ELECTIONS IN ALGORITHMICALLY ALTERED SPACES- THE CHALLENGE TO DEMOCRACY

Neelu Mehra[*]
Monika Yadav[**]

Abstract

The increasing algorithmic mediation in electoral processes has transformed democratic participation into a data-driven phenomenon, raising profound normative and constitutional concerns. The use of algorithmic recommender systems and predictive analysis for behavioral targeting, micro-profiling, hyper-nudging, and manipulative communications to exploit cognitive vulnerabilities of voters in the digital realm has eroded the foundations of democracy, *i.e.,* free and fair elections. This deliberate cultivation of epistemic chaos through hyper-personalized disinformation and the malicious use of generative AI has undone the decades of good carried out by the Supreme Court and various high courts towards making informed voters and deliberative democracy. The extant Indian legal framework and advisories of the Election Commission remain ill-equipped to address algorithmic influence operations. Tracing the *sine qua non* relation between democracy and information, the paper highlights how algorithmic manipulative tactics deployed by stakeholders unmakes voters' informed choices. Drawing upon the global regulatory responses, in particular the EU Artificial Intelligence Act and United States AI legislation, the paper evaluates normative strategies, constitutional safeguards, and techno-legal interventions for India. It argues for a model that combines categorization of algorithms, platform accountability, and robust data protection under the Digital Personal Data Protection Act, 2023, to ensure electoral integrity in algorithm infested democracy.

I Introduction

THE IDEA of democracy, as it stands today, is a method of collective decision-making characterized by equality among participants at an essential stage of the decision-making process.[1] However, in modern-day democracy, this collective decision-making has largely been limited to select occasions, wherein the collective will of people[2] determines the outcomes of these occasions. A few of these occasions are legitimation of political authority, shaping policy formulations and promises, referenda and plebiscites, judicial interpretation of public morality, participatory governance,

[*] Professor of Law, University School of Law and Legal Studies (USLLS), GGS Indraprastha University (GGSIPU), New Delhi, India.

[**] Doctoral Scholar, University School of Law and Legal Studies (USLLS), GGS Indraprastha University (GGSIPU), New Delhi, India.

[1] Democracy, *available at*: https://plato.stanford.edu/entries/democracy/ (last visited on Aug. 05, 2025).

[2] Though debatable, for both 'will' does not exist as such (from the perspective of Free Will) and it's not collective will, but will of majority, which again drops down from majority, once multi-party system comes into picture.

narrative formation, and national solidarity. In the context of India, speaking strictly, the people's collective will, which shapes the decisions and decision-making process of the state's functions, has been relegated to mere participating and voting in the elections.[3] It is through participating in elections as the electorate (and seldom as a candidate) that people collectively decide the course of their own and their nation's future.

In a democratic setup, their participation in elections, choices they consider and make while selecting a candidate to vote for, and most importantly, the information they rely on while making their *informed* choices, establishes their agency as rational individuals. The key to choice architecture, *i.e.,* information, determines what sort of choices an individual is going to make and to what extent her rational and logical cognitive faculties will be deployed in narrowing and acting on the choices. Now that access to information has been democratized, the real question zero-downs to the quality of information that envelops an individual.

The process of decision-making by an individual, whether to participate in the process of democracy or any other cognitive process, inevitably demands access to information, and once an individual has access to information, thereupon, at the cost of her cognitive faculties, she can mould it to her utilization.[4] At this juncture, two aspects come into play in determining the quality of an individual's decision: first, the stage of development of cognitive faculties in terms of genetics, social and cultural interactions, education, economic status, and other interplay of myriad factors. Matured the cognitive faculties, better the quality of decisions that one makes.[5] Second comes the information levels, *i.e.,* low-quality and high-quality information.[6] Low-quality information here means information that an individual acquires in close-knit environments, where either the information in circulation is merely an echo of one's existing beliefs or lacks dimensions, thereby denying one the ability to think in multiple contexts for the issue at hand. In contrast, high-quality information is acquired through quality education, evoking critical and rational thinking, developing ideas, informed peers, and the ability to decipher abstractions. In the decision-making process, usage of low-quality information involves less spent of cognitive faculties and therefore an

[3] This shift can be seen as a move from a participatory democracy to a representative democracy that is not functioning effectively. In a truly participatory system, citizens have avenues to influence government beyond the ballot box, such as through public hearings, referendums, and direct citizen assemblies. The Indian system, while representative, often lacks these robust mechanisms for continuous public input.

[4] The Carter Centre, "Access to Information- A Key to Democracy" (2002).

[5] Daniel Kahneman, *Thinking, Fast and Slow* (Penguin Books, London, 2011). Daniel Kahneman shows how System 2 thinking (slow, deliberate) is associated with cognitive maturity and yields more accurate judgments than impulsive System 1.

[6] For the issue at hand, the use of low- and high-quality information has no relation to low- and high-level information as used in the terms of computer programming languages.

easy and ready choice for brain to make. Meanwhile, the use of high-quality information in arriving at decisions takes a toll on cognitive faculties, to such an effect that brain, unless tuned to use such information, tries to evade this route.[7]

Elections and voting, the fundamentals of democracy falls under low-information rationality[8], for the fact that in modern democratic societies, with a complex labour division, people have a scarce amount of time, money, and energy to expand on politics and therefore must conserve on information gathering costs.[9] Further, factors like brain being cognitive miser[10], volumes of political data, media simplification, rational ignorance[11], heuristics over-reliance, and institutional trust deficits strengthen the case for reliance on low-information rationality. Cumulatively, these factors led to plunge in cognitive entropy and reinforces cognitive biases. It is at this juncture, a voter can be swayed in the desired direction, using highly sophisticated and targeted use of micro targeting and hypernudging algorithms.

These algorithms deployed on social media platforms targets a user, which in this case also happens to be a voter, with highly curated information specific to the parameters collected on that user.[12] Now, as discussed above, the reins of individual choice essentially are in the hands of information curators. The finer the parameters on an individual, the more curated the information, and eventually more potential for skewing one's choices. The picture gets grimmer as social media platforms are becoming the largest source of 'information' for billions of users.[13] Now, using recommendatory and predictive algorithms, even if 0.01 per cent of platform users are swayed in favour of either party or candidate, the effect on results will be drastic,

[7] The Cognitive ease v. Effort- The human brain often prefers low-effort cognitive strategies (what Daniel Kahneman calls *System 1 thinking* in his book Thinking fast and slow)- fast, intuitive, and based on low-quality or readily available information (*e.g.,* stereotypes, heuristics, media headlines). While using high-quality information usually requires *System 2 thinking*-slow, analytical, deliberate, which is cognitively taxing and often avoided unless necessary.

[8] Thomas Christiano, "Algorithms, Manipulation, and Democracy" 52 *Canadian Journal of Philosophy* 116 (2022).

[9] Anthony Downs, *An Economic Theory of Democracy* (Harper & Row Publishers, New York, 1957).

[10] Riccardo Viale (ed.), *Routledge Handbook of Bounded Rationality* 196 (Routledge, New York, 2021).

[11] A concept from public choice theory, rational ignorance suggests that individual voters have little incentive to deeply educate themselves on political issues because they believe that single vote is unlikely to change the outcome and the cost of being informed outweighs the personal benefit. So, it becomes rational to stay uninformed or semi-informed and rely on cues.

