Indian Journal of Law and Technology

Volume 18 | Issue 2 Article 1

2022

Data Privacy And Elections In India: Microtargeting The Unseen Collective

Sayantan Chanda

Follow this and additional works at: https://repository.nls.ac.in/ijlt

Recommended Citation

Chanda, Sayantan (2022) "Data Privacy And Elections In India: Microtargeting The Unseen Collective," *Indian Journal of Law and Technology*: Vol. 18: Iss. 2, Article 1.
Available at: https://repository.nls.ac.in/ijlt/vol18/iss2/1

This Article is brought to you for free and open access by Scholarship Repository. It has been accepted for inclusion in Indian Journal of Law and Technology by an authorized editor of Scholarship Repository. For more information, please contact library@nls.ac.in.

DATA PRIVACY AND ELECTIONS IN INDIA: MICROTARGETING THE UNSEEN COLLECTIVE:

*Sayantan Chanda**

ABSTRACT *In India, the usage of social media to reach hundreds of millions of active online users is common across political parties. With revelations regarding data mining being undertaken by political parties across the world, there is a need for robust data privacy not only to protect individuals, but also to ensure free and fair elections. In this context, the importance of the data collected lies in the inferences it allows a data fiduciary to draw about the person whose data is collected. Access to private details, through mining of online data from social networks and other sources, allows individuals to be aggregated into unseen collectives, purely on the basis of specific data points, and for them to be given targeted and even false messages. Most importantly, this is a problem on a societal scale, as micro-targeting occurs across large groups of people and not merely at the individual level. Thus, the individual centric focus of data privacy law is insufficient when the target of manipulation is not one individual, but entire groups or collectives of people.*

This paper will highlight how both the Data Privacy Bills as introduced by the Indian Government in 2019 and 2022, fail to account for the collective privacy of citizens and how the rights provided do not address the problem of inferences. To that end, a move away from individual privacy and toward collective privacy will be proposed which can protect individuals who are assimilated unknowingly into collectives that are based on mined data.

* Judicial Law Clerk-cum-Legal Researcher at the Supreme Court of India. Undergraduate law degree from O.P. Jindal Global Law School. Views are personal. Feedback is welcome at: sayantan122194@gmail.com. The author would like to acknowledge the efforts and assistance of the peer-reviewer and the editorial team at the NLS Indian Journal of Law and Technology, whose comments were invaluable in refining this work. The author would additionally like to acknowledge Muskan Tibrewala (Advocate at the Delhi HC & Supreme Court), Aiswarya Murali (Judicial Law Clerk-cum-Legal Researcher at the Supreme Court) and Ashish Matthew (former Analyst at a political consultancy) for their advice/assistance on this project.

I. Introduction

The Personal Data Protection Bill, 2019 ('PDP Bill' or 'Bill') went through various revisions over the past 3 years. It was the subject of much discussion among privacy activists, industry heads, government departments, and consumers. The discussions led to a second draft being issued in 2021, however, this also proved inconclusive. It seemingly did not address the entire ambit of concerns that were raised regarding some of the potential drawbacks of the Bill. Hence, on August 3, 2022, the PDP Bill, was withdrawn and it was announced that a new bill would be tabled soon with substantial alterations.[1] It is disappointing that after 3 years of deliberations, the PDP Bill has now returned to the drawing board. However, this turn of events presents an opportunity as well to highlight certain issues with the Bill. The revised Personal Data Protection Bill, 2022 ("2022 Bill" or "PDP Bill, 2022"), was duly introduced which cut out various excessive measures such as data residency requirements within the country and criminal penalties.[2] However, these changes are primarily targeted toward easing the privacy related compliance requirements for commercial actors.[3] In other contexts, various loopholes remain that require addressing. Specifically with regard to

---

1 The Hindu Bureau, 'Union Government Rolls Back Data Protection Bill' *The Hindu* (New Delhi, 3 August 2022), <https://www.thehindu.com/news/national/union-government-rolls-back-data-protection-bill/article65721160.ece> accessed 8 August 2022.

2 Sourabh Lele, 'Nasscom Hails Draft Data Protection Bill for Dropping Contentious Rules' *The Business Standard* (New Delhi, 5 December 2022), <https://www.business-standard.com/article/economy-policy/industry-body-nasscom-welcomes-draft-digital-personal-data-protection-bill-122120501098_1.html> accessed 9 December 2022.

