



Artificial Intelligence and the Emotional Manipulation Economy

Saumya Dwivedi^{1,*}

¹ Research Scholar, Amity Law School, Amity University, Noida, Uttar Pradesh, India

* Corresponding author: dwivedisaumya1318@gmail.com

ARTICLE INFO

Article history:

Received: 23-07-2026

Received in revised form:
03-08-2026

Accepted: 04-08-2026

Available online: 05-08-2026

Keywords:

Affective Computing

AI Persuasion

Emotional Manipulation

India Digital Regulation

Platform Governance

ABSTRACT

Artificial intelligence has turned digital communication systems into behavioural ecosystems that predict, influence and commercialise human emotion. This paper examines what it terms the emotional manipulation economy: the commercial arrangement in which algorithmic recommendation, behavioural profiling, emotional analytics and persuasive design act upon autonomy and decision-making in digitally mediated environments. It argues that emotional response has become an extracted resource within platform capitalism, monetised through engagement maximisation, advertising targeting and behavioural prediction. The analysis traces the relationships between artificial intelligence, surveillance capitalism, emotional dependence and algorithmic amplification, and then examines the ethical and legal problems that emotionally manipulative technologies raise for mental privacy, informed consent, democratic participation, digital addiction and the commodification of emotional life. The method is doctrinal and analytical, drawing on law, behavioural psychology, political economy, media studies and technology ethics, with sustained attention to Indian regulatory practice. The paper finds that consent-based frameworks are structurally insufficient because they cannot reach the invisible behavioural manipulation embedded in algorithmic processes, and that accountability, transparency and enforceable mental privacy protections are required instead. It concludes that protecting emotional autonomy is a condition of democratic freedom in algorithmic societies.

1. INTRODUCTION

Artificial intelligence has significantly changed the dynamics between technology, business, and human behaviour. No longer do digital platforms act as passive sites of communication that facilitate mere exchange of information. On the contrary, they act as behavioural prediction systems that analyse people's emotions, affect decision-making processes, and manipulate human interactions. The recommendation engine, predictive algorithms, targeted ads, emotional analysis, and behavioural profiling have become the bedrock of current digital economy. Emotions have thus become economic commodities in this new economic model.

Attention is extremely important in the business models of the major tech firms since it produces behavioural data, advertising dollars, and predictive certainty. Artificial intelligence has enabled the companies to track behavioural signals such as scrolling speed, pause duration, reaction patterns, searching history, typing patterns, facial expression, and online interaction in order to predict emotional reactions. Emotional responses like anger, fear, excitement, insecurity, loneliness, and validation produce increased levels of engagement and thus increase profits. Therefore, digital environments are designed to provoke emotional behaviour.

The advent of emotionally manipulative digital ecosystems has been instrumental in the creation of what may be referred to as the “emotional manipulation economy.” In this economy, emotional experiences are transformed into quantifiable behavioural data that can be studied, commoditised, and monetised using algorithmic models. The transition highlights an overall shift from attention capitalism to emotional capitalism, whereby human emotions are treated as raw materials for predictive commercial systems [1].

The aims of this paper are as follows: Firstly, to explore the ways through which artificial intelligence systems are employed in emotional manipulation by means of behavioural profiling, recommendation algorithms, emotional analytics, and persuasive design architectures. Secondly, to investigate the ethical, social, psychological, and legal dimensions of emotionally tailored digital systems, especially regarding their effect on autonomy, mental privacy, democratic engagement, and consumer protection. Thirdly, to assess existing governance regimes and develop accountability models that strike a balance between technological progress and human dignity.

Governing algorithmic systems that shape behaviour, influence decisions and mediate the future of professional and democratic work therefore raises questions of accountability, fairness and institutional design related work by Pandey and collaborators provides complementary context [2, 3].

2. MATERIALS AND METHODS

2.1 Study design

The research method used in this paper is doctrinal and analytical in character. The paper draws mostly on secondary sources such as academic works, legal literature, policy documents, judicial opinions, cross-disciplinary research, and regulatory regimes concerning artificial intelligence and digital governance. The study is based on insights from behavioural psychology, media studies, political economy, and technology ethics to achieve a multi-dimensional analysis of the emotionally manipulative digital systems.

