LEGAL AMENITY LAW JOURNAL – ISSN 3107-9148

Volume II Issue III May – June 2026

ARTIFICIAL INTELLIGENCE: DEEPFAKES, BEHAVIORAL TARGETING, AND THE DIMENSIONS OF AI MISUSE

By: Aryan Mehta

Law Student, Government Law College, Mumbai

ABSTRACT

In the contemporary world, Artificial Intelligence (AI) occupies a peculiar position (celebrated as a tool of limitless possibility but also often condemned as an instrument of unprecedented misuse). This article aims to examine the tension between these two realities by tracing the trajectory of AI abuse across various domains: surveillance and facial recognition, behavioural targeting, and the creation of synthetic media through deepfake technology.

At the core of these abuses lies a common denominator i.e. data. For AI systems to function, it requires pre-existing data to train it, and in an increasingly interconnected world, the movement of data occurs at a scale that most users neither comprehend or consent to. The consequences of the structural opacity ranges from state overreach (in the context of facial recognition software), to the manipulation of democratic processes through AI backed micro targeting, and non-consensual sexualisation of individuals through the use of deepfake technology.

India's case presents a particularly interesting case study. This article examines certain aspects of the NITI Aayog's National Strategy for Artificial Intelligence (2018), the Fact Check Unit controversy, the Digital Personal Data Protection Act, 2023, and the proposed AI (Ethics and Accountability) Bill, 2025. India's legislative and judicial response to the governance, development, and misuse of AI reflects a broader global struggle to keep pace with the technology that states seek to govern.

I. UNDERSTANDING AI AND THE DATA PROBLEM

Artificial Intelligence cannot be boiled down to a single technology. Rather, it is to be understood as a broad field comprising of several areas – machine learning, natural language processing, deep learning, etc. And at its core, is the capacity of AI to learn and improve over time. This capacity is grounded in data, where AI systems require training on pre-existing datasets in order to generate outputs, make decisions, identify patterns, etc.

While it could be argued that methods like unsupervised learning and reinforcement learning offer paths to training AI by eliminating the need of large datasets, this is only partially accurate. Unsupervised learning requires raw, unlabelled data without pre-defined answers to be fed into the system and the AI relies on patterns and relationships within the data to learn. On the other hand, reinforcement learning involves providing the AI with a goal and allows it to interact with the environment through a trial-error method, the AI interacts with the environment and learns. In both the cases, some form of data remains a prerequisite.

In a world shaped by the Internet of Things, practically every interaction comprises movement of data either internally or externally, where devices communicate with one another by converting real-world variables into transmissible data, the volume and intimacy of information shared by everyday interactions is staggering. In this context, most users remain unaware that the files they produce: photographs, videos, audio recordings carry not just content but also metadata. Metadata includes information about the device used, settings applied, location of the user at the time of file creation, and even the user details. Permissions for such data collection are often buried in fine lining of terms of service agreements, or activated by default. And the average user often lacking the awareness to identify, or technical literacy to object, inadvertently allows the collection of such personal data.

The problems examined in this article surrounding the abuse of AI technologies are not anomalies in an otherwise well-functioning digital ecosystem. It is rather this technological architecture where extraction and use of data is the norm, combined with technical illiteracy, and lacuna in the law governing such technologies that creates conditions for AI misuse.

II. SURVEILLANCE, FACIAL RECOGNITION, AND GOVERNMENT OVERREACH

THE ARCHITECTURE OF CLEARVIEW AI

The case of Clearview AI in the United States of America offers an illustration of how AI surveillance tools can erode civil liberties. Clearview's methodology was relatively simple: it assembled a database utilising over 3 billion photos from publicly accessible websites, generated vector representations of the faces in those images, and allowed users to conduct a reverse image search by uploading a target photograph and the AI would retrieve and display all nearby vector matches.