[12] Liang Zhang, Katherine Jijo, *et.al.*, "Enhancing Large Language Model Performance to Answer Questions and Extract Information More Accurately", *available at*: https://arxiv.org/html/2402.01722v1 (last visited on Aug. 05, 2025).

[13] Pew Research Centre, "Social Media and News Fact Sheet" (2024).

particularly in democracies where winning percentage historically remains below 1-2 per cent.[14]

The amalgamation of algorithmic recommendations on social media platforms, low-information rationality, legal vacuum, and large demographic base makes the case of India quite interesting. While under Article 19(1)(a), a citizen have freedom of speech and expression, which essentially requires information to exercise the said right, and it is through this article, a citizen among other, makes *informed* choice and act on these choices (choosing and voting for a candidate during an election, being one such exercise of Article 19(1)(a)). Now, if the information based on which choices are formed is skewed to suit the person's heuristic biases, the resultant manifestation will be nothing less than a sham exercise. Given the plethora of information generated, consumed, and shared on social media platforms, the extant regulatory framework[15] remains distant in addressing this latent challenge. These algorithmically determined communications operate on the sheer scale of data, targeting people with precision, and their swift upgradation to new sets of parameters are challenging the foundations of democracy in unforeseen ways.[16]

II Democracy and information- The *sine qua non* relationship

The key operationalizing factors for the abstract idea of democracy are codification of legal and constitutional framework, regular electoral processes, free press and information flow, and creation of democratic institutions. While all factors hold weight in making a democracy resounding, the role of information has been time and again brought to the spotlight by courts, institutions, and civic society. The role of information assumes significance in a representative democracy, in selecting and retaining those politicians who are competent, hard-working, and effective in representing their constituents' desires.[17] In practice, elections appear to generate a vast amount of information as candidates compete for office through speeches, debates, advertisements, media engagements, canvassing, and the sheer scale of activities on social media platforms. Candidates make statements about their competency, integrity, trustworthiness, and, of course, their positions on issues. Further, the stress

[14] Carissa Boerboom, "Cambridge Analytica: The Scandal on Data Privacy" *Augustana Center for the Study of Ethics Essay Contest* (2020).

[15] Sec. 66C and D of Information Technology Act, 2000; 123(4) of The Representation of the People Act, 1951; 171G, 465, 469, and 505 of Indian Penal Code, 1860, and Paragraph (I)(2) of the Model Code of Conduct.

[16] Gloria Marchetti, "The Role of Algorithms in the Crisis of Democracy" 6 *Athens Journal of Mediterranean Studies* (2020); See also, Masabah Bint E. Islam, Muhammad Haseeb, *et.al.*, "AI Threats to Politics, Elections, and Democracy: A Blockchain-Based Deepfake Authenticity Verification Framework" 2 *blockchains* (2024).

[17] Joseph E. Harrington Jr., "Modelling the role of information in elections", 16 *Mathematical and Computer Modelling* 133 (1992).

on the information is for the fact that, in choice architecture, information operates as a cognitive force that shapes the contours of autonomy, agency, and deliberative judgment. Further, it is access to and quality of information which determines 'whether an individual is exercising her agency through reflection or is merely juggling with the illusion of choice?'

Information, when curated transparently, empowers rational self-governance; else, distorts volition, reducing choice to mere illusion. In sum, information is not passive data, but an epistemic filter through which an individual interprets reality, controls alternatives, and constructs contexts. The essence of the very idea of 'I think, therefore I am' lies in access to information, and more importantly, access to uncurated and unfiltered information, to have real thinking, not mere illusion of thinking, where outcome is pre-determined.[18]

The access to uncurated and unfiltered information is essential to have autonomy and agency, when looked from an individual's perspective. But when an individual's cognitive autonomy and agency are staked on information and her functional part in a democracy as a voter are clubbed together, the amalgamation becomes worthy of exploration. In a democracy, information simultaneously has twin roles of foundational and disruptive power. As a foundational power, it is a great equalizer, enabling rational autonomy of all citizens; on the other hand, as a disruptive power, it unsettles the complacencies of power. As an equalizer, individuals, through access to credible, pluralistic, and intelligible information transcends the passivity of subjects and become reflexive agents capable of deliberation, dissent, and decision, thereby ensuring that democracy remains not only procedural, but is epistemological- where demand is not just voting, but knowing.

However, in a political economy, 'knowing' seldom bears the fruits of success for political candidates and party.[19] And therefore, this calls for information commodification, algorithmic curation, and ideological filtration to ensure a manufactured consent. With this, democracy, which should be a system of collective

[18] When information is filtered, curated, or controlled, it limits the perspectives and facts available to an individual. This can lead to a state where people believe they are thinking critically and forming their own conclusions, when in reality they are just processing a predetermined set of inputs. This creates a false sense of autonomy in thought. Real thinking involves genuine critical analysis, synthesis of diverse ideas, and the potential to arrive at an unexpected conclusion. The illusion of thinking, on the other hand, is a process where the "thinker" is led down a specific path, and the conclusion they reach is the one intended by the information provider.

[19] Dave Ellenwood, "Information Has Value": The Political Economy of Information Capitalism", *available at*: https://www.inthelibrarywiththeleadpipe.org/2020/information-has-value-the-political-economy-of-information-capitalism/ (last visited on Aug. 9, 2025); see also, Govindan Parayil (ed.), *Political Economy and Information Capitalism in India- Digital Divide, Development Divide and Equity* (Palgrave Macmillan, New York, 2005).

self-rule, anchored on collective will formed through autonomy and agency, plunges into a spectacle of illusion and manipulation. Therefore, in a democracy, the integrity of information is not a peripheral concern, but the very living condition for epistemic symmetry and cognitive enfranchisement- the twin pillars of meaningful democracy.

To address the issue of access to *information*[20] in a democracy for the sake of free and fair elections, the Supreme Court of India, through a series of judgements, interpreted the right to information as integral to the democratic process. To begin with, in *Union of India v. Association for Democratic Reforms*[21], the court while addressing the issue of "Whether voters have the right to know about candidates' background, including criminal records, assets, and educational qualifications", held that, "voters have a fundamental right to know the antecedents of electoral candidates under Article 19(1)(a)" and directed the Election Commission of India (hereinafter ECI) to compel candidates to disclose their criminal records, educational qualifications, assets and liabilities, and pending criminal cases.

The court established the right to information as part of the right to vote, making it a constitutional imperative in electoral democracy. Later, in *People's Union for Civil Liberties (PUCL)* v. *Union of India*[22] court while expounding on the validity of a Representation of the People Act (Amendment)[23] that sought to override the 2002 ADR judgment, firmly held that voters have the right to information as a fundamental right. The court reasoned that the secrecy of the ballot and the right to know are not contradictory but complementary in a democracy, and therefore, the amendment that diluted mandatory disclosure needs to be struck down. In effect, both cases strengthened electoral transparency and held that uninformed choice is no choice at all. Again, in *Resurgence India* v. *Election Commission of India*[24], the court ruled that non-disclosure or incomplete disclosure of information in election affidavits can lead to rejection of nomination, ensuring voters' right to complete and accurate information. Taking a stride further, in *Lok Prahari* v. *Union of India*[25], the court held that candidates, spouses, and dependents must disclose sources of income, with rational that transparency curbs misuse of power and disproportionate wealth accumulation.

