminding us that Human Rights are not abstract ideas but a guide to every aspect of our life, including the digital realm. The Centre for Human Rights and Disability Studies at the Himachal Pradesh National Law University endeavoured to understand the role of justice in a digitized landscape that questions age-old norms. This commitment springs from a collective belief that technology must serve humanity, not vice versa.

It is with great pride that I contribute to this edition of Amulya Adhikar, which explores the timely and crucial theme of "Digital Justice: Safeguarding Human Rights in the Age of Artificial Intelligence." I wish to commend the authors for the astute work submitted, delving into the complexity of the effects of AI on human rights. Your endeavour exemplifies the spirit of advocacy and discussion that these revolutionary times demand.

MESSAGE FROM THE DIRECTOR

Dr. Sachin Sharma is an Associate Professor of Law at Himachal Pradesh National Law University, Shimla, with a keen interest in philosophy, disability jurisprudence, Healthcare, and gender-sexuality studies, with his primary area of interaction being Legal Theory and Public Law. He used to interact with students on subjects including Law and Weaker Sections; Law, Poverty, and Development; Jurisprudence; Law of Tort; Interpretation & Constitutional Studies; Law and Justice in Globalizing World, etc. Apart from his academic interests, he is also associated with action research in disability studies.

The rapidly evolving field of technology, which usually adopts the 'move fast and break things' approach presents itself with both transformative opportunities and significant challenges thereby making the arena of Human Rights increasingly important. Along these lines, the 2024 UDHR theme, "Your Rights, Your Future, Right Now" is a wake-up call for all of us to act with urgency and vision thereby underscoring the necessity to be fully aware and vigilant about one's human rights and all those principles guiding their development, with Artificial Technology being one of them.

This first edition of 'Amulya Adhikar' features diverse and insightful contributions that intend to unravel the complex yet pressing intersections between technology and human rights. The submissions have helped not only raise awareness about Human Rights but have also inspired actionable solutions to uphold the dignity and equity of all individuals in this digital age.

I would like to extend my heartfelt gratitude to all contributors, whose thoughtful writeups have engendered and enriched this edition. Furthermore, special thanks to the members of the Centre for Human Rights and Disabilities Studies for their outstanding dedication and efforts in bringing this magazine to life.

INDEX

NAVIGATING THE PARADOX: AI MASS MISINFORMATION, FREE SPEECH, AND THE CASE FOR A CONSTITUTIONAL AMENDMENT

-Saksham Rai & Ishita Kumar

The proliferation of misinformation has reached unprecedented levels, fuelled by the rapid evolution of Artificial Intelligence (AI). Generative AI, with its ability to produce high-quality, realistic content at scale, has revolutionized the ways misinformation is created and disseminated. From deepfakes of global leaders manipulating voter sentiments to AI-altered images used in political advertisements, the quality and quantity of false narratives have surged. In India, the 2024 general elections were a striking example of this trend.

For instance, during India's 2024 general elections, Meta approved AI-manipulated political advertisements that incited religious hatred and spread baseless claims. Examples included ads falsely alleging an opposition leader's intent to "erase Hindus from India" alongside a fabricated image of the Pakistani flag. Another approved ad used AI-generated visuals to incite communal violence, containing slurs like "let's burn this vermin" directed at Muslims.[1] The impact of AI-generated misinformation is deeply troubling for Indian democracy. It pollutes the information ecosystem, blurring the line between fact and fiction. During the 'Delhi Chalo' farmers' protest, AI-generated images falsely depicted modified tractors intended to break police barricades, stoking public outrage.[2] Fact-checking revealed these visuals to be fabrications, yet their rapid spread showcased the power of misinformation to undermine trust in democratic movements. Similarly, deepfake videos of political figures swayed voter opinions before they could be flagged as false, leaving behind a trail of misinformed voters and fractured democratic discourse.[3]

Despite its evident dangers, AI-generated misinformation that does not defame individuals or incite immediate violence enjoys protection under Article 19(1)(a)[4] of the Indian Constitution, which guarantees the right to free speech. This constitutional safeguard creates a paradox: while the state can restrict speech on grounds such as defamation or public order under Article 19(2)[5], misinformation that indirectly threatens democratic processes by shaping public opinion remains unregulated. In this essay, the authors will explore how this legal gap permits the unchecked proliferation of AI misinformation and make a case for amending Article 19(2) to address this critical challenge.

Misinformation Not Saved by Article 19(2)

In an era marked by the explosive growth of misinformation, the Indian government sought to address this challenge through the establishment of State Fact Checking Units (FCUs). These units were introduced via the 2023 amendments to Rule 3(1)(b)(v) of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.[6] The FCUs were tasked with identifying and flagging misinformation targeting the central government, effectively compelling intermediaries to take down flagged content. However, these amendments faced widespread criticism for their perceived overreach and potential impact on free speech.

The constitutional validity of the FCUs was challenged by comedian Kunal Kamra in the Bombay High Court. Mr. Gautam Bhatia in the case, contended that misinformation, in itself, is not a recognized ground for imposing restrictions on free speech under Article 19(2) of the Indian Constitution. In a pivotal judgment, the Bombay High Court struck down the amendments as unconstitutional, with Justice Chandurkar's tie-breaking opinion emphasizing that the restrictions imposed by the FCUs could not be traced to any of the grounds enumerated under Article 19(2). This ruling reaffirmed a well-established principle in Indian constitutional law: restrictions on the fundamental right to free speech under Article 19(1)(a) must be justified solely on the grounds explicitly provided in Article 19(2).

The jurisprudence underpinning this principle has been articulated in a long line of Supreme Court decisions. The decision in Romesh Thapar v. State of Madras[7] laid the foundation for this jurisprudence. Here, the Supreme Court struck down a law that imposed prior restraint on the circulation of a magazine, emphasizing that restrictions on free speech could only be justified if they were necessary to address one of the concerns enumerated in Article 19(2). This judgment underscored the paramount importance of free speech in a democratic society. Drawing from Romesh Thapar, the Apex Court has reiterated that restrictions on free speech must be narrowly tailored and cannot extend beyond the boundaries set by Article 19(2) in multiple cases.

---

[1] Revealed: Meta approved political ads in India that incited violence, The Guardian, https://www.theguardian.com/world/article/2024/may/20/revealed-meta-approved-political-ads-in-india-that-incited-violence. (last accessed on 25 November 2024).

[2] Unravelling AI-Generated Misinformation in the 'Delhi Chalo' Farmers' Protest, Cyber Peace https://www.cyberpeace.org/resources/blogs/unravelling-ai-generated-misinformation-in-the-delhi-chalo-farmers-protest. (last accessed on 23 November 2024).

[3] Misinformation fueled by Gen AI threatens democratic elections worldwide, warns CSDI report, The Indian Express, https://indianexpress.com/article/technology/artificial-intelligence/generative-ai-us-elections-ai-deepfakes-csdi-report-9564138/. (last accessed on 29 November 2024).

[4] India Const. art 19, cl. (1)(a).

[5] India Const. art 19, cl. (2).

[6] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Rule 3(1)(b)(v).

[7] Romesh Thapar v. State of Madras, 1950 AIR 124 (India).

In Sakal Papers Pvt. Ltd. v. Union of India[8], the Supreme Court invalidated the Newspaper (Price and Page) Act, 1956, which sought to regulate the price and page count of newspapers. The Court held that such restrictions, aimed at ensuring equitable distribution of newspapers, could not be justified under Article 19(2) as they did not relate to any of its specified grounds. The case further solidified the principle that the State cannot impose restrictions on free speech for reasons not explicitly mentioned in Article 19(2), even if such restrictions are purportedly for the public good. The principle was upheld in Express Newspapers v. Union of India[9], where a constitutional challenge to the Working Journalists (Conditions of Service) and Miscellaneous Provisions Act, 1955, was examined. The Court emphasized that any law infringing Article 19(1)(a) must satisfy the requirements of Article 19(2) and cannot extend beyond its permissible scope. This sentiment was echoed in Indian Express Newspapers (Bombay) Pvt. Ltd. v. Union of India[10], which struck down a similar regulatory attempt, and Sodhi Shamsher v. State of Pepsu[11], which maintained that any restriction not traceable to Article 19(2) would be unconstitutional.