3. RESULTS AND DISCUSSION

3.1 Decoding the emotion-driven influence economy

The emotion manipulation economy is the concept of digital-commercial system, where emotions are analysed, commoditized and used for economic benefits. In contrast to the classical capitalistic system which relied on material production and consumption for economic gains, modern digital platform economy is based on behavioural prediction and manipulation which can bring considerable economic benefits.

Human emotions today are becoming an important source of data and can be used for economic gains. As Zuboff describes the phenomenon as ‘surveillance capitalism’, where human experiences are used as raw materials for behavioural prediction markets [1]. Large tech companies collect massive amounts of behavioural data during digital transactions, which can be transformed into predictive tools of behaviour manipulation. Emotional reactions are especially important because when stimulated emotionally, a person becomes more predictable and controllable.

As Srnicek states, platform capitalism is inherently based on the continual extraction of data as an economic resource [4]. Social media is a prime example of how economic structures manipulate emotions. Some algorithms constantly track user interactions to discover what kind of content elicits stronger emotional responses. Outrageous, fearful, anxious, funny or validating material gets more exposure simply because it is able to sustain engagement better than neutral content does. This is how there arises an economic motive for platforms to amplify emotionally provocative material irrespective of its wider social impact.

Emotional commodification constitutes a substantial development of digital capitalism [5]. Until recently, advertisements were designed mainly to appeal to the desires of consumers through demographic targeting and mass media communications. The use of artificial intelligence makes behavioural targeting much more precise. Vulnerable periods, emotional weaknesses and psychological traits can be identified by means of constant behavioural tracking. As a result, personalized commercial messages are based on emotional vulnerability.

Predictive analysis is another important aspect of the emotional manipulation economy. With the help of artificial intelligence systems, behaviour patterns can be analysed to predict people’s decision making and their emotions. Such predictability becomes economically significant since companies are able to

target consumers when they are most emotionally responsive. A third distinguishing feature of the emotional manipulation economy is behavioural dependence. Digital platforms tend to create behavioural dependencies because extended interactions allow for greater data harvesting and more profitable advertising. Behavioural loops are created via persuasive design techniques such as rewards, autoplay functions, notifications, and validation loops that can be optimized through artificial intelligence based on behavioural analysis in real time.

It is important to note that emotional manipulation in digital systems is invisible rather than coercive. In contrast to explicit coercion, algorithmic manipulation occurs invisibly through environmental structuring and behavioural nudging. Even when recommendation systems guide user choice and emotionality, users may feel that their decisions are autonomous, which makes it difficult to define consent and autonomy.

Finally, the appearance of emotionally manipulative economies should be viewed against the backdrop of changes occurring within neoliberal capitalism. Increasingly, economic value is extracted not only from physical production but also from data harvesting, behavioural manipulation, and prediction [6]. Human attention and emotional responsiveness therefore become strategic commercial assets within platform-driven economies [7].

3.2 Algorithmic frameworks for emotional monitoring and influence

3.2.1 Emotion-sensing capabilities of intelligent systems

The ability for AI systems to understand emotions based on behaviour and physiological data has evolved quicker than many would have predicted. Since the introduction of the term affective computing by Picard in 1997, a significant amount of research has been conducted in this field, focusing on automatic analysis of emotional state from facial expression, voice, text, physiological measures and behaviour [8]. The spectrum of interests that slowly became an agenda for research has been increasingly commercialised into systems that go beyond the mere recognition of emotions to their response in a way that is useful for institutional goals and not for human welfare.

What most people don't realize is that the input modalities that today's AI emotion inference systems have access to are much more diverse. Video-based facial action coding systems trained on video data are able to classify facial expressions into a small set of emotional categories, and accuracy of these systems has been reported to exceed human accuracy on standardised facial expression benchmarks, although the validity of standardised benchmarks is debated [9]. The voice analysis systems are designed to identify the prosodic features, speaking rate, and spectral features of the voice to infer affective state from audio. Sentiment, emotional valence, and psychological state can be reliably detected from written text using NLP models, to the point of being commercially viable and scalable. Perhaps most importantly for platform applications, scrolling speed, pause length, click hesitation and replay behaviour are all behavioural signals that can give a continuous stream of data to a machine learning model to infer emotional engagement states without asking the user to self-report [10].