Police departments across several American states adopted the technology on an experimental basis. The use of Clearview AI would be a departure from the conventional government databases and would grant them access to a database sourced from social media platforms and enabled them to identify individuals who had never been formally processed by the state.

THE PITTSBURGH CONTROVERSY AND POLITICAL SUPPRESSION

The use of Clearview technology by police departments raised concerns of government overreach amongst the public, where such applications could potentially be used to target dissenting individuals and suppress opposition. The concerns of Clearview's potential misuse crystallised into reality (allegedly) in 2021, when it was revealed that officers within the Pittsburgh Police Department had accessed Clearview's facial recognition technology during the Black Lives Matter protests. While the records did not establish whether the use of such technology was done in direct relation to the protest, the disclosure exposed the ease with which AI backed surveillance tools could be turned against citizens engaged in constitutionally protected expression to suppress free speech and crush dissent.

TELANGANA POLICE DATA LEAK

While Clearview AI controversy depicted the potential for abuse of AI Surveillance technology, with respect to government overreach, the breach of Telangana Police's TSCOP application demonstrates a distinct risk. Telangana Police's TSCOP application was a facial recognition

system, often used to identify criminals, which was trained using internal government databases. In 2024, when hackers penetrated the system, they accessed details such as offender records, police gun licenses, and other sensitive law enforcement information. Samples of user information such as officer names, police station affiliations, designations, and images were subsequently posted on forums to attract prospective buyers.

The incident underscored a structural irony in the deployment of AI for security purposes where a system designed and built to identify potential threats itself became a source of threat placing law enforcement personnel and integrity of criminal records at risk. It raised concerns around the adequacy of cybersecurity frameworks governing state-operated AI systems, particularly when those systems contain sensitive personal information of offenders and officials.

III. BEHAVIOURAL TARGETING, POLITICAL MANIPULATION, AND THE CAMBRIDGE ANALYTICA SCANDAL

THE MECHANISM OF BEHAVIOURAL TARGETING

Another significant AI driven technology is Behavioural Targeting. It is typically a user centric feature based on an algorithm that personalises recommendations and gives ‘smart’ results aimed towards providing users with tailored content based on their likes/dislikes. It functions as a system of inference and influence, mapping user behaviour to build predictive profiles which can be used to shape the information or content to which the users are exposed.

Behavioural targeting mechanism includes aggregation of user data which includes but is not limited to posts, likes, patterns of engagement, duration of attention and translation of this behavioural pattern into profiles with scores, assessment of personality traits, political inclinations, and positions on key issues. With respect to the abuse of this mechanism, the value of these profiles lies in the synthetic delivery of customised content to reinforce beliefs, and nudge opinion in a particular desired direction.

CAMBRIDGE ANALYTICA SCANDAL AND THE INTERFERENCE IN DEMOCRATIC PROCESS

A pressing example highlighting abuse of behavioural targeting is the Cambridge Analytica Scandal. The scandal involved unethical harvesting of Facebook user data through a personality quiz app on Facebook that exploited a loophole in Facebook's application programming interface to harvest not just the user's data but also their friends' data. This massive data which included posts, their likes, their patterns, etc was used to create their profiles which were given scores to assess their personality traits.

What the users failed to realise was that this profile was used to predict a user's political inclination and their stance on key issues and the data was used to target the users with customized ads, which were used to reinforce existing beliefs or sway opinion. This is said to have played a significant role in influencing behaviour of public during democratic processes including BREXIT campaign and US presidential elections.

With just about 2,70,000 people consenting to the use of their data by using the quiz, the data of over 87 million people was tapped into. The Cambridge Analytica case is significant not merely for the scale of its violation but for the clarity with which it demonstrates how AIenabled behavioural analysis can be operationalised as a tool of political engineering. The manipulation was effective precisely because it was invisible: users encountered content that appeared organic and personally resonant, without any indication that it had been crafted to produce a specific response.