3 Hemant Kashyap, 'Data Protection Bill: From Deemed Consent to Exemptions, Lack of Clarity May Hurt the Cause' *Inc42* (New Delhi, 5 December 2022), <https://inc42.com/buzz/data-protection-bill-deemed-consent-exemptions-lack-clarity-hurt-cause/> accessed 8 December 2022.

microtargeting, the 2022 Bill, in fact, constitutes a step in the wrong direction.

The importance of data privacy rights has increased with the progress of technology and the increasing digitization of society. The fact that almost all forms of economic activity and human interaction have shifted online has raised several concerns regarding the security of peoples' personal data. Not just protestors, but consumers, researchers, academics, and governments themselves, have taken a keen interest in regulating privacy rights. The interests of these stakeholders are often at odds. While members of civil society desire greater privacy coverage, law enforcement would like access to as much information as possible. While consumers browse the products available to them on Amazon and eBay, they worry about the amount of personal data these tech giants are accumulating. Moreover, while governments take a dim view of their own citizens' privacy rights, they themselves wish to maintain utmost secrecy regarding their own activities with such data.

These complex inter-relationships between stakeholders lead to significant legal and policy implications. One relationship which requires particular scrutiny is that between political parties and the electorate. An example of why this is relevant may be drawn from the Cambridge Analytica scandal associated with the 2016 United States Presidential Elections.[4] The culpability of Facebook in failing to protect private information led to widespread condemnation of the social media giant.[5] Cambridge Analytica took advantage of Facebook's Open Graph platform to harvest information about millions of users.[6] Having created profiles of these individuals based on this information which included their social background, the posts they 'liked', the comments they made and put on their respective Facebook Wall etc., this data was then sold to different political campaigns, including Donald Trump's and was then exploited by these Presidential candidates to target these individuals with targeted messages.[7] Consequently, having this infor-

---

4 Scott Detrow, 'What Did Cambridge Analytica Do During the 2016 Election?' *NPR* (20 March 2018) <https://www.npr.org/2018/03/20/595338116/what-did-cambridge-analytica-do-during-the-2016-election> accessed 28 October 2022.

5 Emma Graham-Harrison & Carole Cadwalladr, 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach', *The Guardian* (London, 17 March 2018) <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> accessed 24 September 2021.

6 David Ingram, 'Zuckerberg Apologizes for Facebook Mistakes with User Data, Vows Curbs' *Reuters* (21 March 2018 <https://www.reuters.com/article/us-facebook-cambridge-analytica-idUSKBN1GX0OG> accessed 24 September 2021.

7 Nicholas Confessore,'Cambridge Analytica and Facebook: The Scandal and the Fallout So Far', *New York Times* (New York, 4 April 2018) <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html> accessed 26 September 2021; Allan Smith, 'There's an Open Secret about Cambridge Analytica in the Political World: It

mation at hand, gives a prospective party/candidate, an advantage over competitors. This allows the campaigner in question to target each individual with the kind of messaging and advertising that is most persuasive to them.[8] This process is known as 'Micro-Targeting'.[9]

The use of personal data for political campaigning has not received the scrutiny that it deserves, in India. For example, the state assembly elections in Bihar featured mass usage of online platforms for campaigning.[10] The importance of this in the electoral space is such that almost every political party now requires a comprehensive database of personal information regarding voters in order to be effective on election day.[11] One may even say that now personal data is the primary commodity for those wishing to contest elections.

Therefore, the legal framework for privacy is of great importance for regulating the use of data in elections. While the PDP Bill, 2019 is no more, it is necessary to identify the potential loopholes in the Bill that could have allowed parties to collect data *enmasse* (in large numbers) for their campaigns. Further, it is even more important to highlight the advantages the 2019 Bill had over the current 2022 iteration. The inception of the analysis will focus upon how the 2019 Bill, though not without its shortcomings, remains a far better standard for data privacy laws and its ability to address issues such as microtargeting. However, while noting the primacy of the 2019 Bill over the 2022 Bill in this context, the primary argument is that it would have also be necessary to examine the gaps in the former that required redressal from the perspective of microtargeting and political manipulation.