These capabilities are not necessarily used in a commercial application with absolute accuracy. Even though the system misclassifies 40 percent of users, it will outperform a uniform baseline much of the time when it correctly infers a higher emotional vulnerability in 60 percent of cases. The economic rationale for emotional targeting is not about precision, it's about the 'marginal advantage', but taken across an awful lot of interactions. If you look at the scale at which major platforms operate in India, with over 500 million users on their social platforms and over 500 million users on YouTube, then any marginal enhancements on emotional targeting equate to a difference in revenue which is far bigger than what would cost to develop the targeting systems [11].

Just as the inference side is important, so too is the output side of emotional targeting. After the emotional state has been inferred, the system chooses the content, timing, framing and presentation format to get the desired response. The desired outcome of an ad campaign is usually a purchase impulse or association with a brand. To engage with social media, it's ongoing time on platform. Attitude change, donation behaviour, or mobilisation of voters is the political communication target. The selection process is trained on historical data reflecting the outcome that maximized platform/ advertiser's revenue. Therefore, it increases the selection of patterns that led to the highest levels of historical engagement for the platform/ advertiser, without regard for whether those patterns are beneficial to users [12].

3.2.2 Behavioural profiling and algorithmic persuasion

Recommendation algorithms represent one of the most influential forms of algorithmic persuasion. Platforms such as YouTube, Instagram, TikTok, Netflix, and Amazon continuously analyse user behaviour

to determine which content is most likely to sustain engagement. These systems prioritise behavioural outcomes rather than informational neutrality. Emotionally stimulating content therefore receives algorithmic amplification because it generates stronger interaction rates [13, 14].

Emotional targeting is made much more effective if it is used alongside psychological profiling, which allows for the matching of content with more stable emotional traits of the personality that make it susceptible to specific persuasive approaches. The academic underpinning of this approach is the Five Factor model of personality, which describes individual differences between the openness, conscientiousness, extraversion, agreeableness, and neuroticism. Youyou, Kosinski and Stillwell [15-16] showed that machine learning models based on Facebook likes can predict personality dimensions as accurately as judgments by a participant's close friends, with just a few digital behavioural signals.

The results of this research were popularized in the Cambridge Analytica controversy, when the use of psychographic profiles generated from Facebook data for the purpose of political advertising was reportedly done on voters in the UK and U.S. [17]. Political actors in India had the same powers. The 2019 General Election in India was exceptional in terms of the sophistication of data-driven political communication efforts of various parties. Journalists and researchers reported that voter-level information was used and that this information was sent via messaging platforms, such as WhatsApp, effectively being the main means of political polarization and emotional information in both urban and rural areas [18].

The most relevant factor for the emotional manipulation economy is the fact that the dimensions of the personality can be used to predict not only content preferences, but even emotional vulnerabilities. People with high levels of neuroticism react to threat appeals. Low agreeableness persons are more sensitive to in-group solidarity content. Persons with high conscientiousness are more susceptible to appeals based on duty and responsibility. An AI system that has deduced such features could then choose content to elicit the emotions in which a particular user is most receptive, without any human agent making such a selection explicitly [19].

Table 1. AI emotional inference modalities and applications, including the Indian platform context

Input modality	Technical method	Inferred state	Application / Indian platform context
Facial expression	CNN on video	Discrete emotions	Retail and advertising testing; emerging in OTT platforms used by Hotstar and Amazon Prime India
Voice prosody	Spectral feature extraction	Stress, sentiment	Call centre analytics widely deployed in India's BPO sector; BFSI customer profiling
Written text	Transformer sentiment models	Emotional valence	Social media targeting across Hindi, Tamil, Telugu and Bengali content; WhatsApp message analysis
Behavioural signals	Clickstream modelling	Vulnerability, receptivity	YouTube Shorts and Instagram Reels retention optimisation; dominant mode for Indian users
Physiological data	Wearable sensor integration	Arousal, stress	Health and insurance apps growing in the Indian market; regulatory oversight absent

Source: Compiled by the author

Table 1 shows the diversity of emotional inference modalities in commercial use (which was recorded), and the rightmost column shows the extension of the table to account for the context of the Indian platform. There are certain vulnerabilities in the Indian digital environment. This is because short video platforms are the leading platforms for new Internet users, WhatsApp is used as an information channel by users of all ages, and because digital consumption in India is multilingual, emotional targeting systems exist in contexts where the users' media literacy and algorithmic awareness are highly variable [20].