IV. SYNTHETIC MEDIA AND DEEPFAKES

THE TECHNOLOGY AND ITS APPLICATIONS

The term 'deepfake' is a combination of deep (from Deep Learning) and fake. It utilizes Deep Learning models and creates images, videos, or audio mimicking a real human or an inanimate object, making them appear to say or do something that never happened. It has the capability of creating media of significant verisimilitude, where its application can span from legitimate and useful to profoundly harmful.

DEEPFAKES IN BROADCASTING

Deepfake technology has proven genuine utility in broadcasting and accessibility. The India Today Group in India which runs news channels like Aaj Tak has made use of an AI news anchor called as Sana to present news which has not only made it possible for 24/7 news delivery, with multi-lingual capabilities, ability to deliver consistent and rapidly updating content, and even increasing cost-efficiency.

IMPACT OF DEEPFAKES ON THE ELECTORAL INTEGRITY IN INDIA

In the recent past, social media has played a significant role in spreading democratic awareness and influencing political opinions. In this context, where social media plays a crucial role in electoral outcomes, elections have become a theatre for deployment of deepfake technology. In one such instance is when a video of a woman going by the name of Duwaraka was circulated, but what is intriguing is that the said woman was considered to have been killed in an airstrike in 2009. In another instance, deepfakes of famous actors like Ranveer Singh and Amir Khan were circulated prior to the General Elections criticising the Prime Minister, and endorsing an opposition party.

It is a known fact that elections are often accompanied by misinformation campaigns, but the use of such AI driven technology escalates the challenge of controlling misinformation. Election experts have recognised that, in a sufficiently saturated environment of synthetic media, the very authenticity of real footage becomes contestable, creating conditions in which even legitimate reporting can be dismissed and uprisings become conceivable.

NON-CONSENSUAL DEEPFAKE PORNOGRAPHY AND VIOLATION OF DIGITAL PERSONHOOD

At the helm of the most serious dimensions of deepfake abuse lies the deployment of deepfakes as a vehicle for gender based digital violence. According to UN Women, about 90-95% of all online deepfakes are non-consensual pornographic images with about 90% depicting women. The consequences of such deepfakes extend beyond the digital realm; they often result in loss of jobs, separation from family, isolation by the society, and destruction of reputation.

In the case of Shilpa Shetty vs. Getoutlive.in, where deepfakes of the actress Shilpa Shetty were created and shared online, the Bombay High court while granting an injunction held that such non-consensual deepfake pornography was an assault on a woman's digital personhood. It held that no person can be portrayed in a fashion that would affect their right to life as guaranteed under Article 21 of the Constitution of India and that their right to live with dignity cannot be compromised in any manner. Such an act which amounted to reconstructing a person's identity without consent was a violation of their digital personhood.

V. INDIA'S ACTIONS: POLICY, LEGISLATION, AND THE WAY FORWARD THE NITI AAYOG NATIONAL STRATEGY (2018)

In June, 2018, NITI Aayog the Think-Tank of India released a National strategy For Artificial Intelligence. The document acknowledged the possibilities of AI deployment across key sectors, identified structural impediments to AI development in India, and offered recommendations for a national approach to the technology.

Among its observations and reasoning, was the identification of a 'black box' problem surrounding AI which highlighted the structural opacity of AI systems where the users could not trust the results given by the AI due to lack of understanding of the process for the output solution given by the AI.

Furthermore, it also recognised the importance of creating a level playfield for startups and research with respect to data access and suggested the establishment of a global marketplace for AI, where the data should be stored in a decentralised format, stored and accessed through blockchain technology instead of a centralised repository to prevent abuse and alleviate privacy or security concerns. The strategy paper also recommended the use of manual labour in annotation mechanisms to ensure the verifiability of data and balancing costs.

THE FACT CHECK UNIT: GOVERNMENTAL OVERREACH AND ITS CONSTITUTIONAL CORRECTION

In 2019, a Fact Check Unit was formed under the aegis of the Press Information Bureau (PIB) with a mandate to act as a deterrent to creators and disseminators of misinformation. The Fact Check Unit was empowered to label content as fake, false, and misleading, and could compel digital platforms to remove such content. And the Government, by a press release dated 20th

March 2024, notified the Fact Check Unit as a fact check unit of the central government which extended it powers to censor content.