One of the strongest reiterations of this principle is the case of Shreya Singhal v. Union of India[12], where the Supreme Court invalidated Section 66A of the Information Technology Act, 2000[13], on the grounds that it violated Article 19(1)(a) and was not protected under Article 19(2). Section 66A, which criminalized certain forms of online speech, was criticized for its vagueness and overbreadth, allowing it to be misused to curtail legitimate expression. The Court held that the provision went beyond the eight permissible grounds for restricting free speech under Article 19(2).

More recently, in Anuradha Bhasin v. Union of India[14], the Supreme Court reinforced the principle while evaluating restrictions on internet access in Jammu and Kashmir. The Court recognized the internet as a critical medium for exercising free speech under Article 19(1)(a) and held that restrictions on its use must conform to the grounds outlined in Article 19(2). The judgment also highlighted the importance of ensuring that any restrictions imposed on speech, or its mediums do not have a disproportionate impact on the fundamental right.

In this evolving jurisprudence, the Bombay High Court's judgment in Kunal Kamra v. Union of India[15], represents a contemporary reaffirmation of the constitutional limits on the State's power to regulate speech. While the FCUs were established in response to the real and pressing issue of misinformation, the Court held that the grounds for such restrictions must be constitutionally sanctioned. The decision aligns seamlessly with the precedent that Article 19(2) is exhaustive and cannot be expanded to include new grounds and if there exists a restriction beyond the grounds mentioned, the same is unconstitutional.

Alternatives in the Public Interest

The established jurisprudence surrounding Article 19(1)(a) of the Indian Constitution firmly confines restrictions on free speech to the grounds enumerated in Article 19(2). However, alternative perspectives have been proposed in judicial opinions that question the rigidity of this framework, suggesting a broader interpretation to accommodate evolving societal and national interests. These alternatives, though non-binding, offer a lens through which the contours of free speech could potentially be redrawn, especially in the context of misinformation and hate speech.

Justice Reddi's concurring opinion in PUCL v. Union of India[16], provides a nuanced departure from the strict confines of Article 19(2). Drawing upon Justice Jeevan Reddy's observations in The Secretary, Ministry of Information v. Cricket Association of Bengal[17], Justice Reddi argued that certain inherent limitations could be read into the right to free speech and expression if necessary to serve the broader interests of the nation or society. Justice Jeevan Reddy had previously posited that while Article 19(2) does not explicitly refer to "national interest" or "public interest," the enumerated grounds such as sovereignty, security of the State, and public order inherently align with these broader concepts. He referred to the U.S. Supreme Court's ruling in FCC v. National Citizens Committee for Broadcasting[18], where the denial of a broadcasting license in public interest was upheld despite the First Amendment's absolute guarantee of free speech. Justice Reddi adopted a similar rationale, emphasizing that freedom of expression must not endanger societal or national interests, even if these terms are not expressly used in Article 19(2). This perspective advocates for implied limitations on free speech, expanding the framework to address emerging challenges like misinformation.

---

8 Sakal Papers (P) Ltd., and Ors v. Union of India, 1962 AIR 305 (India).
9 Express Newspapers (Pvt) Ltd and Anr. v. Union of India and Ors, 1986 AIR 872 (India).
10 Indian Express Newspapers (Bombay) v. Union of India, 1986 AIR 515 (India).
11 Sodhi Shamsher Singh and Ors. v. The State of Pepsu and Ors., AIR 1954 SC 276 (India).
12 Shreya Singhal v. Union of India (2013) 12 SCC 73 (India).
13 Information Technology Act, No. 21 of 2000, §66A.
14 Anuradha Bhasin v. Union of India, AIR 2020 SC 1308 (India).
15 Kunal Kamra v. Union of India, 2024 SCC OnLine Bom 3025 (India).
16 PUCL v. Union of India, (2003) 4 SCC 399 (India).
17 Secretary, Ministry of I & B, State of W. B v. Cricket Association (1995) 2 SCC 161(India).
18 FCC v. National Citizens Committee for Broadcasting, 436 US 775 (1978) (U.S.).

A complementary yet restrained approach was articulated by Justice P.B. Sawant in his opinion in Cricket Association of Bengal. Justice Sawant unequivocally affirmed that restrictions on Article 19(1)(a) must strictly conform to the grounds specified in Article 19(2). However, he also acknowledged that the unique nature of broadcasting media might warrant regulatory measures within the framework of public interest. While rejecting the Union's submission that free speech could be curtailed on grounds beyond Article 19(2), Justice Sawant emphasized that the right to telecast or broadcast carries an inherent duty to ensure access to the widest possible audience. Regulation, therefore, could address licensing and content control to safeguard public interest, provided such regulation operates strictly within the bounds of Article 19(2).

Justice Nagarathana's dissent in Kaushal Kishor v. State of Uttar[19], Pradesh further explored the limits of the protective perimeter of Article 19(1)(a). She adopted a structured analysis of rights and correlative duties, drawing on Hohfeld's framework to argue that the extent of the State's duty to refrain from interfering with speech depends on its content and societal value.

While affirming that Article 19(2) remains the constitutional basis for imposing restrictions on free speech, she proposed a distinction between speech that constitutes a "propagation of ideas" and speech devoid of social value, such as hate speech. Justice Nagarathana argued that such derogatory or vitriolic speech, which undermines the ethos of a civilized society, falls outside the protective ambit of Article 19(1)(a) and can be restrained without recourse to Article 19(2). Her dissent posited that hate speech does not necessitate a balancing act between competing rights but rather represents an abuse of free speech. Collectively, these judicial opinions present an alternative paradigm for interpreting restrictions on free speech, one that extends beyond the exhaustive list in Article 19(2) and can be used to curb the rampant AI misinformation which is threating Indian democracy. While Justice Reddi and Justice Jeevan Reddy advocate for implied limitations rooted in national and societal interests, Justice Sawant and Justice Nagarathana suggest a content-sensitive approach that differentiates between socially valuable speech and possibly harmful expressions like AI based misinformation.

Judicial Response to Alternative Approaches

While alternative interpretations of Article 19(1)(a) discussed in cases like PUCL v. Union of India, and Cricket Association of Bengal have proposed expanding the grounds for restricting free speech, the judiciary has unequivocally rejected these approaches in subsequent landmark judgments. These rulings reaffirm the principle that restrictions on free speech must strictly adhere to the grounds enumerated in Article 19(2). Through majority opinions in Kaushal Kishor v. State of Uttar Pradesh[20], and ADR v. Union of India (the electoral bonds judgment), the courts have fortified the rigidity of Article 19(2) as an exhaustive and non-negotiable framework, effectively rendering the earlier alternative views redundant.

The majority in Kaushal Kishor categorically held that courts could not expand the grounds for restricting free speech by employing interpretative tools to go beyond the eight grounds listed in Article 19(2). This decision directly addressed and negated the perspectives advanced by Justice Jeevan Reddy in Cricket Association of Bengal and Justice Reddi in PUCL. While these opinions had proposed implied limitations on free speech in the broader interest of society or national welfare, the majority in Kaushal Kishor reasserted that the Constitution does not contemplate such implied limitations. The Court emphasized that judicial creativity cannot substitute or augment the explicit language of the Constitution.