3.3 Sectoral applications and the mechanics of harm

3.3.1 Digital advertising and consumer exploitation

Emotional targeting is a slow, progressive practice in the digital advertising space and most often unrevealed to the people who are being targeted. The underlying technology is contextual and behavioural advertising, whereby ads are chosen based on browsing history, purchase history and demographic inference. Emotional targeting is one level up from this: In addition to who the user is and what he has done, what emotional state did he appear to be in at this moment, and what content is likely to be most useful to the advertiser from a productive standpoint?

The real-world effect is an advertising context that is specifically targeting users at their emotional weak points. An individual who has been viewing content about financial strain, health worry, or family problems is algorithmically categorized as being in a heightened affective state, and may be exposed to ads that are based in that affective state: ads for financial products to people in a state of health anxiety, ads for food delivery to people in a state of loneliness, ads for aspirational content to people in a state of economic insecurity. This is coupled with a high level of digital illiteracy in India. The first-generation smartphone user in a tier-two city who gets directed to an algorithmically optimised feed is unlikely to be cognizant of the targeting architecture used to deliver the content to them, and lacks the tools and knowledge to assess the content being delivered to them [11].

The moral dilemma is not the idea that ads respond to emotions in general. Contextual ads have been doing something like this all along. The issue, however, is that with AI emotional targeting, they are more specific than general contextual placement, which doesn't leave the user to the subconscious level of evaluation. The difference between advertisements being placed in an emotionally relevant setting and ads specifically tailored to elicit an individually inferred psychological vulnerability is meaningful [21].

There is no specific regulatory framework for emotional targeting using AI in India, with the Advertising Standards Council of India (ASCI) and the Consumer Protection Act 2019 providing a general guideline. While the ASCI's 2023 guidelines for influencer advertising and AI-generated content are a measure of accountability, they do not address the targeting architecture that drives the platforms [22]. As a consequence, the Indian market is virtually unregulated for practices that would attract regulatory attention in Europe.

3.3.2 Social media platforms and engagement optimisation

Social media platforms are a unique but related use case for emotional AI. Platform engagement systems use emotional inference to engage users with the platform and cause users to spend more time on the platform, which creates the platform advertising inventory that is sold through advertising. The two systems are not independent; the engagement optimization system sets the emotional ground that advertising targeting plays in, and advertising revenue sets the incentives that the engagement system's goals are based [23].

There has been significant research into the link between algorithmic content curation and emotional harm. Bail and colleagues discovered that when participants are exposed to material that causes outrage, they become more polarised in their political attitudes when viewing this content through algorithms. This has been more pronounced in India in the circulation of emotionally triggering content around religious and caste lines. In the lead-up to the Shaheen Bagh protests in 2019-20 and the violence in Delhi in February 2020, researchers and journalists have documented the presence of weeks of algorithmically amplified inflammatory content on Facebook and WhatsApp, which preceded both events. In 2021, Facebook's internal research, revealed in the Wall Street Journal's Facebook Papers investigation, found that content encouraging group identity and outrage was effective at driving engagement and was a danger that the company had yet to properly mitigate [6-7].

Recommendation algorithms do not cause emotionally destabilising content to be recommended in an intentional way; it's a structural process. Strong emotional responses lead to these behaviours, and these behaviours will be amplified based on the system's optimisation for engagement metrics. Emotion evoking anger, outrage, fear and anxiety is reliably more engagement-generative than contentment or satisfaction [5]. An algorithm that could maximise engagement, without being specifically "optimised" for that, will therefore be automatically programmed to show content that elicits those conditions most likely to increase platform use, even when they're harmful to people. The multilingual nature of this amplification in India, which occurs in several languages including Hindi, Urdu, Tamil, Telugu, Marathi, as well as other languages, adds a significant challenge for monitoring and intervention compared to the monolingual setting.

3.3.3 Political communication and the manipulation of democratic feeling

Beyond consumer-level concerns, the use of emotional AI in political communication has sparked an understanding of the potential implications for the integrity of democratic deliberation itself. Emotionally-charged rhetoric has been a part of the political's arsenal since the dawn of time. The issue of personalised emotional targeting via artificial means is asymmetrical and opaque: it is targeted at different psychological profiles, it works at a scale that makes it impossible for human control, and the targeted person is not aware that their feelings are being manipulated with AI for political ends [24].