The constitutionality of this action came to be questioned. This essentially made the Government an arbitrator of truth of the content relating to its own conduct, with the overextending power to compel obedience and demand compliance, a clear violation of the fundamental right of free speech.

This notification was stayed by a division bench of Chief Justice D. Y. Chandrachud and Justice J. B. Pardiwala of the Hon'ble Supreme Court of India until final decision of the Bombay High Court on the pending matter. Justice Chandurkar acted as a tie-breaking judge to break the tie of the split verdict delivered in this matter, and struck down the 2023 amendment, which empowered the central government to establish the said fact check unit. He stated that the amendments in his opinion were violative of Article 14 and 19 of the Constitution of India.

The Fact Check Unit episode is a significant case study highlighting the limits of state sanctioned regulatory response to combat misinformation. While not specifically created to combat AI driven misinformation campaigns, the Fact Check Unit would have played a significant role in combating the same as evident from the scale of misinformation documented above. The instinct to arm the government with the power to police such misinformation in the digital landscape is understandable but it faces constitutional barriers when the state deploys that power to insulate itself from criticism.

THE DIGITAL PERSONAL DATA PROTECTION ACT, 2023

The Digital Personal Data Protection Act (DPDPA), 2023, is currently India's most substantive legislative intervention in the government of data. It establishes a consent-centric mechanism requiring informed consent for processing of personal data, and lays down data fiduciary obligations and penalties for violation of such obligations.

While the DPDPA is a step forward, it is not the comprehensive solution to address the problem of AI misuse. While the DPDPA's consent architecture addresses regulated data processing relationships, there remains a lacuna in addressing the practice of scraping publicly accessible data particularly surrounding section 3(c) and section 17(2)(b) which outline exemptions where

the Act would and would not apply and the exemption for research, archiving, and statistical processing respectively.

THE DIGITAL INDIA ACT AND AI (ETHICS AND ACCOUNTABILITY) BILL, 2025

The erstwhile Information Technology Act, 2000 is set to be replaced by the Digital India Act, which is expected to contain specific provisions for AI and emerging technologies, increasing accountability, enhanced intermediary obligations, and also advancing towards India's goal of becoming a $1 trillion economy.

Another significant piece of legislation that is awaited is the AI (ethics and accountability) Act which was introduced in 2025 in the Lok Sabha as a bill. It is proposed to establish a framework for responsible AI use, ensuring transparency and accountability, and is largely aligned with the proposed solutions and the roadmap for the creation of an ecosystem for the development of AI technology in India as envisioned in the Responsible AI for All, an approach document for India released by NITI Aayog in 2021.

Whether the Bill, once enacted, will translate its stated commitments into enforceable obligations remains to be seen. Another issue that may potentially create issues is the use of AI in various fields, and accountability for any incidents arising therefrom. It is also necessary to enforce certain measures, such as mandatory registration of AI technologies and regular audits to be conducted to assess their safety and track potential breaches.

CONCLUSION

The technologies examined hereinabove are not inherently malign, but have demonstrated legitimate application with potential to address genuine needs. And the problem lies not in the technology itself but rather in the conditions in which it is deployed and governed, and most of all the tendency of state and commercial actors to prioritise benefits and capability over ethics, morality and accountability.

India's regulatory response is evolving. However, the difference in pace of legislative development and technological advancement remains wide. What is required is not merely judicial intervention, while important as that is – there is a greater need for clarity in the

legislation to eliminate the potential grey areas surrounding the governance of these technologies.