With these case laws, the jurisprudence is now well established that voters have right to information under Article 19(1)(a), regarding the antecedents of candidates and

[20] The emphasis of access to information was relating to credentials of a candidate, in particular criminal antecedents, assets, education, *etc.* not on other incidental and ancillary facets of information like funding of political rallies, canvassing, third-party virtual campaigning, etc.

[21] *Union of India* v. *Association for Democratic Reforms* (2002) 5 SCC 294.

[22] *People's Union for Civil Liberties (PUCL)* v. *Union of India* (2003) 4 SCC 399.

[23] The Representation of the People (Second Amendment) Act, 2003.

[24] *Resurgence India* v. *Election Commission of India* (2014) 14 SCC 189.

[25] *Lok Prahari* v. *Union of India* (2018) 4 SCC 699.

political parties. On a closer scrutiny of 'Right to Information', one finds that the availability of information ensured by the court is just the beginning of steps in the direction of cognitive enfranchisement. Thus laudable, but need to desist from complacency at this juncture. As the information available in the public domain by means of election affidavits filed by candidates and political parties, before and after the elections, presents true information, but the problem was never in this set of information.

The real problem lies with the information disseminated by candidates, political parties, influencers, and anonymous entities during the course (and even before) of political campaigning, in the virtual space. And with the changing societal dynamics and masses readily accessible in virtual spaces (better called platformized polity), campaigners are shifting to virtual campaigning, thereby making digital platforms the new battleground for political campaigns. This can be buttressed by the surge in amount spent by political parties on online advertisements *via* Google Ads, Meta, and X (formerly Twitter).

Unfortunately, despite the vast amount of verbiage in elections, it does not immediately follow that voters are well-informed. The immense information in virtual space only creates the clutter which only happens to confuse the voter and detracts her from real issues[26]– a byproduct of information incoherency.[27] As established in preceding sections, the epistemic architecture of democracy demands for voter's rational agency, which eventually hinges on the coherence and intelligibility of available information. However, modern algorithmically driven ecology of social media platforms, information is no longer transmitted as a linear or hierarchically ordered stream but as an incessant deluge- fragmented, sensationalized, and often contradictory.[28] These transmissions using well-documented faults in cognitive faculties of humans[29], ensure that voter, instead of indulging in reasoned deliberation, becomes susceptible to heuristic shortcuts, biases, and polarizations.

[26] Clever Reach, "More options, more confusion – the paradox of choice. Psychology meets email marketing", *available at*: https://www.cleverreach.com/en-de/push-magazin/email-marketing-strategy-tips/more-options-more-confusion-the-paradox-of-choice-psychology-meets-email-marketing-part-5/ (last visited on Aug. 11, 2025); see also, David Bawden and Lyn Robinson "Information Overload: An Introduction" *Oxford Research Encyclopedia of Politics* (2020) and European Commission, "Study on the impact of new technologies on free and fair elections" (2021).

[27] Refers to the phenomenon where a vast amount of information, often contradictory and disjointed, prevents a clear and sensible understanding of candidates, issues, and policies. It's not just that there's too much information or information overload, but that the information itself lacks cohesion and clarity, making it difficult for voters to form a coherent picture.

[28] Donghee Shin and Emily Y. Shin, "Cascading falsehoods: mapping the diffusion of misinformation in algorithmic environments" *AI and Society* (2025).

[29] Chunpeng Zhai, Santoso Wibowo, *et.al.,* "The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review" 11 *Smart Learning Environments* (2024).

The virtual space frequented by voters as routine, where their interaction with content smeared by misinformation, disinformation, fake, and all other sorts of information far from the truth, erodes their capacity for reflective judgment, by collapsing the distinction between information and noise. This epistemically destabilization of voters made them incapable of synthesizing a coherent narrative from the barrage of disjointed inputs. Therefore, a movement which was started to inform voters to participate meaningfully in a deliberative democracy resulted in a voter incapacitated in cognitive comprehension, particularly when it comes to information coherence.

Amidst this, to ensure free and fair elections, ECI in 2013 extended existing legal provisions related to election campaigning to apply to social media in the same manner it applies to other forms of electoral campaigning using other media platforms. Further, in January 2022, the ECI added a new column in candidates' election expenditure returns to report the money spent on digital campaigning, which is also applicable to political parties. This expenditure shall include all expenditure on campaigning, including expenditure on advertisements on social media, payments to internet companies, expenditure on content development, etc. Lately, taking cognizance of the directions of the High Court of Delhi in *Lawyers Voice* v. *Union of India Through Ministry of Electronics and Information Technology*,[30] ECI in May 2024 issued an advisory regarding "Responsible and ethical use of social media platforms and strict avoidance of any wrongful use by political parties and their representatives during MCC period in General Elections and byelections".[31] The advisory was merely directive (or *persuasive*) in nature, with no penal provision or penalty attached for cases of non-compliance. Taking a stride further, ECI in January 2025, issued an advisory for "labelling synthetic/ AI-generated content used by Political Parties for election campaigning"[32] – again without any sort of penal or penalty provisions attached in case of non-compliance.

Despite these concerted efforts by ECI to ensure quality information is disseminated to voters, the grey areas persisted. Political spending on digital platforms is not only

[30] W.P.(C) - 6186 / 2024.

[31] Election Commission of India, "Responsible and ethical use of social media platforms and strict avoidance of any wrongful use by political parties and their representatives during MCC period in General Elections and by elections-regd." *available at*: https://www.eci.gov.in/eci-backend/public/api/download?url=LMAhAK6sOPBp%2FNFF0iRfXbEB1EVSLT41NNLRjYNJJP1KivrUxbfqkDatmHy12e%2FztfbUTpXSxLP8g7dpVrk7%2FeVrNt%2BDLH%2BfDYj3Vx2GKWdqTwl8TJ87gdJ3xZOaDBMndOFtn933icz0MOeiesxvsQ%3D%3D (last visited on Aug. 12, 2025).

[32] Election Commission of India, "Advisory for labelling synthetic/AI generated content used by Political Parties for election campaigning- reg." *available at*: https://www.eci.gov.in/eci-backend/public/api/download?url=LMAhAK6sOPBp%2FNFF0iRfXbEB1EVSLT41NNLRjYNJJP1K ivrUxbfqkDatmHy12e%2FzGjJMI0%2FjETs7fjrM8lYn4ipTqYtDEvVosG8Bae5QB8%2Fj5TBF9Esc2hlzORgYtkmzyKzGsKzKlbBW8rJeM%2FfYFA%3D%3D (last visited on Aug. 12, 2025).

done directly by political parties or candidates (intended targets of 2024 and 2025 advisories) but also indirectly by related affiliates or sympathetic groups, which may or may not necessarily be linked to the contestants. The ECI's guidelines do not adequately regulate third-party expenditure on online political advertising. Even the Model Code of Conduct and the Voluntary Code of Ethics do not adequately cover this area.[33] E.g., an entity named 'Ulta Chashmaa' whose Facebook page describes its goal as "shaping political discourse with a twist" spent 3.2 crore and also funded ads for other pages like MemeXpress, Political X-Ray, Tamilkam, and Malabar Central.[34] Candidate spending is capped, and this includes money spent by third parties. If a political party spends in favour of a candidate, it is treated as a third party spent. However, this has not been successfully extended to social media platforms. The term 'online political advertisement' is also not defined by either the ECI or in the Representation of Peoples Act, 1950 and 1951.