India offers one of the most important and under-studied examples of political emotional targeting on a large scale. Sophisticated digital political operations were seen during both the 2014 and 2019 general elections. The Bharatiya Janata Party's IT cell, as well as other affiliated organisations, created a network of WhatsApp groups in the hundreds of thousands that were leveraged to provide politically charged material at the constituency level, with an emphasis on emotion [13]. This content was deliberately optimised for emotional targets: pride, aspiration for some time, threat and insecurity for others, according to analysis of what was most effective in previous distribution cycles. Comparable, though less resourced operations, have developed on the same timeframe among the Congress party and regional parties.

These features have also been enhanced in the 2024 Indian general elections, with AI-generated audio and video content being employed by various parties to create personalised outreach materials in regional languages. Examples of documented cases included audio recordings of party leaders' speeches appearing in constituencies where the party leaders had not visited in person [25]. In 2024, the Electoral Commission of India issued guidelines for disclosure of AI-generated political content, but they were limited in their scope and coverage and do not cover the architecture of emotional targeting of content that is used for content selection [26].

Table 2. Emotional targeting in political communication: comparative cases

Election	Country	Primary channel	Emotional strategy	Documented source
2016 US presidential	United States	Facebook psychographic ads	Fear, identity threat, in-group solidarity	Cadwalladr & Graham-Harrison, 2018 [17]
2016 Brexit referendum	United Kingdom	Facebook micro-targeting	Economic anxiety, sovereignty fear	Cadwalladr & Graham-Harrison, 2018 [17]
2019 general election	India	WhatsApp group networks	Pride, threat to identity, religious fear	Banaji et al., 2019 [27]
2020 US presidential	United States	Facebook and YouTube A/B tested ads	Real-time emotional register optimisation	Diakopoulos, 2020 [28]
2024 general election	India	AI audio in WhatsApp and regional platforms	Personalised aspiration and threat messaging	Election Commission of India, 2024 [26]

Source: Compiled by the author

Table 2 compares the Indian political emotional targeting cases with those in the U.S. and the U.K. The comparison shows that the mechanisms are the same in both contexts, but the Indian cases introduce new layers of complexity: the distribution model used in WhatsApp is beyond the control of standard transparency tools on the platforms; the multilingual landscape makes it more difficult to keep track of the different content ecosystems; and the electoral integrity regulatory framework has not advanced to match the capabilities of AI being used.

3.4 Regulatory frameworks and the barriers to existing protection

3.4.1 Scope and limitations of existing regulatory frameworks

The regulatory response to emotional AI has been uneven, with different jurisdictions taking varying approaches and the number of instances in which these AI systems are deployed lagging significantly behind the number of instances being regulated. The European Union's General Data Protection Regulation (GDPR) came into effect in 2018 and set key guidelines, including the need for a lawful data-processing purpose, that special category data cannot be used for processing without explicit consent, and the right to object to automated decision-making [29]. The principles of these provisions might in principle seriously limit emotional profiling. In practice, GDPR's limitations have been offset by the legitimate interests provision, by consent architectures relying on dark patterns, and by enforcement capacity that has

proven to be insufficiently powerful in comparison with the complexity of what is expected to be regulated by GDPR [30].

The EU AI Act (2024) provides more focused provisions. It defines AI systems capable of manipulating behaviour using subliminal methods, which makes it difficult for humans to make autonomous choices, or that take advantage of psychological vulnerabilities, as prohibited uses of AI [31]. This is important, but it poses challenging definitional questions which the text of the Act does not comprehensively answer, and it only applies to manipulation which results in, or is likely to result in, a significant harm. This leaves us with a large grey area of emotionally targeted practices below the threshold of significant harm.

The Federal Trade Commission in the USA has the jurisdiction to take action against unfair or deceptive practices, but has not established rules specifically concerning emotional targeting. California has passed privacy laws that include the California Consumer Privacy Act as well as several amendments, including a recent update to the definition of personal data that add the sensitive data categories discussed in the California Consumer Privacy Act, and those laws impose some restrictions on the commercial use of sensitive personal data, however emotional inferences based on behavioural data are not explicitly identified as falling under the sensitive data categories described in the California Consumer Privacy Act.

3.4.2 India's regulatory framework and its gaps

The regulatory landscape for emotional AI in India is very underdeveloped when compared to the problem size at hand. The Information Technology Act 2000 offers a framework for combating online fraud and harmful content, but was not designed to cover AI-generated or AI-targeted content and does not specifically mention harmful effects of emotional manipulation [32]. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021 provide guidelines on grievance redress mechanisms and specified deadlines for removal of harmful content from online platforms; however, they do not cover the recommendation systems and targeting mechanisms that cause emotional harm on a systemic basis as opposed to individual pieces of harmful content.