Justice Nagarathana's dissenting opinion in Kaushal Kishor also came under implicit critique by the majority. Her method of categorizing speech into those that propagate ideas or possess social value and those deemed unfit for a civilized society, such as hate speech, was effectively dismissed as inconsistent with the constitutional framework. The majority firmly held that all restrictions on speech must derive their validity from Article 19(2) and no classification of speech beyond this framework could justify interference. As a result, her argument that certain forms of speech could be restrained without recourse to Article 19(2) was left without constitutional footing under the binding majority ruling.

The position articulated in Kaushal Kishor was further cemented in ADR v. Union of India, where Chief Justice D.Y. Chandrachud directly addressed the scope of Article 19(2). In this case, which dealt with the constitutionality of the electoral bonds scheme, the Court scrutinized whether curbing black money could justify restrictions on free speech. Chief Justice Chandrachud categorically observed that considerations such as public interest or national welfare, while significant, cannot form the basis of restrictions under Article 19(1)(a) unless explicitly included in Article 19(2).

The judgment critically examined the opinions of Justice Sawant and Justice Jeevan Reddy in Cricket Association of Bengal, which had entertained the idea of public interest as a ground for regulating access to airwaves, a public good. Chief Justice Chandrachud clarified that these observations must be understood in the specific context of broadcasting, where the use of public airwaves necessitated regulations to ensure plurality and fair access. However, he rejected any broader interpretation that would extend the scope of restrictions under Article 19(2) to include public interest or similar considerations in other contexts.

---

[19] Kaushal Kishore v. Union of India (2016) 10 SCC 295 (India).
[20] Association for Democratic Reforms v. Union of India (2024) 5 SCC 1 (India).

This decisive stance effectively dismissed not only Justice Jeevan Reddy and Justice Sawant's acknowledgment of broader considerations in Cricket Association of Bengal but also Justice Reddi's reliance on these ideas in PUCL. Chief Justice Chandrachud reiterated that the exhaustive nature of Article 19(2) does not permit implied grounds or flexible interpretations, even when justified by pressing societal concerns such as curbing black money or misinformation.

Consequently, these alternative interpretations now remain as academic exercises, devoid of legal standing. The judiciary's firm rejection of these approaches underscores a commitment to preserving the sanctity of the constitutional text and maintaining a clear and narrow framework for restricting free speech. This rigid adherence to Article 19(2) leaves no room for accommodating emerging challenges like misinformation within its existing structure.

Thus, the Bombay High Court's decision in Kunal Kamra v. Union of India, which struck down State Fact Checking Units on the ground that misinformation is not a recognized basis for restricting speech under Article 19(2), stands as a natural extension of this judicial philosophy. By reaffirming the exhaustive nature of Article 19(2), these rulings collectively highlight the constitutional protection afforded to misinformation, however troubling its implications may be. The alternatives proposed in earlier judgments have been relegated to the realm of unimplemented ideas, leaving Article 19(1)(a) as a robust, albeit rigid, shield for even the most contentious forms of expression.

Conclusion

The exponential rise of AI-generated misinformation poses a grave threat to Indian democracy, undermining public trust and distorting informed decision-making. Despite these dangers, India's well-established constitutional jurisprudence—most recently reaffirmed in Kunal Kamra v. Union of India—rigorously protects free speech under Article 19(1)(a), allowing restrictions only on the grounds explicitly listed in Article 19(2). Courts have consistently held that no additional grounds, including public or societal interest, can justify curbing free speech, as demonstrated by the Bombay High Court's rejection of State Fact-Checking Units on constitutional grounds. While this jurisprudence upholds the sanctity of free expression, it leaves the nation vulnerable to the unchecked proliferation of AI misinformation, which threatens democratic processes and societal harmony. To address this critical challenge, legally and constitutionally, a constitutional amendment is essential to include misinformation as a ground for restriction under Article 19(2).

SUBSTITUTION OF "JUDICAL MIND" WITH ARTIFICAL INTELLIGENCE: CHARTING CONCERNS FROM A HUMAN RIGHTS PERSPECTIVE

--Chetana Goud and Koustubh Sharma

The term 'judge' in legal parlance refers to any authority vested with the power to interpret and apply the law. The duty to act as a judge or 'judicially' arises whenever an authority is entrusted with an obligation to determine questions that affect the rights and liabilities of an entity. Such duty is not merely the power to conduct an inquiry rather a comprehensive power to hear a case, weigh evidence and consideration of submissions of both parties in order to finally settle a lis.[1] Generally, such duty is principally presided over by the Courts of law administered by judges. However, in certain situations, this duty has also been delegated to bodies which are non-judicial in nature such as administrative and government bodies. When the duty to act judicially has been delegated to an authority, which is not a Court in the ordinary sense, any decision made by such authority in respect of its aforesaid duty would be considered to be a quasi-judicial act.[2]

The Twenty-First Century has witnessed the adoption of technology-driven approach by several professional domains including the legal industry. The advent of Information Communication Technologies established the humble beginnings of this adoption, introducing to the field of law digital libraries, e-database, internet-based research tools, digital record storage devices etc. The most recent technological phenomenon to grip the legal field, particularly the judiciary is the use of Artificial Intelligence ("AI"). Although, there exists an ongoing debate on the most appropriate definition of AI, keeping in mind the context of the legal domain, a suitable definition is "An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action."[3]

In several jurisdictions, AI tools have found application and use in the judicial realm, for purposes such as but not limited to automated decision-making, commission of offence and risk prediction, investigative assistance, automated transcription. The most concerning aspect from the perspective of human rights, is the use of AI in replacement of a human judge. The substitution of the human 'judicial mind' with artificial intelligence poses several concerns impacting fundamental human rights such as right to privacy, right to a fair hearing, right to work and non-discrimination. Human rights are those which are essential for the protection and maintenance of dignity of individuals, they are inalienable rights which are inherited by birth. Therefore, any act that affects the dignity of an individual, such as a judicial decision falls within the scope of human rights, and any external fact that influences the nature of such act, such as the impact of AI tools, ought to be critically analysed from a human rights perspective.

Right to a fair hearing, compromised?

When an authority has been entrusted with the power to act judicially, it is obligated to perform two necessary actions. Firstly, such authority must give an opportunity to the parties to make their representations, weigh the evidence submitted, consider all submissions of fact and law before arriving at a decision. Secondly, apart from following a mere procedure that is 'judicial', common law jurisprudence extended the obligation of an authority vested with judicial power, to 'act fairly'.[4] For instance, it was held in A.K. Kraipak v. UOI, "the requirement of acting judicially in essence is nothing but a requirement to act justly and fairly and not arbitrarily or capriciously."

In light of this, an AI technology replacing a judge, would have to act judicially i.e., conduct all procedures considered judicial along with conducting those procedures in a fair manner. Here, it is pertinent to mention that any AI tool, delivers the function it is programmed to deliver, by its creator training it on large amounts of data. It identifies patterns and gives an output on basis of pattern-prediction. The question arises, whether an AI tool built on data analysis, can act judicially and fairly?

Art. 10 of the Universal Declaration of Human Rights (UDHR) guarantees to every individual the right to a fair hearing by an independent and impartial tribunal.[5] Art. 10 also implicitly reflects the principle of equality enshrined in Art. 7 UDHR.[6] Thus, a crucial effect of Art.10 is to ensure a fair trial which prohibits discrimination on all grounds mentioned in Art. 2 (race, colour, sex etc.) and other grounds specific to the context of a trial such as any inappropriate distinctions made on the basis of crime committed, gravity of offence, or relationship with claimant or accused/defendant.[7] Further, Art. 10 also enforces the right to be heard as an essential component of a fair trial.