There has been limited investigation of this in an Indian legal context,[12] even though the conversation on this topic has reached an advanced stage

---

Doesn't have the "Secret Sauce" it Claims' (*Business Insider*, 21 March 2018) <https://www.businessinsider.in/tech/theres-an-open-secret-about-cambridge-analytica-in-the-political-world-that-sheds-new-light-on-the-facebook-data-scandal/articleshow/63402917.cms> accessed 26 September 2021.

8 Sandra C Matz and others, 'Psychological Targeting as an Effective Approach to Digital Mass Persuasion' (*Proceedings of the National Academy of Sciences* 2017) <https://www.pnas.org/content/114/48/12714> accessed 25 September 2021.

9 Ira Rubinstein, 'Voter Privacy in the Age of Big Data' (2014) 5 Wisconsin Law Review 861, 882.

10 Amita Tagore, 'The Digital Campaign', *Indian Express* (New Delhi, 27 October 2020) <https://indianexpress.com/article/opinion/the-digital-campaign-bihar-assembly-election-6901647/> accessed 23 September 2021.

11 Nikhil Pahwa, 'The Election Commission of India Needs to Restrict Political Usage of Data' (*Medianama*, 20 June 2019) <https://www.medianama.com/2019/06/223-the-election-commission-of-india-needs-to-restrict-political-usage-of-data/> accessed 20 September 2020.

12 ibid.

in other parts of the world.[13] It is necessary for the Indian legal sphere to address this issue with greater urgency, given the hundreds of millions of Indians who lead active lives online. The question that this paper will seek to address is whether the PDP Bill, 2019, and to a lesser extent the newer version introduced in 2022, provide sufficient protection against mass collection of personal data by political parties and how privacy rights may be shaped to protect against microtargeting. The utility of this exercise is to highlight deficiencies in the now withdrawn Bill that may, hopefully, be addressed in the scheme and provisions of the future data privacy bill that is currently being considered. However, in order to do so, it is also necessary, as touched upon, to point out the numerous ways that the PDP Bill, 2022 is a regression from the earlier model of the Bill.

This paper will attempt to address this question by first providing an overview of the level of digital penetration in Indian society. It will also focus on the increasing usage of data analytics for the purpose of refining campaign advertising by political parties in the country. In this first part, it will conclude with looking at the both the promises and drawbacks of microtargeting in the context of elections.

The second part will go on to examine the state of the law vis-à-vis data privacy in the electoral sphere, in other jurisdictions before examining the problem that inferences pose to proper data regulation and enforcement of data rights in India. In doing so, it will elaborate on how the PDP Bill in 2019 did not sufficiently address the issue of inferences and merely including it under the ambit of personal data in Clause 3(28) of the PDP Bill is insufficient. Clause 3(28) refers to the different forms of data that come under the ambit of 'personal data' and lists inferences under it. The protection for such inferences under the 2022 Bill has been negated completely. Clause 2(13) of the PDP Bill, 2022,[14] makes no reference to inferences specifically and reduces personal data to only those forms of data which make a person identifiable. Hence, though flawed, Clause 3(28) of the PDP Bill, 2019, still provided a form of recognition to inferences which has now been done away with entirely.

The third and final part will propose the steps that can be taken, both by the Data Protection Authority that was to be set up under the PDP Bill, and the Election Commission ("EC"), to properly deal with the threat of data collection and political microtargeting. Fundamental to this, will be the

---

[13] Normann Witzleb, Moira Paterson and Janice Richardson, *Big Data, Political Campaigning and the Law* (Taylor & Francis 2018) Part 3.

[14] Digital Personal Data Protection Bill 2022, cl 2(13) ("Personal Data Bill 2022").

need for the DPA to lay out codes of good practice, and for the recognition of the concept of 'collective privacy'. This part will show that the recognition of "collective privacy" is the most appropriate way to protect constituents and that collective privacy should be incorporated into the revised PDP Bill which is currently being considered.