The Digital Personal Data Protection Act (DPDP) 2023 is the most crucial legislative development in recent times. It sets out standards for data principal rights, standards for data fiduciary obligations and standards for the processing of sensitive personal data, and the standards on consent and the limitation of purpose have the potential to limit the commercial use of emotionally inferred data in principle [33]. Emotional/psychological data was not, however, explicitly designated as a sensitive category, and the rules for the implementation of the Act were still incomplete as of mid-2024, while the Data Protection Board created by the Act was still to implement enforcement mechanisms. The implementation potential of the legislated measures is far from realised.

The Telecom Regulatory Authority of India has been dealing with the issue of algorithmic content governance, albeit intermittently, especially when it comes to misinformation on OTT platforms, without providing any specific guidelines on emotional targeting [34]. As the Competition Commission of India has pointed out in its 2022 order, India's competition regulatory regime could be used to address data-related platform practices such as forced data sharing arrangements in its 2021 order finding that WhatsApp's Privacy Policy changes of 2021 contravened competition law. This is a possible way to tackle emotionally targeted advertising that uses information collected through coercive and deceptive consent processes, but one that has not been systematically explored.

The outcome is a regulatory system that allows emotional manipulation on a large scale to happen with little regulation. AI technology is used in emotional targeting, affecting India's vast youthful and tech-savvy population, who are vulnerable to such targeting by global platforms, which are more heavily regulated in European jurisdictions than in India. It's not just about regulatory lag. It illustrates the "dialectic" between India's desire to have a large domestic digital economy and to limit the harms associated with the business models that feed the digital economy.

3.5 Recommendations

A number of recommendations can be drawn from this analysis. First, transparency requirements in relation to recommendation systems, emotional analytics, and behavioural targeting should be mandatory. Secondly, independent algorithmic audit frameworks should be established to analyse discriminatory effects and exploitative behavioural architectures. Thirdly, the use of emotional inference techniques should be regulated with enhanced mental privacy protections. Fourthly, persuasive design techniques that are designed to foster compulsive behaviours should be subject to stricter consumer protection policies.

Finally, digital literacy campaigns should be launched to enable users to understand how their behaviour is being manipulated by algorithms.

4. CONCLUSION

The emotional manipulation economy signifies a radical shift in the nature of the intersection between technology, business, and human behaviour. Artificial intelligence systems have started functioning as behavioural infrastructures with the ability to predict, influence, and structure emotions. The commercial value of emotional engagement lies in its ability to increase the efficiency of advertisements, behavioural predictability, and platform dependence. This paper has shown that emotionally manipulative systems depend on behavioural profiling, recommendation systems, emotional analysis, persuasive architecture designs, and predictive modelling. All of these systems work under economic conditions in which emotionally intense systems are economically rewarding, as emotionally stirred users are more likely to be engaged and commercially viable.

Ethical considerations are significant here. Emotion manipulation infringes on autonomy since decisions are made under hidden influencing techniques. Emotional surveillance breaches mental privacy and psychological integrity. Algorithmic amplification is one of the reasons behind misinformation, polarisation, digital addiction, and loss of trust. Children and psychologically vulnerable people have extra risk factors in an algorithmically designed environment. Existing regulations are ineffective as consent-based paradigms cannot regulate behavioural manipulation. This calls for new regulatory approaches that should go beyond privacy concerns.

Artificial intelligence should not be considered only from the perspective of technological pessimism. Indeed, there are many positive aspects of AI which can help us make progress in healthcare, education, accessibility, and communication. What is at stake is to ensure that innovation does not get ahead of human dignity, which is at stake. Protecting emotional autonomy is crucial for maintaining democracy and individual freedom in an algorithmic age. If our emotions start being considered as resources that need to be continually harvested, analysed, and manipulated for profit, democratic citizenship might start to erode slowly but surely.

ACKNOWLEDGEMENT

None.