In a trial conducted by an AI tool, both the right to be heard and right against bias stand compromised. In the United States, judges, probation and parole officers have been allowed to use algorithms to assess a defendant's likelihood of becoming a re-offender.[8] One such tool is the COMPAS[9] Needs and Risk Assessment Tool being used by the New York State Parole Board.[10] It assesses whether a defendant, or incarcerated individual shows likelihood of recidivism based on factors such as education level, age at time of conviction, their plans for re-entry into society etc.

---

1 Province Of Bombay vs Kusaldas S. Advani and Others, Opinion of Fazl Ali. J., 1950 AIR 222.
2 Id.
3 Kelly M. Sayler, Cong. Rsch. Serv., RL45178, Artificial Intelligence and National Security, 1-2 (2020).
4 Art.10, UDHR.
5 Art.7, UDHR.
6 David Weissbrodt and Mattias Hallendorff, Travaux Preparatoires of the Fair Trial Provisions--Articles 8 to 11--of the Universal
7 Declaration of Human Rights, 21 HUM. RTS. Q. 1061 (1999).
8 Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin, How We Analyzed the COMPAS Recidivism Algorithm, Pro Publica, May 23, 2016, https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.
9 COMPAS: Correctional Offender Management Profiling for Alternative Sanction.
10 Allyson Brunette, Humanizing Justice: The transformational impact of AI in courts, from filing to sentencing, Thomson Reuters, Oct. 25, 2024

Determining the likelihood of a person to commit future crimes on indicators like education and age in itself is a prima facie discriminatory practice, as it eliminates the right to a hearing and substitutes it with decisions that are devoid of socio-economic contexts. It violates another crucial right guaranteed by UDHR i.e., presumption of innocence until proved guilty[11] as an incarcerated person is being punished for a future offence which has neither been committed nor subjected to trial. Since AI relies on data for its functioning, any bias in the input data will result in biased outputs when processed by it. It was discovered that the COMPAS tool's accuracy was questionable, as an analysis of its decisions revealed that Defendants belonging to the black race were twice as likely as defendants of white race to be wrongly classified as high-risk for recidivism despite not reoffending within two years.[12]

Further, an AI tool inherently is a product of creation. Such tool would extend an IP right to its creator. In order to judge, one must be "independent". An AI tool's independence and transparency is highly questionable. Even if, a creator relinquishes their IP rights in an algorithm, the algorithm still would be a reflection of the creator's psyche, meaning a culmination of all biases and archetypes possessed by the creator. If an AI tool is expected to replace the 'judicial mind', it would be impossible to find a creator, who is able to program legal principles evolved over time immemorial without any personal biases of their own. For instance, the developers of UK's HART similar to USA's COMPAS in purpose, did not include race as an indicator for determining commission of repeat offenses.[13] However, that would also not be holistic decision-making, since judges often acknowledge positive discrimination due to intrinsic link between social factors and crimes. Complete exclusion or inclusion of certain factors while judging a case, is not a fair trial, since judges are not expected to give uniform decisions rather empowered to give a variety of reasonable, different responses within an ethical framework. Thus, the control of an AI by any other entity be it a corporate, an educational institution or a programme-developer, dilutes the principle of independence of judiciary.

AI and Employment Concerns

Several Indian High Courts have begun using AI tools in transcribing oral arguments and translating judicial documents, beginning from February 2023.[14] This project has already undertaken the transcription and translation of over forty-five thousand court orders in Hindi, as well as other regional languages.[15] While this move has been described by the Ministry of Law and Justice as a cost-cutting measure, it also focuses on making the Court's decisions more accessible to the common man. Yet, it may have greater ramifications.

Every technological development leads to the upheaval of the existing status-quo, and with it the security which is provided by the extant system. It has been approximated that the implementation of latest technology into the large-scale operations of most sectors would affect about 1.2 billion jobs, costing about 14.6 billion USD of salaries–[16] indicating the severe impact of transition to AI on an individual's right to work and free choice of employment.[17]

With the existing pitfalls of the 'machine learning' which has been made common within the erstwhile AI industry, the biggest concern remains the informational blind spot of the AI; and implementing relevant checks to ensure it does not hinder the ratio which has to be dispensed by the Court.

One of the principles propagated by the UNESCO for the ethical consumption of Artificial Intelligence is that of 'Human Oversight and Determination', intending to keep the steering wheel in the hands of humans, favourably professionals of the particular field, in order to ensure that the AI is being used in the desired direction. The Recommendation directs the Member States to ensure that the responsibility for the legal and ethical actions of AI tools can always be related to humans working in relation with such tools. This would ensure that any oversight would be attributable to the individual working in tandem with the AI tool, in order to reduce non-liability.

The Recommendation also suggests that the option for ceding control over such tools remains in the hands of humans themselves, and in connection with the use of such tools in decision-making processes. However, it must be ensured that AI systems must never replace ultimate human responsibility and accountability. "As a rule, life and death decisions should not be ceded to AI system".[18]

---

[11] Art. 11, UDHR.

[12] Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks, Pro Publica, May 23, 2016.

[13] V.A. Laptev & D.R. Feyzrakhmanova, Application of Artificial Intelligence in Justice: Current Trends and Future Prospects, Hum-Cent Intell Syst. 394, 394–405 (2024).

[14] Ministry of Law and Justice, Measures to Translate and Publish Proceeding and Judgments of Supreme Court and High Courts (Nov. 28, 2024), https://pib.gov.in/PressReleaseIframePage.aspx?PRID=2078399#:~:text=Measures%20to%20Translate%20and%20Publish,Supreme%20Court%20and%20High%20Courts&text=The%20Supreme%20Court%20of%20India,in%20translation%20of%20judicial%20documents.

[15] https://judgments.ecourts.gov.in/pdfsearch/index.php.

[16] Michael Chui et al., The Countries Most and Least likely to be affected by automation (2017) Harv. Bus. Rev. (Apr. 12, 2017).

[17] Art. 23, UDHR.

[18] UNESCO, Recommendation on the Ethics of Artificial Intelligence, 2022, ¶35-36

Therefore, the only ideal way of continuing with the usage of AI technology within the judiciary would consist of the jobs being structured in synchronicity with the technology, so that professionals can utilise such technology to its utmost. This would include limiting the usage of the AI to simply expediate the clerical tasks while ensuring there are no technical errors which pollute the proceedings, and which leave the humans in a position of scrutinising and determining the implementation of the works of the AI.

CONCLUSION

Humans are entitled to be judged by humans. Judgements are not merely an analysis of evidence and rules rather generously supplemented by human experience.AI, obviously shall become ubiquitous in the coming future, however it ought to be used to assist rather than replace humans.[19]

The integration of AI in the judicial realm albeit promotes efficiency, at the same time it necessitates careful navigation of concerns that strike at the heart of human life. The judiciary's role as a guardian of fairness and impartiality cannot be substituted by AI tools given their heavy dependence on, not so objective, data. While the technological progresses witnessed within the AI industry in the past decade must be lauded, they should also be perceived in a manner which does not threaten the largely human judicial rubric or the employment of great portions of the global population, and must certainly not be propagated with the vision of cutting costs in favour of the industrialists.

The future of AI lies in its use as a tool of assistance, which would ensure that the human element comprising empathy, ethical and holistic reasoning is preserved while embracing the fruits of innovation in dispensation of justice.

[19] M-RCBG Associated Working Paper No. 220, AI, Judges and Judgement: Setting the Scene, Rt Hon Sir Robert Buckland KBE KC MP, Harvard Kennedy School.