Money thus spent on social media platforms is difficult to monitor and is not reported.[35] Social media algorithms and the use of anonymous third parties and proxies to spread political messages make it hard to track spending.[36] There are also definitional challenges when it comes to identifying and regulating third-party accounts. In sum, virtual space is a great nuisance when it comes to free and fair elections in terms of information access by voters, but there is another factor which is muddying the water further, *i.e.,* hyper nudging and micro targeting of voters by algorithms developed and deployed by social media platforms, solely for commercial purposes.

III Algorithms, information, and manipulation- the *un*making of choices

The paradox is here- the expansion of technology in digital space has kept the promise of democratization of information, but algorithmic structuring of information has rendered voter choices opaque, curated, and manipulable. Algorithms, which are essentially designed to perform a task, operate on the set of instructions embedded in their code, have emerged as epistemic agents of digital platforms that selectively filter, amplify, and suppress information flows in accordance with economic, political, and psychological logics embedded in platform capitalism. These algorithms, designed to promote engagement, retention, and monetization of user attention, have profoundly reshaped the architecture of public reason.[37] Steadily, the cherished 'choice', which

[33] Association for Democratic Reforms, *available at*: https://adrindia.org/content/elections-digital-spends-on-advertisements-lack-transparency? (last visited on Aug. 12, 2025).

[34] *Ibid.*

[35] *Ibid.*

[36] Often these parties switch allegiance depending on nature of elections, region in which election is being held, political coalitions, and monetary benefits.

[37] Jocelyn Maclure, "AI, Explainability and Public Reason: The Argument from the Limitations of the Human Mind" 31 *Minds and Machines* 427 (2021).

makes us human[38] is becoming merely a residue of behavioural nudges, computational predictions, and psychometric profiling, leaving our decision- making capacity redirected, if not wholly unmade.

This algorithmic mediation has created an environment where our experience with information is in itself engineered. This collapse of the core pillars of traditional epistemological scaffolding, *i.e.,* authority, credibility, and deliberation, has led to a new regime of epistemic fragmentation. In a democracy, at the centre of this fragmentation stands the 'voter'- who has now ensconced in an astroturf world, with a curated worldview that simply mirrors his emotional predispositions, eroding the pluralistic engagement. The epistemically robust political content on social media platforms, to which a user (*voter*) subscribes for the sake of staying informationally relevant and updated, has paved the way for algorithmically tailored emotionally intensified content, privileging outrage, sensationalism, and tribal identity over reason, moderation, and complexity. Cumulatively, information and environment, rather than building informed consent, simulates a participatory illusion, with the end result of democracy without epistemic substance.

The algorithms feed on data and the more granular the data, the better the targeting. The commodification of user data has provided this opportunity to algorithms and opened the pandora box. With user base of over a billion and recognizing them in thousands of parameters, social media platforms have granular access to our likes, dislikes, and all that lies in between the two.[39] The commodification of personal data to this extent has facilitated microtargeting and hypernudging, that personalize not only advertisements, but political ideologies. On these platforms, adult users forming

[38] From the perspective of choice architecture, an individual's ability to make a choice is what defines their humanity, but this ability is not exercised in a vacuum. Choice architecture is the deliberate design of the environment in which people make decisions. This design influences, or "nudges," people toward certain choices without taking away their freedom to choose. Therefore, our humanity lies in our capacity to choose, but the exercise of that capacity is profoundly shaped by how the options are presented to us.

[39] What all lies in between- I am neither theist or atheist, but prays to God on occasions, or neither strict vegetarian or non-vegetarian, switches between two on whims, or likes policies/ approach of Party A on issue X and policy/ approach of Party B on issue Y. On these issues, an individual may face difficulty in understanding how and why my choices are shaped in that particular manner, but platforms who have details like, which type of content I like, dislike, comment on (and what was the tone of that comment), shared, reported, spent less or more time, type of content I prefer or search etc. can create a complex yet decipherable pattern. Which helps them in decoding an individual in a manner that even individual may not understand, for the reason that these patterns are taking activities which an individual perform in haste or without application of rationality or choices deep buried in consciousness in consideration while building the patterns. So, apart from white and black, platforms through algorithms know a lot of greys about an individual- which essentially provides granular data about that individual.

a voter base are no longer addressed as rational entities, but as behavioural subjects segmented by their vulnerabilities, preferences, and predicated responses.

Following are some illustrative case examples for reference:

(i) A pivotal moment in understanding the manipulative potential of these algorithms was the Facebook Emotional Contagion Experiment (2012), a study that demonstrated the non-conscious transfer of emotional states between users. This experiment, while controversial, laid bare the inherent power of platforms to influence psychological states. The experiment aimed to investigate *whether emotional states could be transmitted via social networks*, specifically through the manipulation of content displayed in users' News Feeds. The study concluded that *"emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness"*. While the Emotional Contagion Experiment utilized relatively simple algorithmic filtering rules, it established the principle of algorithmic influence over user psychology.[40] The intervening years have seen an explosion in the sophistication and pervasive application of Artificial Intelligence (AI) within social media, including Facebook (now Meta).

(ii) The 2016 US Presidential Election was a watershed moment, not least for the unprecedented role of digital data and social media in political campaigning. At the heart of a major controversy that erupted two years later was Cambridge Analytica (*hereinafter* CA), a firm that purported to use "data science methodologies" and "psychological profiling" to sway voter behaviour. CA's claimed innovation lay in its application of psychographic profiling, moving beyond traditional demographic segmentation (age, gender, income) to categorize voters based on psychological traits. Using the "Big Five" personality model (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism), CA utilized the collected Facebook "likes" and other digital behaviours to infer individual personality profiles. The scandal brought into sharp focus the opaque practices of data acquisition, the potential for sophisticated computational techniques to manipulate public opinion, and the urgent need for robust ethical and regulatory frameworks in the digital age.[41]

[40] L. Floridi, and M. Chiriatti, "GPT-3: Its nature, scope, limits, and consequences" 30 *Minds and Machines* 681-694 (2020).

[41] Campaign Legal Center, "*Newly Published Cambridge Analytica Documents Show Unlawful Support for Trump in 2016*", *available at*: https://campaignlegal.org/update/newly-published-cambridge-analytica-documents-show-unlawful-support-trump-2016, (last visited on April 27, 2025), *see also*, Down To Earth, "*Psychographics: the behavioural analysis that helped Cambridge Analytica know voters' minds*" *available at*: https://www.downtoearth.org.in/science-technology/psychographics-the-behavioural-analysis-that-helped-cambridge-analytica-know-voters-minds-59999 (last visited on Aug. 14, 2025).

(iii) At home, in the 2019 Indian general election, organizers used a network of WhatsApp groups to manipulate Twitter trends through coordinated mass postings. With over 600 WhatsApp groups linked to the political parties' IT cell that systematically instructed members to post specific content on Twitter, thereby generating artificial trends. These trends gave the illusion of organic public opinion, manipulating Twitter's trending algorithm to amplify political narratives, showcasing that the method of 'Human Astroturfing' is both effective and scalable, making it harder for automated systems to detect manipulation.[42] The social media platforms' overreliance on engagement metrics and hashtag velocity renders them vulnerable to coordinated behavior-particularly when such behaviour is engineered by members of the ruling dispensation. The case presents a textbook example of swaying voters under the guise of grassroots support.