REFERENCES

1. Reddy, US 2024, *Data leak: Cops shut sites, apps for security overhaul*, The Times of India, The Times Of India,

   https://timesofindia.indiatimes.com/city/hyderabad/telangana-police-data-leak-sitesand-apps-shut-down-for-security-overhaul/articleshow/110807834.cms
2. Won Kyung Jung and Hun Yeong Kwon. 2024. Privacy and data protection regulations for AI using publicly available data: Clearview AI case. In Proceedings of the 17th International Conference on Theory and Practice of Electronic Governance (ICEGOV '24). Association for Computing Machinery, New York, NY, USA, 48–55. https://doi.org/10.1145/3680127.3680200
3. Canon, and Marcus Hawkins. "Understanding EXIF and Metadata." Canon Europe, 2023. https://www.canon-europe.com/pro/infobank/all-aboutexif/#:~:text=Every%20time%20you%20take%20a,keywords%20and%20captions%2 C%20and%20more..
4. Staff, Coursera. "What Is the Internet of Things (IoT)? With Examples." Coursera, 2024. https://www.coursera.org/in/articles/internet-of-things.
5. Unsupervised learning. "Can an AI Be Trained without Data? | Natif.ai." natif.ai, May 24, 2024. https://natif.ai/can-an-ai-be-trained-without-data/.
6. Connell, Eileen. "How AI Works | CSU Global." Colorado State University Global, August 11, 2025. https://csuglobal.edu/blog/how-does-ai-actually-work.
7. Stryker, Cole, and Eda Kavlakoglu. "Artificial Intelligence." Ibm.com, August 9, 2024. https://www.ibm.com/think/topics/artificial-intelligence.
8. Shireen Yachu. "Explained: The Digital India Act 2023." Vidhi Centre for Legal Policy, August 8, 2023. https://vidhilegalpolicy.in/blog/explained-the-digital-india-act2023/.
9. "Approach Document for India Part 1 -Principles for Responsible AI FEBRUARY

2021 RESPONSIBLE AI #AIFORALL," n.d. https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.

10. "National Strategy for Artificial Intelligence #AIFORALL June 2018" https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf

11. Laws, AI. "AI Laws and Regulations in India as of 2026." Prashantmali.com, 2026. https://www.prashantmali.com/cyber-law-blog-india/ai-laws-and-regulations-in-indiaas-of-2026.

12. Whitecase.com. "AI Watch: Global Regulatory Tracker - India | White & Case LLP," May 13, 2024. https://www.whitecase.com/insight-our-thinking/ai-watch-globalregulatory-tracker-india#:~:text=Other%20laws%20affecting%20AI%20*%20The%20Information,publication%2C%20is%20yet%20to%20come%20into%20force..

13. Rihl, Juliette. "Emails Show Pittsburgh Police Officers Accessed Clearview Facial Recognition after BLM Protests." Pittsburgh's Public Source. Pittsburgh's Public Source, May 20, 2021. https://www.publicsource.org/pittsburgh-police-facialrecognition-blm-protests-clearview/.

14. Reality Defender — Enterprise-Grade Deepfake Detection. "Good Uses of Deepfakes," July 31, 2025. https://www.realitydefender.com/insights/good-uses-of-deepfakes.

15. Ritu. "Bombay High Court Orders Immediate Takedown of AI-Generated Deepfake Content Infringing Shilpa Shetty Kudra's Privacy and Dignity." SCC Times, December 29, 2025. https://www.scconline.com/blog/post/2025/12/29/bom-hc-shilpa-aigenerated-deepfake-content-scc-times/.

16. UN Women – Headquarters. "AI-Powered Online Abuse: How AI Is Amplifying Violence against Women and What Can Stop It | UN Women – Headquarters," November 18, 2025. https://www.unwomen.org/en/articles/faqs/ai-powered-onlineabuse-how-ai-is-amplifying-violence-against-women-and-what-can-stopit#:~:text=Image%2Dbased%20abuse%20through%20deepfakes,the%20individuals% 20targeted%20are%20women..