SAFEGUARDING PRIVACY AND DIGNITY IN THE ERA OF DEEPFAKES

-Annanya Sharma

Technology has a significant impact on the human life in terms of both intensity and complexity. With the advancement of civilization, man has become more sensitive to publicity making solitude and privacy a more essential part of an individual's life. However, modern enterprise and invention have, through invasions upon privacy, also subjected man to mental pain and distress that are far greater than those which could be inflicted by mere bodily injury.[1] This is an age where artificial intelligence is becoming capable here-and-now and has enormous capabilities to outsmart human intelligence.

The origin of deepfakes

The proliferation of AI and machine learning has both beneficial and detrimental effects. Several privacy and security issues are associated with Artificial Intelligence, yet countries and governments worldwide are investing and developing Artificial intelligence technologies. The interconnectivity of AI systems, which optimize every aspect of our lives, including our genomes, faces, finance, emotion and environment, has further added to the problem of privacy protection.[2] A notable consequence of this AI technology is the emergence of deepfakes which are based on machine learning algorithms that use face mapping software to produce fabricated content of an individual's identity without their permission.

The term 'deepfake' was initially coined by a Reddit user who used 'Face-swapping' technology to superimpose celebrities' faces on pornographic videos.[3] However, today, deepfakes have extended to creating convincing fake content, including videos, audio content, and images, which can be used to spread false or misleading information. Inaccurate depictions, which can be indistinguishable from genuine content, can be manipulated to defame, deceive, and engage in other forms of misconduct.[4] Deepfake technology's most alarming use is its potential to disrupt reputations and privacy. Misrepresentation of individuals in undignified or harmful contexts undermines trust and creates societal harm, necessitating robust legal and ethical frameworks.

Legal framework for deepfakes

The term 'deepfake' has been defined under Article 3(60)[5] of the European Union Artificial Intelligence Act, 2024 (EU AI Act) as 'AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful'. Article 50(4),[6] of the EU AI Act states that providers of AI technology creating or altering image, audio, or video content to produce deepfakes must openly acknowledge that the content has been synthetically generated or modified making them responsible for the same.

U.S. has enacted several federal legislations to address the concern of Deepfakes. The Identifying Outputs of Generative Adversarial Networks (IOGAN) Act, introduced in 2020,[7] was proposed to establishing a task force within the Department of Homeland Security to study deepfake technology and develop strategies to counter its harmful effects. The DEEPFAKES Accountability Act, 2023[8] also aims to protect national security against the threats posed by deepfake technology and to provide legal recourse to victims of harmful deepfakes. Lastly, the Defiance act of 2024[9] has been enacted with an aim of improving the relief system for individuals affected by non-consensual intimate digital forgery protection.

India's legal framework lacks specific definitions or measures to address Deepfake threats, but existing laws such as the Information Technology Act, 2000 [IT Act], the Bharatiya Nyaya Sanhita, 2023 and the IT Rules offer potential remedies.[10] Deepfake attacks fall under the purview of S. 66D[11] of the IT Act, which addresses punishment for cheating through personation using computer resources. The perpetrators of deep fake attacks could face imprisonment up to three years and fines.[12] S. 66E[13] of the IT Act of 2000 is also applicable in cases pertaining to deepfake offenses encompassing the capturing, dissemination, or transmission of an individual's visual representations through mass media, thereby infringing upon their right to privacy. S. 51[14] of the Indian Copyright Act of 1957, also establishes penalties for specific offenses related to copyright infringement.

---

[1] Mritunjay Kumar, Right to Privacy in a 'Posthuman World': Deconstructing Transcendental Legacies and Implications of European Renaissance in India, 1 SML L Rev 52 (2018).
[2] Simran A. Jain & Sunitha Abhay Jain, Artificial Intelligence: A Threat to Privacy?, 8.2 NULJ 21 (2019).
[3] Bhanusshre Sivaramachandran & Vaishnavi Kulkarni, Tackling the Multifaceted Legal Dilemmas of Deep Fake Technology, 4.3 JCLJ 217 (2024).
[4] Anuragini Shirish & Shobana Komal, A Socio-Legal Inquiry on Deepfakes, 54 CAL. W. INT'l L.J. 517 (2024).
[5] The European Union Artificial Intelligence Act 2024, art. 3(60).
[6] The European Union Artificial Intelligence Act 2024, art.50(4).
[7] The Identifying Outputs of Generative Adversarial Networks (IOGAN) Act, 2020.
[8] The DEEPFAKES Accountability Act, 2023.
[9] The Defiance act, 2024.
[10] Shinu Vig, Regulating Deepfakes: An Indian perspective, 17 JSS 79 (2024).
[11] The Information Technology Act, 2000, S. 66D.
[12] The Information Technology Act, 2000, S. 66E.
[13] Id.
[14] The Indian Copyright Act, 1957, S. 51.

Deepfakes, which encompass the unsanctioned manipulation or modification of photos and videos come under the unauthorized utilization of an individual's property. Section 318(4)[15] of the BNS deals with cheating and prescribes punishment with imprisonment up to seven years with fine for those found guilty of cheating. Section 356[16] of the BNS deals with defamation offenses, wherein individuals spreading false and damaging information about others could be subject to imprisonment up to two years, or fines, or both.

Lastly, The DPDP Act, 2023[17] which came after the Puttaswamy[18] decision protects personal data such as photos, videos, etc., and the deepfake makers who use such personal data could be liable or guilty of personal data breach as the confidentiality is being compromised. A deepfake can be covered under the definition of personal data because it can be used to identify the person being featured in it.

Deepfakes and the right to privacy

The right to privacy is a cornerstone of human rights and is recognized under UDHR and ICCPR. Article 12[19] of the UDHR provides that "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." Similarly, article 17[20] of the ICCPR, which has been ratified by 167 States, provides that "no one shall be subjected to arbitrary or unlawful interference with his or her privacy, family, home or correspondence, nor to unlawful attacks on his or her honour and reputation." Other regional bodies such as the EU, recognises the right to privacy as a fundamental human right in Article 8[21] of the Charter of Fundamental Rights of the European Union.

In India, the Supreme Court has, in a number of decisions, recognized the right to privacy as a subset of the larger right to life and personal liberty under Article 21 of the Constitution of India.[22] In K S Puttaswami v. Union of India the right to privacy was declared as a fundamental right guaranteed by Indian Constitution on various philosophical-juridical grounds, including the right to be left alone, the right to life with human dignity, the right to liberty, security, and autonomy, the right to have an identity, the right to anonymity, the right to repose and sanctuary, the right to make intimate decisions, the inalienability of rights, and the reasonable expectation of privacy, etc. Deepfakes infringe upon the personality rights of an individual, which are recognized under the right to privacy, due to their potential use in creating images, videos and audio that depict individuals participating in certain activities considered harmful, inflammatory or undignifying.

The Supreme Court has explicitly recognized publicity rights in the form of the right to privacy in the case of R. Rajagopal v. State of T.N.,[23] where it was held that the first aspect of violation shall be using a person's name or likeness for advertising without his consent. Similarly, In the case of Sonu Nigam v. Amrik Singh,[24] while recognising the celebrity rights of Sonu Nigam in promotional posters, the Bombay High Court granted him an injunction for false representations made against him in the said posters. Similarly, to combat deepfake pornography and deepfake parody videos created to disrupt the image rights of a celebrity, relief can be sought from the courts.

Measures taken by the government

The government has acknowledged the issue of deepfakes and undertaken several regulatory measures. On January, 2023, the Ministry of Information and Broadcasting issued a cautionary directive to media entities urging them to exercise prudence when disseminating content that may have been manipulated or tampered. MeitY and MHA keep close contact with social media platforms under the provisions of the IT Act, 2000 to effectively remove objectionable content. Initiatives such as 'promotion of fact-checking' and 'Information Security Education & Awareness' (ISEA) were introduced to educate the citizens to not to fall for and to refrain from sharing/spreading information.[25]

Recommendations

India needs to develop a comprehensive and effective approach to tackle the challenges posed by deepfakes. To defend an individual's right to privacy and dignity, countries around the world are waking up and battling against this menace caused by the deepfakes. Having specific legislation or provisions in our legal framework to address deepfakes is the need of the hour.