(iv) At regional level, the 2021 West Bengal elections, witnessed digital media teams using AI-enabled sentiment analysis to monitor online discourse across Bengali-language content on Facebook, YouTube, and Twitter. Then gathered behavioral insights guided daily adjustments in tone, imagery, and issue focus.[43] Later in the 2022 Uttar Pradesh elections, political parties focused on behavioral segmentation grounded in caste, regional affiliation, and economic class. Here WhatsApp served as the primary delivery mechanism for personalized scripts, short videos, and calls to action. Messages were tailored based on behavioral attributes, including past voting history, issue sensitivity, and local leadership credibility.[44] Similarly, the 2023 Telangana elections saw parties heavily invested in automated WhatsApp bots and Customer Relationship Management (CRM)-integrated outreach platforms to deploy behaviorally segmented messages at scale. Campaigns tested micro-narratives, including irrigation schemes, power subsidies, and regional pride, based on real-time sentiment data gathered from social media interactions and voter feedback loops.[45]

[42] Maurice Jakesch, Kiran Garimella, *et.al.,* "Trend Alert: A Cross-Platform Organization Manipulated Twitter Trends in the Indian General Election", *available at*: https://arxiv.org/pdf/2104.13259 (last visited on Aug. 14, 2025).

[43] Smruthi Nadig, "As Elections Approach, AI Reshapes Electoral Analysis in India", *available at*: https://analyticsindiamag.com/ai-features/as-elections-approach-ai-reshapes-electoral-analysis-in-india/(last visited on Aug. 17, 2025).

[44] Paul Petritsch, "AI in Politics – The Example of the 2024 Indian Elections", *available at*: https://dentroai.com/ai-in-politics/ (last visited on Aug. 17, 2025); see also, D Dhanuraj, Sreelakshmi Harilal, et.al., "Generative AI and its Influence on India's 2024 Elections- Prospects and Challenges in the Democratic Process", *available at*: https://www.freiheit.org/sites/default/files/2025-01/a4_policy-paper_ai-on-indias-2024-electons_en-4.pdf (last visited on Aug. 17, 2025).

[45] Rahul Batra, "Elections, Accountability, and Democracy in the Time of A.I.", *available at*: https://www.orfonline.org/research/elections-accountability-and-democracy-in-the-time-of-a-i (last visited on Aug. 17, 2025).

These personalized and targeted campaigns, using algorithmic communications prioritize emotional exploitation and strategic disinformation to distort voter perceptions and reduce informed-decision making. The absence of an operational and comprehensive data protection framework tailored to dictate the use of voters' data for political purposes complexes the issue of electoral gerrymandering in virtual spaces. Though the Data Protection Act,[46] promises to regulate the handling of personal data, its implementation is still a far-fetched idea. At ground level, ECI lacks technical resources and legal authority to audit political advertising algorithms, verify compliance with data use, or mandate transparency disclosures. These gaps risk eroding trust in democratic institutions and processes- a phenomenon currently witnessed through lower voter turnout in each passing election.

IV The legal vacuum between virtual and physical

The advent of generative AI tools in the domain of politics by parties with vested interests and otherwise has introduced the problem of reality distortion. A deepfake image or video of a political candidate saying or doing something that they never did is not merely misinformation, but a malicious fabrication that poisons the information ecosystem and erodes trust in all forms of media. This opens up a new frontier of what Michel Foucault termed *biopolitics*, *i.e.*, the administration and regulation of human life by a sovereign power, but in altered realities of digital iteration, the power is algorithmic, and the object of regulation is the cognitive and affective life of the voter.[47] Thus, the collective harm is not just that a voter might believe a lie, but that they may cease to believe that truth is ascertainable at all. It is this profound threat to voter autonomy and epistemic integrity that any meaningful regulation must address.

The key legislation for the conduct of elections in India, *i.e.,* The Representation of the People's Act, 1951 (*hereinafter*, The Act), is largely silent on these concerns. Under chapter III of The Act, Sec 8 and 8A laid down the offences, in which if a member of parliament or state legislature gets convicted, then such member will stand disqualified from the membership of said house for the term of conviction and further for a period of 6 years since his/her release. Likewise, section 9, 9A, 10, and 10A call for disqualification on the grounds of corruption or disloyalty, Government contracts, office under a government company, and failure to lodge an account of election expenses. Further, section 123 defines 'Corrupt Practices' for the Act, ranging from bribery to gratification, undue influence to booth capturing, and more, section 125 to 136 detailing other electoral offences. Amidst these detailed provisions covering all facets of possible crimes associated with elections, not even a single provision

[46] Digital Personal Data Protection Act, 2023.

[47] Hazel Marie M. Vitales, "Foucault and Beyond: From Sovereignty Power to Contemporary Biopolitics", 9 *Mabini Review* 161-178 (2020).

deals with the algorithmic manipulation of voters by micro targeting, behavioural manipulation, and hyper nudging to sway the election.

However, to address the issue, ECI in 2013 extended existing legal provisions related to election campaigning to apply to social media platforms in the same manner it applies to other forms of electoral campaigning using other media platforms. Despite this, the problem remained unaddressed for the simple reason that issues and challenges arising in physical campaigning and virtual campaigning are poles apart and therefore, remedial measures efficient in physical space remain ineffective in virtual space.

Take the case of hate speech, at a physical rally, it may evoke an FIR under section 153A of IPC or Sec. 196 of BNS, but in virtual space, hate-laden political messages are amplified not by explicit intent but by engagement-driven algorithms favoring polarizing content. The candidate claims no direct role, arguing that the algorithm autonomously amplified content based on user engagement metrics. Similarly, bribery of voters may result into candidate's disqualification under section 123 of Act 1951, but in case the same candidate hires a data analytics firm to exploit voter data from social media, to send personalized narratives subtly reinforcing voter-specific biases, then though the act amounts to manipulation of electoral choice, it does not squarely fall under the bribery provision. In essence, mere extrapolation of provisions to virtual space won't do much to curb the associated challenges. Also, the 2013 guidelines are pre-AI phase, so their operation and efficacy remain dormant in the age of AI.

The recent advisories on use of AI in elections by ECI, *i.e.,* "Responsible and ethical use of social media platforms and strict avoidance of any wrongful use by political parties and their representatives during MCC period in General Elections and bye lections"[48] and "labelling synthetic/AI generated content used by Political Parties for election campaigning"[49] are indeed important but largely administrative signals; they need statutory teeth, technical specificity, and integration with platform-side obligations. Also, the targeted audience of both the guidelines is Political Parties and their leaders, candidates, and star campaigners, with no mention of ordinary citizens/voters, private Indian and foreign entities, and social media platforms. While former groups are well within the ambit of established statutes, rules, and regulations, it is the latter stakeholders who are coming at the forefront in virtual space campaigning for the simple reason that these stakeholders carry zero or minimal liabilities.

V Algorithmic manipulation: An international outlook

With Indian legal framework on algorithmic election manipulation succumbing to inefficacy, let's look at how other nations are faring with the issue.

[48] *Supra* note 32.

[49] *Supra* note 33.