REFERENCES

- [1] Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.
- [2] Pandey, J. K., & Kumar, R. (2025). Governing the algorithmic agent: Confronting overt and covert challenges to justice and the future of work. *International Journal of Law Management & Humanities*, 8(4), 1974–1984. <https://doi.org/10.1000/IJLMH.1110648>
- [3] Lalitha Kumari, P., Pandey, J. K., Verma, K., Naidu, G. T., Gupta, A., & Alazzam, M. B. (2023). Impact of artificial intelligence on employees' performance in the information technology sector. *AIP Conference Proceedings*, 2821, 020004. <https://doi.org/10.1063/5.0150403>
- [4] Srnicek, N. (2017). *Platform capitalism*. Polity Press.
- [5] Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralised content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>
- [6] Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 9216–9221. <https://doi.org/10.1073/pnas.1804840115>
- [7] Horwitz, J., & Hagey, K. (2021, September 14). Facebook says its rules apply to all; company documents reveal a secret elite that's exempt. *The Wall Street Journal*. <https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353>
- [8] Picard, R. W. (1997). *Affective computing*. MIT Press.
- [9] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>

- [10] Torous, J., & Hsin, H. (2018). Empowering the digital therapeutic relationship: Virtual clinics for digital health interventions. *NPJ Digital Medicine*, 1(1), 1–3. <https://doi.org/10.1038/s41746-018-0028-2>
- [11] Arora, P. (2019). *The next billion users: Digital life beyond the West*. Harvard University Press.
- [12] Stray, J. (2021). Optimising for engagement: Understanding the use of persuasive technology on internet platforms. arXiv. <https://arxiv.org/abs/1911.09011>
- [13] Rao, M. (2018). *Anatomy of hate*. Westland Books.
- [14] Wu, L., Morstatter, F., Carley, K. M., & Liu, H. (2019). Misinformation in social media: Definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter*, 21(2), 80–90. <https://doi.org/10.1145/3373464.3373475>
- [15] McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139–153). Guilford Press.
- [16] Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 1036–1040. <https://doi.org/10.1073/pnas.1418680112>
- [17] Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- [18] Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences*, 114(48), 12714–12719. <https://doi.org/10.1073/pnas.1710966114>
- [19] Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., & Rentfrow, J. (2018). Musical preferences predict personality better than estimated using self-reports. *Psychological Assessment*, 30(6), 731–747. <https://doi.org/10.1037/pas0000495>
- [20] Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online behavioural advertising: A literature review and research agenda. *Journal of Advertising*, 46(3), 363–376. <https://doi.org/10.1080/00913367.2017.1339368>
- [21] Calo, R. (2014). Digital market manipulation. *George Washington Law Review*, 82(4), 995–1051.
- [22] Advertising Standards Council of India. (2023). Guidelines for advertising involving artificial intelligence and influencer content. ASCI. <https://www.ascionline.in/asci-guidelines>
- [23] Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- [24] Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalisation in American politics*. Oxford University Press.
- [25] Sharma, N. (2024, April 3). AI-generated audio of politicians floods WhatsApp ahead of Indian election. Reuters. <https://www.reuters.com/technology/ai-generated-audio-politicians-floods-whatsapp-ahead-indian-election-2024-04-03/>
- [26] Election Commission of India. (2024). Guidelines on the use of artificial intelligence and social media in election campaigns. ECI. <https://www.eci.gov.in/guidelines-on-ai-social-media>
- [27] Banaji, S., Bhat, R., Agarwal, A., Passanha, N., & Sadhana Pravin, M. (2019). WhatsApp vigilantes: An exploration of citizen reception and forwarding of WhatsApp misinformation linked to mob violence in India. London School of Economics and Political Science.
- [28] Diakopoulos, N. (2020). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- [29] European Parliament and the Council of the European Union. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation). *Official Journal of the European Union*.
- [30] Zuiderveen Borgesius, F. J., Moller, J., Kruikemeier, S., Bol, N., Lelkes, Y., Heinderyckx, F., & De Vreese, C. H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, 14(1), 82–96. <https://doi.org/10.18352/ulr.420>
- [31] European Parliament. (2024). Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.
- [32] Ministry of Law and Justice, Government of India. (2000). The Information Technology Act, 2000 (Act No. 21 of 2000). Gazette of India.

- [33] Ministry of Law and Justice, Government of India. (2023). The Digital Personal Data Protection Act, 2023 (Act No. 22 of 2023). Gazette of India.
- [34] Telecom Regulatory Authority of India. (2022). Recommendations on regulatory framework for over-the-top communication services. TRAI. https://www.trai.gov.in/sites/default/files/Recommendations_OTT_2022.pdf