We can mirror the strategy adopted in the EU AI Act and the legislations enacted in U.S. to have a stringent regulation to govern high-risk AI before the video can be disseminated. The blockchain technology should be used more effectively to distinguish between authentic and manipulated content. It is also essential to mandate the creators and providers to obtain consent from the people being displayed in the videos and an obligation on the service providers to place a system to confirm the users' real identities and develop a database to detect illegal or false information. Intermediaries also need to verify the authenticity of the videos though content moderators and be fast enough to take down the manipulated content.[26]

---

[15] The Bharatiya Nyaya Sanhita, 2023, 318(4).
[16] The Bharatiya Nyaya Sanhita, 2023, 356.
[17] The Digital Personal Data Protection Act, 2023.
[18] K S Puttaswami v. Union of India, (2017) 10 S.C.C. 1.
[19] Universal Declaration of Human Rights, 1948, art. 12.
[20] The International Covenant on Civil and Political Rights, 1966, art. 17.
[21] The Charter of Fundamental Rights of the European Union, 2000, art. 8.
[22] The Constitution of India, 1950, art 21.
[23] R. Rajagopal v. State of T.N., 1994 S.C.C. (6) 632.
[24] Sonu Nigam v. Amrik Singh, MANU/MH/0517/2014.
[25] Mohit Kar & Shreya Sahoo, Deepfakes and its Iniquities : Regulating the Dark Side of AI, 5.1 NLUO SLJ 41 (2020).
[26] Supra note 3.

Conclusion

A person’s reputation and fame can transcend into damaging various rights of a person including his right to livelihood, right to privacy, right to live with dignity within a social structure, etc. The tarnishment, blackening or jeopardises of the individual's personality or attributes associated with the said individual are illegal.

Although India currently lacks legislative provisions with respect to the extent and application of the right of publicity and privacy in the context of deepfakes, judicial precedents such as Anil Kapoor v. Simply Life India and Ors, where the court exclaimed that there can be no justification for any unauthorised website or platform to mislead consumers by using a person's name, voice, dialogues, images in an illegal manner and the same cannot be permitted, still enforce the idea of fighting against such manipulation of identity.

Recently the Bombay High court recognized the necessity of amending the rule arises out of the concern of the government pertaining to an increase in use of social media as a communication medium which has a reach, unparalleled to any other medium of communication and the danger of spread of misinformation and fake information, the negative impacts of which present a real, clear, and specific danger to public order. The threat of disinformation and hoaxes has evolved from mere annoyance to warfare that can create social discord, increase polarisation and in some cases, even influence election outcome. State and non-state actors with geopolitical aspirations, ideological believers, violent extremists, and economically motivated enterprise can manipulate social media narratives with easy and unprecedented reach and scale. This dis-information now also has a new tool in the form of Deep fakes. Therefore, a collective effort between the governments, legal system and technology providers is needed to strengthen the safeguards against the misuse of deepfake technology in this digital age.

Convergence of Human Rights and Artificial Intelligence

-- Kushagra Seth and Deeksha Rao

Abstract

Today, the relativity of human rights and artificial intelligence is seen as conventional. Historically, both fields have existed separately, as any confluence between the rights movements and the development of technology would be deemed absurd and irrelevant. However, these hard-earned rights are posed with threats today due to the progress of artificial intelligence and simultaneously the inability of the institutions to catch up with such developments. While deliberating the threats, the uses and benefits of AI cannot be left untouched and have to go hand-in-hand in discussing the 'Convergence of Human Rights and Artificial Intelligence.' The essay will also discuss ethical, legal, and social considerations in this complex relationship.

Introduction

AI has effects that go deep and erode the barriers that protect the lives of the citizens, i.e., their rights. The convergence and the phases passed by both phenomena must be studied. Rights held by citizens have become more advanced from one generation to another, covering more substance and areas. For instance, the fight for the fundamental rights of equality and democracy has evolved to add to the basket the rights for data protection, privacy, etc., to keep up with the world's advancements. The confluence between the two will be discussed, including the interaction of artificial intelligence and different generations of rights. Here, 'generation of rights' is used to denote the rights secured dominantly by the citizens in that era. The passing of time/generation has not affected their relevance whatsoever.

1. First-Generation Rights: Civil and Political Rights:

The effects of the use of AI are two-fold, as discussed above. AI, in terms of civil and political rights, has the potential to enhance and degrade them. By providing citizens with (i) better access to justice with the help of AI tools, (ii) detecting electoral fraud, misinformation, or manipulation, (iii) analyzing any disinformation, (iv)providing data for advocacy, and (v) lastly, balancing freedom of expression, AI has helped strengthening such rights. AI has an ample influence on these rights, especially when the country is bent on digitalization and the use of technology to improve efficiency.

Another facet of incorporating the use of AI to establish civil and political rights is the risk of-(i) surveillance and privacy violations: These privacy violations with the use of AI-driven surveillance tools could lead to authoritarian practices by the government, such as suppressing dissent. (ii) manipulation and disinformation: Use of AI to spread misinformation, undermining the trust in democratic information. (iii) aggravating existing discrimination and bias: Biased algorithms could lead to further inequality in targeting or favoring something/someone over others.

2. Second Generation Rights: Social, Economic, and Cultural rights:

AI has helped the world in the social sector by aiding the health and educational sectors. The health sector has benefitted from using AI to improve diagnostic capabilities, predict disease outbreaks, help reach out to underserved areas, etc. Concerning the educational industry, AI has not only improved the quality of education but has also made it accessible to the population deprived of such rights. Social welfare programs have also improved owing to the use of AI.

AI has also enhanced the economic rights of the citizens- (i) creating new jobs and opportunities, (ii) helping banking platforms, and (iii) keeping records to improve efficiency- AI-powered systems can be used to optimize resources, improve supply chain management, and enhance productivity.
The benefits of AI have also reached cultural rights where AI can be used to preserve cultural heritage through digital archiving, language preservation, 3D Reconstruction, etc.
However, all these benefits are balanced out by the risks posed by the use of AI, such as (i) cultural homogenization leading to a threat to minorities, (ii) bias and inequality as discussed in the previous section, (iii) threat to the existing job opportunities, etc.

3. Third Generation Rights: Collective Solidarity Rights:

Collective Solidarity Rights, also known as third-generation rights, are human rights that emphasize and focus on the collective interests of groups or communities instead of individual rights. AI can enhance these rights by:

(i) Mediation and diplomacy: AI can back up peace negotiations between two parties, nations, etc., with data, analyze the conflicting grounds, recommend data-based solutions, and even help in language translation to facilitate such meets.
(ii) Conflict prediction and prevention: Various AI algorithms could help identify any conflict that might arise, and early intervention would prevent it based on social, political, and economic data.
(iii) Arms control: Satellite imagery and pattern recognition can help governments monitor control over arms and detect illegal weapon production.
(iv) Environment Monitoring and Cultural Preservation
(v) Helping in disaster response, etc.

The risks of AI concerning collective solidarity rights include- (i) conflicts due to misinformation: With the widening access of AI, it could be used by any person to ignite any conflict between two parties, (ii) weaponization of AI: AI has also helped the defense sector but has simultaneously made itself a threat, etc.

4. Fourth Generation Rights: Rights against challenges posed by AI:

These rights are not officially codified in any law. These rights address the challenges and opportunities arising out of rapid technological advancements. These fourth-generation rights focus on digital technology, artificial intelligence, genetics, etc. Some examples of these rights include:

1. Right to Digital Privacy and Data Protection
2. Right to Access and Control Over Technology
3. Right to Cybersecurity
4. Right to AI Accountability and Transparency, etc.