European Union

The European Union is charting the course on AI by laying out world's first comprehensive and binding AI Act. Its regulatory philosophy is distinctly European, *i.e.,* precautionary, rights-based, and built on a tiered-risk approach. The EU AI Act[50], categorizes AI systems based on their potential for harm and therefore imposes stricter obligations on higher-risk systems.[51] For the purpose of electoral integrity, three articles are of paramount importance, *i.e.*, article 5, article 6, and article 50 dealing with prohibition on certain AI practices, Classification Rules for High-Risk AI Systems, and transparency obligations, respectively.

Art. 5(1)(a) of the AI Act represents the Act's most powerful intervention, which categorises 'Prohibited AI Practices' as:[52]

> *"… an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm."*

The article is a direct attempt to regulate the algorithmic *means* of persuasion. Going by the definition, a political campaign using an AI tool to micro-target voters with emotionally manipulative content designed to suppress their vote or incite outrage could fall under this prohibition imposed by the article. Despite the overarching definition, the challenge lies with the high threshold of proving "significant harm" and the intent to "materially distort behaviour".[53] Whether tweaking or attempt to tweak an individual's cognitive faculties and altering with choice architecture falls under significant harm, will be seen in the due course of time, as courts, legislatures, and civic society ponder on these harms through case laws and other discourses. Further, article 5(1)(b) of the Act bans AI systems that "exploit any of the vulnerabilities of a person or a specific group of persons due to their age, disability or a specific social or economic situation…". This provision directly targets the predatory nature

[50] The EU Artificial Intelligence Act, 2024

[51] The EU AI Act classifies AI systems into four risk categories: unacceptable risk (banned), high risk (strict regulations), limited risk (transparency required), and minimal or no risk (unregulated); see also, High-level summary of the AI Act, *available at*: https://artificialintelligenceact.eu/high-level-summary/ (last visited on Aug. 19, 2025).

[52] The EU AI Act, 2024, art. 5(1)(a).

[53] The definition clause of EU AI Act, 2024, i.e., Art. 3, doesn't provide explanation to significant harm and materially distort behaviour.

of AI-powered micro-targeting and hyper-nudging of demographics which are more susceptible to manipulation and seeks to protect them.

The article 6 of the Act, defines a high-risk AI system and Annex III of the Act with reference to Art. 6(2), among others states that:[54]

> *"… AI systems intended to be used for influencing the outcome of an election or referendum or the voting behaviour of natural persons in the exercise of their vote in elections or referenda. This does not include AI systems to the output of which natural persons are not directly exposed, such as tools used to organise, optimise or structure political campaigns from an administrative or logistical point of view."*

A plain reading of the provision and classification makes it abundantly clear that AI tools used to alter the electoral outcome or voting behaviour of natural persons fall under high-risk systems and are therefore used for specific psephological and other purposes only (mentioned in the later part of the above provision), with laid out risk management system. Providers and deployers of such systems must document risks, test for foreseeable misuse (including the creation or distribution of synthetic political content), and implement meaningful human oversight. In the context of India, being devoid of the AI Act, there are no such classifications of AI tools, which specifically deal with AI harms to voters caused by providers of AI tools and users of these tools.[55]

Lastly, comes the 'Transparency obligations for providers and deployers of certain AI systems' under Art. 50 of the Act, which states that:[56]

> *"…Providers of AI systems, including general-purpose AI systems, generating synthetic audio, image, video or text content, shall ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated…"*

The EU's approach here is one of mandatory disclosure. It does not ban deepfakes outright (except for specific illegal uses like non-consensual intimate imagery) but insists on labelling. For elections, this means that synthetic campaign videos, AI-generated voice-clones of political leaders, and photorealistic composites must carry conspicuous provenance disclosures. Providers and deployers must implement technical means to facilitate such labelling and, where feasible, state-of-the-art methods for

[54] The EU AI Act, 2024, art. 6(2).

[55] Unlike the European Union's comprehensive AI Act, India does not have a single, dedicated law specifically for the regulation of AI. This creates a unique legal landscape where there are no formal classifications of AI tools based on risk, especially concerning potential harms to voters.

[56] The EU Artificial Intelligence Act, 2024, art. 50.

traceability (*e.g.,* watermarking, content credentials). The article is an attempt to mitigate epistemic chaos by providing citizens with the necessary information to critically evaluate the content they consume.

In the EU's AI Act, the liability lies on the AI tool provider and not on the users *per se*, but in the context of India, this is reversed. The January 2025 advisory for labelling synthetic/AI-generated content used by Political Parties for election campaigning[57], put the onus on political parties, candidates, and star campaigners. While the measures taken by ECI here have merit, as it tackles the issue to an extent, it would be more beneficial if an overarching mandatory notice for labelling AI-generated content to providers of AI tools were sent. This will not only have taken care of political parties, candidates, and star campaigners, but also of all the third parties involved directly or indirectly with election campaigning. Certainly, enforcing such a measure will fall outside the executive power domain of ECI, but the same can be in place for a period when the model code of conduct is in force.

Philosophically, the EU's AI framework can be read as an autonomy-centred model. The prohibited practices clauses of "materially distorting behaviour" and "significant harm" translate the Mill's harm principle[58] into a digital register. The deepfake disclosure rule is likewise not a speech ban but an epistemic safeguard, as it treats disclosure as a condition for legitimate persuasion within a democratic public sphere. The overall structure attempts to reconcile pluralism in political persuasion with the baseline that voters must not be secretly programmed by inscrutable systems.

United States of America

Unlike EU's comprehensive legislation, the United States has adopted a more fragmented and market-oriented approach, driven by a combination of executive actions, non-binding policy frameworks, and targeted legislative proposals. It began with 2022 'Blueprint for an AI Bill of Rights' outlining five principles for the design and use of AI systems, *i.e.,* Safe and Effective Systems, Data Privacy, Algorithmic Discrimination Protection, Notice and Explanation, and Human Alternatives, Consideration and Fallback. The principle of 'Notice and Explanation' is of relevance here, asserting that an individual should know when and in what aspects an automated system is being used and understand how it works. In a broader political context, this would mean voters would have a right to know if they are being targeted by an AI-driven advertising campaign.[59]

[57] *Supra* note 33.

[58] Isabel Kusche, "Possible harms of artificial intelligence and the EU AI act: fundamental rights and risk", *Journal of Risk Research* 1-14 (2024) .

[59] Blueprint for an AI Bill of Rights, 2022, *available at*: https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/ (last visited on Aug. 22, 2025).

Again in 2023, President Biden's 'Executive Order on Safe, Secure, and Trustworthy AI'[60] set the path forward for federal policy but not a law. Its primary relevance to election integrity is its strong emphasis on transparency and detection. Section 4.1 of the order on 'Developing Guidelines, Standards, and Best Practices for AI Safety and Security'[61] stresses that the National Institute of Standards and Technology (NIST) to develop standards for detecting and watermarking AI-generated content. The goal is to create a reliable mechanism for distinguishing between authentic and synthetic content. The US philosophy is clear on AI-generated content, *i.e.,* rather than banning the creation of potentially manipulative content, the government aims to create the technical tools to identify it. The onus is then placed on the platforms to implement these tools and on the citizens to be discerning consumers of information. The focus here is on the disclosure model, not the prohibition model.