17. Sebastian, Meryl. "AI and Deepfakes Blur Reality in India Elections." Bbc.com. BBC News, May 16, 2024. https://www.bbc.com/news/world-asia-india-68918330.

18. Pib.gov.in. "Press Information Bureau," 2017. https://www.pib.gov.in/FAQ_fact.aspx?reg=3&lang=2#:~:text=The%20PIB%20Fact%20Check%20Unit%20uses%20various%20technological%20tools%20such,%E2%9D%9A%E2%9D%9A.

19. Pib.gov.in. "Government Notifies PIB's Fact Check Unit under IT Rules 2021," 2021. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2015792&reg=3&lang=2.

20. Arunima. "Bombay High Court Strikes down IT Rules Amendment 2023; Major Setback for Fact-Check Units." SCC Times, September 20, 2024. https://www.scconline.com/blog/post/2024/09/20/bombay-high-court-it-rulesamendment-2023-fact-check-units-article14-article19-legalnews/#:~:text=Fact%2DCheck%20Units,Bombay%20High%20Court%20strikes%20down%20IT%20Rules%20Amendment%202023;%20Major,chilling%20effect%20on%20free%20speech.&text=Bombay%20High%20Court:%20A%20case,chilling%20effect%20on%20free%20speech..

21. Tsaaro. "The Digital India Act: An Analysis," March 29, 2023.https://tsaaro.com/blogs/the-digital indiaact/#:~:text=1.,obligations%20apply%20to%20all%20intermediaries..

22. PRS Legislative Research. "The Digital Personal Data Protection Bill, 2023," March 29, 2026. https://prsindia.org/billtrack/digital-personal-data-protection-bill2023#:~:text=Exemptions%20to%20data%20processing%20by,independent%20funct ioning%20of%20the%20Board..

23. Amoha Basrur. "The Digital Personal Data Protection Act, 2023: Recommendations for Inclusion in the Digital India Act." orfonline.org. OBSERVER RESEARCH FOUNDATION (ORF), March 2024. https://www.orfonline.org/research/the-digitalpersonal-data-protection-act-2023-recommendations-for-inclusion-in-the-digitalindia-act#:~:text=2023%20PDF%20Download-,The%20Digital%20Personal%20Data%20Protection%20Act%2C%202023:%20Recommendations%20for%20Inclusion,October%202023%2C%20Observer%20Research%20Foundation..

24. lexanalytico. "Right to Be Forgotten in India under DPDP Act 2023." LexAnalytico – Techno-Legal & IP Consulting Firm, July 8, 2025. https://lexanalytico.com/right-tobe-forgotten-india-dpdp-act/#:~:text=The%20DPDP%20Act%20introduces%20several,those%20who%20process%20such%20data..

25. The Artificial Intelligence (Ethics and Accountability) Bill, 2025

26. seo835. "Deepfake Regulation India 2025: MeitY's Comprehensive IT Law." Khurana And Khurana, December 16, 2025. https://www.khuranaandkhurana.com/deepfakeregulation-india-2025-meity-s-comprehensive-it-rules-amendment#:~:text=The%20IT%20Act%20has%20multiple,Bharatiya%20Nyaya%20Sanhita%2C%202023.

27. "Digital Personal Data Protection Act, 2023.." *Indiacode.Nic.In*, 2023. https://doi.org/AC_CEN_45_0_00003_2023-22_1763464807080.

28. Law. "Publicly Available Data under the DPDP Act: The Limits of Exemptions in AIDriven Processing." Law School Policy Review, January 13, 2026. https://lawschoolpolicyreview.com/2026/01/13/publicly-available-data-under-thedpdp-act-the-limits-of-exemptions-in-ai-driven-processing/.

29. Website team 05. "Publicly Available Personal Data Exemption Is Unclear." Law.asia, July 31, 2024. https://law.asia/publicly-available-data-dpdpa/.

30. Jan. "Legality of Data Scraping under Indian Law." Law.asia, August 2025. https://law.asia/india-data-scraping-regulation/.