The irony lies here in the fact that to protect the rights of citizens against the use of AI, a well-defined mechanism of laws along with AI has to be incorporated. Without the proper knowledge and use of AI, no protection against its misuse can be granted or guaranteed.

Ethical Consideration:

The intersection of artificial intelligence and human rights is a growing area of concern and opportunity in contemporary society. With AI technologies advancing rapidly, they pose risks and benefits to fundamental human rights, and thus, there is a need for critical examination of the implications. This raises enormous ethical concerns about its impact on human dignity and autonomy. While AI may facilitate better decision-making, it may also infringe upon privacy rights, result in discrimination, and further deteriorate the personal sense of agency. For example, the application of AI in surveillance or social scoring is bound to affect already marginalized groups of people more than others. It is thus more likely to entrench systemic biases that subvert equality. Ethical deployment of AI requires a framework that defines human rights and ensures that technology promotes freedom rather than curtails it.

It requires particular attention to be paid to vulnerable groups such as women, children, older people, and persons with disabilities. Such groups are highly vulnerable to the applications of AI, as their specific needs are ignored or they exacerbate the existing inequalities. AI systems must be designed to make them more inclusive to safeguard the rights of these groups.

The pace of development of AI technologies has been faster than the legal frameworks that exist today, leaving a gap in human rights protections. The EU AI Act is one such legislative effort that addresses these challenges by establishing guidelines for the ethical use of AI and its potential human rights impacts. Such regulations must include accountability, transparency, and oversight provisions to prevent misuse and ensure compliance with international human rights standards.

This complicates accountability for the actions of increasingly autonomous AI systems. Existing legal frameworks must evolve to hold developers and operators responsible for human rights violations through AI decision-making. Lines of responsibility must be established when AI systems produce harmful outcomes or perpetuate biases.

Public trust is the basis for social acceptance of AI technologies.[1] When individuals perceive AI as a threat to their rights or privacy, resistance to its adoption increases, and engaging communities in discussions about the implications of AI can foster understanding and support for responsible innovation. Additionally, promoting transparency on how AI systems function and what impacts they may bring is essential to establishing trust among users.

The global nature of technology means that international cooperation must occur to mitigate AI's challenges. Collaborative efforts between countries will ultimately create universal standards that are safe for human rights and allow technological advancement. For instance, the Council of Europe's Ad Hoc Committee on Artificial Intelligence (CAHAI) strives to develop binding and non-binding legal instruments that safeguard democracy and human rights while deployed with AI.[2]

Conclusion

The convergence of artificial intelligence and human rights poses challenges and opportunities that need urgent attention from policymakers, technologists, and civil society. By prioritizing ethical considerations, enhancing legal frameworks, and fostering public trust through transparency and engagement, AI's benefits can be harnessed while protecting fundamental human rights. As we move across this complex topography, there is an imperative need not to let technological progress engulf all that we hold dear in shared values and freedoms.

---

1. Trust in artificial intelligence, 2023 global study on the shifting public perceptions of AI, KPMG., https://kpmg.com/xx/en/our-insights/ai-and-technology/trust-in-artificial-intelligence.html.
2. The CAHAI fulfilled its mandate (2019-2021) and has been succeeded by the Committee on Artificial Intelligence (CAI), CAHAI - Ad hoc Committee on Artificial Intelligence.

Surveillance vs. Democracy: A Threat to India's Democratic Identity

-- *Aayush Rana*

Privacy is a fundamental right that protects human dignity and is essential for a just and respectful society. It serves as a cornerstone and is important to protect and safeguard other key freedoms, such as the right to freedom of expression and association. Today, almost every part of the world acknowledges the fundamental right to data protection. This right focuses on protecting the private information of individuals and organisations from other governmental or non-bodies and ensuring it is handled with care and respect. Closely linked to the right to privacy, data protection is often looked upon as an important component of privacy rights with respect to the framework of the United Nations human rights system.

Data analysis using AI systems may reveal private information about individuals, which qualifies as secured information and should be treated as sensitive even if derived from big databases fed from publicly available information. An example of the thin line between public and private data is the increased use of government social media monitoring programs, wherein law enforcement agencies collect troves of social media information and feed it to AI-powered programs to detect alleged threats. While isolated checks of a target's public social media may seem to some like a wise policing strategy, these programs instead will involve massive, unwarranted intake of the entire social media lifespan of an account, group of accounts, or more. Bulk collection of this type has been found to inherently violate human rights. Moreover, Government surveillance has expanded with the growth of the internet and the development of new technologies, and AI is enabling more invasive surveillance tools than ever.

A recent article by BBC, titled "How Pegasus' Snooping Threatens Indian Democracy,"[1] sheds light on the implications of government surveillance through spyware software like "Pegasus". The report highlights how the government, citing national security and counter-terrorism measures, uses software to intrude into the lives of its citizens. This raises a big question: can the fundamental right to privacy be compromised in the name of protecting the right to life? Moreover, it raises a serious concern as to whether the people of India are willingly trading their privacy for a sense of security, placing unwavering trust in the government to safeguard their lives.

This debate creates a balance between individual freedoms and collective security, challenging the very basis of democracy and the trust between a government and its people. In India, the evolution of the right to privacy has been very slow and concerning. The framers of the Constitution, at the time of its inception, did not deem it necessary to explicitly include a provision securing the privacy of citizens. It was only after 69 years and numerous legal battles that the Supreme Court of India, in the landmark Puttaswamy v. Union of India case,[2] recognized privacy as a constitutional and fundamental right. This historic judgment firmly established privacy under Article 21 of Part III of the Indian Constitution, marking a significant milestone in the country's legal and democratic framework.

How is the use of spying software like Pegasus killing the very essence of democracy and why it should be stopped?

According to reports, a leaked confidential list having 50,000 phone numbers, potentially subjected to Pegasus surveillance, was uncovered by Forbidden Stories. Among these, 300 numbers were linked to India. Alarmingly, the list reportedly included prominent figures such as Mr Rahul Gandhi, the Leader of the Opposition in Lok Sabha; Mr Ashok Lavasa, a former Election Commissioner who questioned Prime Minister Narendra Modi's alleged poll code violations during the 2019 general election; Mr Alok Verma, the ousted Chief of the Central Bureau of Investigation (CBI); and Mr Umar Khalid, a student activist later charged with sedition. The targeting of opposition leaders, independent officials, and even university students signifies a direct assault on free speech, and the checks and balances essential to a healthy democracy. By employing spyware to create surveillance over the political activities and civil society members, the government risks disturbing the foundations of democracy, bringing it to a system that oppresses opposition voices and acts as an authoritarian state. The citizens of India must recognize the implications of such practices. Spying on opposition leaders, activists, and organisational heads is not just a violation of individual privacy but a big blow to the democratic process altogether. This misuse of heightened surveillance tools silences dissent, weakens democratic institutions, and destroys the trust and support that is vital for healthy and good governance. For democracy to survive and grow, it is important to stop the misuse of spyware like Pegasus and ensure that a proper legal framework is made that safeguards privacy and uphold the democratic rights of all citizens.[3]

The over-surveillance practices of the Indian government bear an unsettling resemblance to those seen in regimes like Taliban-ruled Afghanistan. In Afghanistan, individuals particularly those whose livelihoods or personal circumstances made them potential targets—were forced to frantically delete or alter their social media profiles, aware that the Taliban would likely use surveillance to suppress dissent and consolidate power. Ironically, while India positions itself as a big critic of the atrocities committed by authoritarian regimes, its increasing reliance on surveillance tools risks pushing the nation toward the very practices it condemns. This troubling trend undermines the democratic values India stands for, creating an enviornment where dissent and freedom of expression are obliterated in the name of national security.