While at the federal level, the processes and progress are a bit lackadaisical, but states are at a fast pace when it comes to taming AI harms. As of mid-2025, more than two dozen states[62] regulate political deepfakes; in that, too, two archetypes dominate, i.e., time-bounded prohibitions on deceptive deepfakes proximate to elections (*e.g.,* Texas bans harmful political deepfake videos within 30 days of an election), and affirmative disclosure mandates for synthetic political ads (seen in Minnesota and many others). Remedies against AI induced harms typically include injunctions, takedown orders, and civil liability, with criminal penalties in some jurisdictions. The state trend line shows an American preference for narrow, speech-sensitive rules focused on deception and timing, rather than categorical content bans.

Apart from these enacted and enforced legislations, various bills with bipartisan support have been introduced in Congress to be passed into a comprehensive federal law. The proposed Deepfakes Accountability Act[63] and Protect Elections from Deceptive AI Act[64] (PEDAI), prohibit materially deceptive AI-generated audio or visual media in political ads about federal candidates, with narrow defenses (*e.g.,* satire clearly labelled as such). Even if not yet enacted, PEDAI's design choices are quite instructive, as it targets deception rather than AI per se, leverages labelling as a safe harbour, and creates remedies that dovetail with existing ad-disclosure frameworks. If enacted and

[60] Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 2023, *available at*: https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence (last visited on Aug. 22, 2025).

[61] *Ibid.*

[62] Future of Privacy Forum, "U.S. State AI Legislation- How U.S. State Policymakers Are Approaching Artificial Intelligence Regulation" (2024); see also, Tatevik Davtyan, "The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained" 16 *Journal of Law, Technology, and The Internet* (2024).

[63] Deepfakes Accountability Act, 2023.

[64] Protect Elections from Deceptive AI Act, 2023.

adopted, it holds the potential to become the first federal statute squarely aimed at algorithmic manipulation in campaign communications.

Doctrinally, EU's model is systemic; it regulates upstream (market placement), midstream (risk management), and downstream (transparency and users-facing duties). Election manipulation in the EU's framework is absorbed into this general AI governance canvas, using function-based triggers (Annexe III tothe EU's AI Act, 2024) to switch on heavier obligations. The United States model, by contrast, is modular, *i.e.,* content authentication via executive action, channel-specific disclosure rules (broadcast), and state-level election speech statutes aimed at deception windows. Both converge on transparency as the minimally invasive tool- label synthetic political content, yet diverge on *ex ante* market controls and formal risk classification. At their core, both embody a different view of democratic autonomy. The European Union's prohibitions and risk tiers acknowledge that some techniques so thoroughly bypass reflective agency that they must be precluded before they reach voters. The U.S. treats manipulation as a species of fraud or deception that can be cabined by labels and timing while preserving a wide berth for political speech. The trade-offs are legible: the EU optimizes for systemic safety and predictability; the United States optimizes for speech resilience and incrementalism.

VI The Way for India

The ECI response to algorithmically manipulated elections through May 2024, and January 2025 advisory is more or less an administrative signal, devoid of statutory teeth, technical specificity, and platform-side obligation. Apart from these, India's larger digital law stack, i.e., The Information Technology Act, 2000, The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021 (updated in 2023), and Digital Personal Data Protection Act, 2023 too presents inefficacy when it comes to swaying of voters by means of algorithmically advanced micro-targeting and hyper-nudging propagandas.[65] Though the DPDP Act, have consent-centric data processing norms, which, if operationalized against political profiling and microtargeting, can limit invasive voter manipulation at scale.[66] But the lack of an election-focused framework let loose the synthetic political media and political recommender systems.

[65] Sayantan Chanda, Data Privacy and Elections in India: Microtargeting the Unseen Collective", 18 *Indian Journal of Law and Technology* (2022).

[66] Purushraj Patnaik, "AI at a Crossroads: Navigating Consent-Centric Data in India", *available at*: https://www.orfonline.org/expert-speak/ai-at-a-crossroads-navigating-consent-centric-data-in-india (last visited on Aug. 21, 2025); see also, Aditi Kanoongo and Harshitha Adari, "Unpacking India's Digital Personal Data Protection Act: A New Dawn or a False Start?" *available at*: https://ohrh.law.ox.ac.uk/unpacking-indias-digital-personal-data-protection-act-a-new-dawn-or-a-false-start/ (last visited on Aug. 21, 2025).

To tame the menace, India should consider a *narrow*, content-neutral prohibition on AI systems that materially distort behaviour via manipulative or subliminal techniques when deployed in electoral contexts.[67] The prohibition would not criminalize persuasion[68] or satire, rather, it would target the *mode* of operation- systems designed to bypass deliberation by exploiting cognitive vulnerabilities at scale (*e.g.,* adaptive synthetic avatars tuned to neuro-signals, or psychometric microtargeting that leverages sensitive attributes). This maps onto the EU's autonomy-protective rationale while remaining consistent with Indian free-speech doctrine, which permits reasonable restrictions "in the interests of…public order" and the integrity of elections[69]. Following Annexe III of EU's AI Act approach, India should classify an AI system as "high-risk", used to target, profile, or influence voters or to moderate the visibility of political content during the election period. Consequences of such a measure would include- risk management plans (as per article 9 of EU AI Act), pre-deployment testing (including red-team exercises and sandbox operations for disinformation misuse), audit logs, and documented human-oversight mechanisms within campaign stacks and platforms. The ECI could be given rule-making power (apart from the period when the model code of conduct is in operation), akin to a sector regulator, to define test protocols and to certify third-party auditors.

Building further on ECI's 2025 labelling advisory, the advisory should convert into binding rules that mandate clear, prominent on-artifact labels for AI-generated or manipulated audio-visual political content¸ requires AI content generation platforms to support content-provenance standards such as C2PA credentials[70] and robust watermarking, with anti-stripping measures, create a verifiable disclosure scheme accessible to fact-checkers and journalists, and provide robust and expedited notice-and-action pipelines during Model Code of Conduct periods (e.g., verified political entity fast-lanes with 2-hour removal targets for deceptive deepfakes). A distinct and

[67] NITI Aayog, "National Strategy for Artificial Intelligence- #AI for All" (2018).

[68] Persuasion is the ethical process of influencing beliefs through rational, transparent appeals that respect an individual's autonomy and ability to make an informed choice. Manipulation, conversely, is an unethical form of influence that uses deception or emotional exploitation to bypass rational judgment for personal gain.

[69] Authors believe that sooner or later art. 19(2) should have this among reasonable restrictions. As, it will bolster the protection of democracy- one of the ideal and basic structure of Indian constitution.

[70] C2PA (Coalition for Content Provenance and Authenticity) credentials are a set of tamper-evident metadata that establishes the origin and history of digital content. They act as a "nutrition label" for media, cryptographically signing information about creation, edits, and authorship to provide transparency and build trust; see also, National Cyber Security Centre, "Content Credentials: Strengthening Multimedia Integrity in the Generative AI Era" (2025), *available at*: https://media.defense.gov/2025/Jan/29/2003634788/-1/-1/0/CSI-CONTENT-CREDENTIALS.PDF (last visited on Aug. 21, 2025).

evolving Indian challenge is the deep percolation of recommender systems, short-video 'For You' feeds, and closed messaging virality.