---

1. Pegasus: Why unchecked snooping threatens India's democracy, Soutik Biswas, BBC News, 20 July 2021, https://www.bbc.com/news/world-asia-india-57887300.
2. (2017) 10 SCC 1, AIR 2017 SC 4161
3. IT Minister Shri Ashwini Vaishnaw's Statement in Parliament on "Alleged use of spyware Pegasus to compromise phone data of some persons as reported in Media on 18th July 2021", 19 JUL 2021 4:42PM, https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1736803.

77 years of independence, where do we stand today?

Astonishingly, we find ourselves fighting against our own government for the most basic right to privacy. This is not what our leaders and freedom fighters had for a nation that is home to 17% of the world's population. India's path to be the Vishvaguru is surely full of thorns when its own govermnent is using technology to opress the rights of indivisuals.

The current state of affairs is deeply alarming and calls for an urgent judicial intervention. At the same time it is important for citizens to become more aware about their rights.This awareness is not just about protecting ones privacy but is about protecting the very foundation of Democratic identity that India has carried over the years. If left unchecked, the opression of these rights could pave the way for India to transform into an authoritarian regime similar to Taliban ruled Afghanistan or China. This is a turning point for the nation a moment to uphold the principles of freedom and democracy that countless individuals sacrificed their lives to achieve.

THE ROLE OF PRIVACY AND ETHICS IN SHAPING THE DIGITAL ERA

-Devang Bhardwaj

Data ethics refers to the use of data in accordance with the wishes of the people whose data is being collected. But the omnipresence of digital technology in our daily life, its use and its impact on organisations and individual;[1] raises ethical questions about its role in our society. These concerns include consent and privacy, security, inclusion and fairness, protection from online harm, transparency and accountability. These issues underscore the need for stringent safeguards to ensure that technology operates within ethical and legal boundaries.[2]

Key Ethical Challenges in the Digital World

1**. Consent and Privacy-** Ethics of data emphasizes understanding, autonomy, and security. But, collection and processing of personal data happens at times without the knowledge and acceptance of the users. For example, India's Aadhar scheme, which aimed to facilitate easier access to government services, was said to violate an individual's right to privacy owing to collection of biometric data. In a momentous judgment the aforementioned case was named as Puttaswamy case (2017)[3] whereby the Indian Supreme Court ruled that privacy is a basic aspect of right to life and liberty and therefore is a fundamental right provided under Article 21 of the Constitution of India. This pronouncement reiterated previous decisions that privacy is an integral part of the fundamental freedoms protected under Part III of the Constitution. Of late, Property Owners Association and Ors. Vs. State of Maharashtra and Ors, 2024[4] has also reiterated the principle of privacy which was promulgated in the Puttasawamy case.

2. Security and Mitigating Misuse- Invasion, digital platforms are increasingly under the threat of breaches, surveillance, and manipulation rooms. Despite the threats in any of the aforementioned avenues, it is true that a lot is derived from the digital services consumers engage themselves in. For instance, TikTok has over the years been a popular app in India, however in the year 2020, India banned the app owned by the Chinese grime App, ByteDance, because of fears concerning data's protection as well as the country's national security. They included measures such as prohibition on cross border data transfers and unauthorized use of other sensitive cross border data.

3. Transparency and Accountability- Algorithms and artificial intelligence come up with digital systems with lack of transparency which prevent identification of bias, mistakes, or even unfair practice. Such a silence can abuse inequalities, inhibit possibilities, and breed misconceptions about technology. For example: ChatGPT although famous and quite useful, at times engages in providing inaccurate information to the users which is a concerning factor.

4. Online Harm and Social Impact- With the aid of technology, some forms of harm can be aimed for, that is through the use of social media. It is crucial for platforms to serve users of these social networking sites so as to protect users from such violence as hate speech and bullying but also protect their diversity. For instance: Young people use Instagram for creating accounts to enjoy some features only to be assaulted for sexual abuse, bullying, etc.

India's Legal Framework: Digital Personal Data Protection Act, 2023

India's response to these challenges has seen the implementation of the Digital Personal Data Protection Act, 2023 (DPDP Act).[5] Such a law provides a framework for protecting data privacy and the ethical handling of information as the world becomes more digital. Some of the provisions include Section 2(u), which explains "personal data breach" as such unauthorized processing of personal data or more specifically, such as an invasion of privacy, accidental loss, exposure, gain, sharing, utilization, alteration or destruction of personal data which poses a risk in the confidentiality, integrity or availability of that data. The concerns in this case further include – privacy of an individual, provisions for cyber security, restriction to the abuse of data, and liability etc. This Act does enable India to comply with the normal international requirements for privacy protection and at the same time deal with local issues, again a case of striking a nexus between growth and regulatory aspects. Recently, Telecommunications (Regulatory Sandbox) Rules, 2024[6] mentioned to ensure compliance by itself and by every participant with the provisions of the DPDP Act.

International Perspectives on Privacy

Privacy rights have been recognized globally as fundamental to human dignity and autonomy. Article 12 of the 1948[7] Universal Declaration of Human Rights asserts that: "No one shall be subjected to arbitrary interference with his privacy, family, home, or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks." This international acknowledgment underscores the need for nations to align their legal frameworks with these principles, ensuring that technological advancements do not undermine individual rights.

---

[1] 23rd Biennial Conference of the International Telecommunications Society (ITS): "Digital societies and industrial transformations: Policies, markets, and technologies in a post-Covid world, Ethical issues in digital technologies, 21st-23rd June, 2021. https://digitalprivacy.ieee.org/publications/topics/digital-ethics-and-privacy-technology-how-to-ethically-manage-data

[2] K.S. Puttaswamy v. Union of India, AIR 2017 SC 4161

[3] Property Owners Association and Ors. Vs. State of Maharashtra and Ors, 2024 INSC 835

[4] The Digital Personal Data Protection Act, 2023.

[5] Telecommunications (Regulatory Sandbox) Rules, 2024 - DRAFT - 27.11.2024 - Ministry of Communications : MANU/MCOM/0104/2024

[6] The Universal Declaration of Human Rights (UDHR), 1948.

[7]

Thus, History has witnessed several waves of technology but what the recent technological advances have had a bearing on is the nature of data around us, and more so, on our digital identity. Misuse of the technology has led to ethical discussions and questions. In India, the Aadhaar case, the ban of Tik Tok, the implementation of the DPDP Act, demonstrate how the country is trying to navigate these problems while addressing the challenges of globalization. However, there still remain disagreements on how privacy rights on the internet should be regulated. These disagreements should be placed in the context of local self-regulation, in the face of techno libertarianism, and be actively dealt with through assertive vires policy-based approaches.

SPECIAL THANKS

SAUMYA RAJPAL	CONVENOR	IV th Year
ZORAWAR SINGH RATHORE	CO-CONVENOR	IV th Year
ANNANYA SHARMA	SECRETARY	III th Year
AKASH SHARMA	SOCIAL MEDIA HEAD	III th Year
MANISHI	MEMBER	III th Year
SHASHANK SAHAI	MEMBER	III th Year
ABHYUDYA SINGH RATHORE	MEMBER	II nd Year
PALAK THAKUR	MEMBER	II nd Year
UMANG NITHARAWAL	MEMBER	II nd Year
SUMAN SIDDH	MEMBER	II nd Year
DISHA CHAUHAN	MEMBER	II nd Year
DYUKSHA CHAUHAN	MEMBER	II nd Year
SALONI SHARMA	MEMBER	II nd Year
AASTHA GOYAL	MEMBER	I st Year
MANTHAN PATHANIA	MEMBER	I st Year