Platform governance suggestions

To address this, ECI can mandate platforms to oblige with - (i) publication of down ranking and demotion criteria for detected or likely synthetic political media[71], (ii) enable provenance-aware friction[72] (*e.g.,* interstitials before re-shares of flagged items), (iii) expose ad-targeting taxonomies, look-alike audience features, and political ad libraries with real-time API access, and (iv) offer an MCC "civic integrity mode" that temporarily disables sensitive-attribute like microtargeting and limits optimization to reach/engagement criteria that have undergone harm testing. These measures can very well be grounded in IT Rules 2021 due-diligence duties, elaborated by ECI rulemaking and MeitY advisories.

As said, it is always better to cure the ailment at the root. At the root of the ailment at hand lies the poor data protection rules and guardrails. If, along with biometrics and other personal information, 'political preferences' are categorised as 'sensitive' information, then the collection, possession, processing, and targeting will fall under legal no-go zone, and the problem would not have arisen in the first place.[73] On this facet, the DPDP Act's forthcoming implementation rules hold remedy, as they should clarify that political opinions and inferences about political preferences fall under 'the sensitive information' segment during the election periods, which then trigger stricter processing conditions and meaningful transparency about automated decision-making that materially affects civic participation (*e.g.*, ad delivery that determines what

[71] Bekir Tolga Tutuncuoglu, "Algorithmic Memory: How AI is Rewriting What Societies Choose to Remember or Forget" *available at*: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5271673 (last visited on Aug. 21, 2025); see also, Jide Alaga, Jonas Schuett, *et.al.*, "A Grading Rubric for AI Safety Frameworks" *available at*: https://arxiv.org/html/2409.08751v1 (last visited on Aug. 21, 2025).

[72] Marilyn Zhang, "Strengthening Information Integrity with Provenance for AI-Generated Text Using 'Fuzzy Provenance' Solutions", *available at*: https://fas.org/publication/strengthening-information-integrity-provenance/ (last visited on Aug. 21, 2025); see also, Natalia Díaz-Rodríguez, Javier Del Ser, *et.al.*, "Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation" *available at*: https://www.sciencedirect.com/science/article/pii/S1566253523002129 (last visited on Aug. 21, 2025).

[73] The Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011 (SPDI Rules) provided a narrow list of data considered "sensitive." This list included passwords, financial information, health conditions, biometric information, and sexual orientation. Notably, political preferences were not explicitly on this list. This created a legal grey area, allowing political parties and other entities to collect and process data on citizens' political affiliations without the heightened protections required for other types of sensitive data. This lack of protection raised concerns about voter manipulation, profiling, and discrimination, as highlighted by numerous incidents.

information a voter sees).[74] Also, the consent interfaces for political ads should be explicit, revocable, and non-coercive, with dark patterns in campaign sign-ups and data collection should be treated as "unfair trade practices." The normative aim of these measures is not to ban all targeting, but to cabin the forms that undermine equal access to information and amplify epistemic asymmetries.

As a measure, taking cue from the Unied States, ECI, along with social media platforms should wipe out all the explicit and implicit campaign material in circulation, before 48 hours of voting. And oblige platforms for expedited takedowns and demolition of such content if it originated or circulated on their platform. These measures should encompass all sorts of content originated by politicians, political parties, voters, and other private entities, with no exception. However, platforms here need to be conscious enough to not to take down which calls for voters' literacy and voters' participation in the elections.[75]

Data protection suggestions

Finally, ECI should institutionalize a technical directorate with authority to (i) certify provenance technologies[76], (ii) maintain a secure clearinghouse for cryptographic signatures used by parties and authorized creators, (iii) publish election-period risk bulletins (analogous to CERT-In advisories) for deepfake typologies & other manipulative practices, and (iv) coordinate with CERT-In and MeitY on rapid responses. The 2024-25 advisories show intent; it's time to go for durable capacity and binding standards.

It's high time for India to augment and synthesise these democracy saving measures, as in liberal constitutionalism that India beholds, which protects speech, must also regulate the conditions of speech- provenance, explainability of amplification, and the accountability of systems that intermediate between speakers and listeners. Its constitutional vision of democracy, that is both robustly procedural and substantively

[74] Shamita Islur, "How India's DPDP Act could change digital campaigns", *available at*: https://www.socialsamosa.com/experts-speak/how-india-dpdp-act-could-change-digital-campaigns-8606658 (last visited on Aug. 23, 2025).

[75] Vaishali U. Gongane, Mousami V. Munot, *et.al.*, "Detection and moderation of detrimental content on social media platforms: current status and future directions" 12 *Social Network Analysis and Mining* (2022); see also, Katja Müller-Helle, "Napalm Girl as Data Set. On the Reappraisal of Iconic Images by Automated Filtering Technologies", *available at*: https://journals.openedition.org/transbordeur/2718 (last visited on Aug. 23, 2025).

[76] Provenance in artificial intelligence refers to the complete, verifiable, and transparent history of an AI system. It's a comprehensive record that tracks every component and decision, from the origin of the data to the final output of the model. The primary goal is to provide a clear audit trail that answers the fundamental question: "Why did the AI system do that?" ; see also, Shayne Longpre, Robert Mahari, *et.al.*, "Data Authenticity, Consent, and Provenance for AI are all broken: what will it take to fix them?", *available at*: https://arxiv.org/html/2404.12691v2 (last visited on Aug. 23, 2025).

egalitarian, has taken shape as a platformized polity that requires both *ex ante* design constraints and *ex post* remedies. The immediate task therefore, is modest yet foundational- make the synthetic content visible, make the targeting accountable, and ensure that the most powerful political speakers are answerable to the public during the hours when democracy speaks.

VII Conclusion

Undoubtedly, the transition of political campaigning from physical spaces to virtual space in a platformized polity made individuals and private entities campaigners in one form or another. Where earlier political campaigning and canvassing were limited to political parties, candidates, and supporters, now with the opening of virtual spaces for one and all at minimal cost, these spaces have become virtually a 'fish market'- with noise, chaos, and disorganised content taking much of the space. To aggravate the issue came the 'Generative AI'-a potent tool for nuisance, which at scale blurred the line between information and mis-, dis-, and fake information. As an effect arises trust deficit in the information economy, institutional autonomy, and in the election process itself. It is imperative, then, to re-theorize democracy in light of these developments.

A critical philosophy of information must interrogate not only the content of political communication but the infrastructures that mediate its circulation. The question must be raised over who designs the architectures of attention? Whose interests do these architectures serve? And what modes of resistance are available to reclaim epistemic agency? While digital technologies hold the potential to expand participatory democracy, at the same time ,their current configuration deepens epistemic inequality and cognitive manipulation. If democracy is to survive the algorithmic age, it must evolve beyond procedural formalism to embrace an ethics of information- a normative framework that foregrounds transparency, accountability, and cognitive justice. Such an ethics would then require a radical reconfiguration of platform governance, where algorithmic transparency is not a concession but a democratic right. It would demand educational initiatives that cultivate critical media literacy, enabling citizens to navigate and interrogate the information environments they inhabit. Most importantly, it would necessitate a revalorization of the political subject, not as a consumer of content, but as a co-creator of the democratic public sphere. Only through such epistemic vigilance can we resist the unmaking of choices and reclaim the possibility of informed and autonomous political action. In doing so, we not only protect the integrity of elections but reassert the philosophical promise of democracy itself: that governance belongs to those who think, decide, and act together, in full knowledge and shared understanding.