For undertaking this evaluation, there will be a focus upon the 2019 and 2022 versions of the PDP Bill. These two draft legislations, being drafted and tabled by the Government of India itself, are the most appropriate for analysing the evolution of the focus and thought process behind data privacy law in India. The critiques of the 2019 Bill led to the unveiling of the more recent 2022 iteration. However, as will be elaborated upon, the latest incarnation of the PDP Bill constitutes a decline in the level of protection that is necessary to appropriately address the problem of microtargeting. Conversely, the earlier 2019 Bill, even with its drawbacks, was preferable. Before proceeding, an acknowledgement of the 2018 Draft by Justice B.N. Srikrishna and the Committee of Experts on a Data Protection Framework in India, and the Joint Parliamentary Report in 2021 on the 2019 version of the PDP Bill, is necessary. However, this paper will not discuss these, given that these suggestions while valuable, were never endorsed by the government of India and did not directly address the specific problems of data privacy in the context of elections. In fact, it appears that this particular danger with regard to microtargeting and manipulation of voters during elections has largely evaded attention so far. Consequently, to maintain focus on this specific subject matter, the actual Bills which have been officially tabled and considered by the Government of India, will remain the centre of attention in this paper.

Part I

II. Elections in a Digital India

India's digital presence is significant. Over 400million people are estimated to be owners of smartphones. WhatsApp recorded 487.5 million active users in India as of June 2021,[15] with Facebook recording 329.65 million profiles as of 2022.[16] Comparatively, Twitter has a fairly limited following in India

[15] Statista, 'Number of WhatsApp Users in Selected Countries Worldwide as of June 2021' (*Statista*, October 2021) <https://www.statista.com/statistics/289778/countries-with-the-most-facebook-users/> accessed 5 November 2022.

[16] Statista, 'Leading Countries Based on Facebook Audience Size as of January 2022' (*Statista*, January 2022) <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/> accessed 5 November 2022.

with approximately 23.6 million users in 2022.[17] A large amount of online campaigning utilizes the two former services for information collection and advertising. In the aftermath of the Cambridge Analytica scandal, multiple accusations were made by political parties in India against each other for having exploited Cambridge and similar services.[18] The ability to engage with voters in the online sphere has been shown to play a major role in a party's success.

The COVID-19 pandemic further compelled parties to grow their digital footprint due to restrictions on physical campaigning.[19] Parties which had invested in building their online infrastructure were in a position to take advantage of this. The Bihar election demonstrated the strength of the Bharatiya Janata Party (BJP) in this sphere, where high ranking party members were attuned to online campaigning, unlike their Mahagathbandhan ("MGB") opponents which included the Indian National Congress, Rashtriya Dal, and various Left Parties.[20] The MGB moved the ("EC") to restrict the amount that parties could spend on their online campaigns.[21] This demonstrated the MGB's apprehension regarding the BJP and Janata Dal's expertise in the digital sphere in terms of the resources they had available to harvest data on voters and indulge in micro-targeting.

However, the act of campaigning through microtargeting is merely the final stage of an elaborate process whereby data is harvested and utilised. The important work that is done in the background is obtaining the information regarding voters. The raw data allows analysts to make inferences/predictions regarding the biases and opinions of each voter. The ability to discern the most effective message to sway voters and then send such messages via social media is the object of this exercise. Shivam Shankar Singh, a data analyst who worked directly on a number of political campaigns, went

---

[17] Statista, 'Leading Countries Based on Number of Twitter Users as of January 2022' (Statista, January (2022) <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> accessed 5 November 2022.

[18] Yahoo News, 'Congress was Caught in Alliance with Cambridge Analytica, Facebook to Weaponise Data: RS Prasad', *Yahoo News* (New Delhi, 16 August 2020) <https://in.news.yahoo.com/congress-caught-alliance-cambridge-analytica-115007570.html> accessed 23 September 2021.

[19] Vijai Laxmi, 'Bihar Elections: Parties Go Full Throttle on Online Prachaar for Victory Amid Coronavirus Pandemic', *India TV News* (5 July 2020) <https://www.indiatvnews.com/politics/national-bihar-elections-2020-congress-bjp-online-campaign-ljp-rjd-631841> accessed 21 October 2021.

[20] Tushar Dhara, 'BJP's Social-Media Dominance is Shaping Mainstream Media Narratives Ahead of Bihar Polls', *Caravan Magazine* (Patna, 3 October 2020), <https://caravanmagazine.in/politics/bjps-social-media-dominance-is-shaping-mainstream-media-narratives-ahead-of-bihar-polls> accessed 21 September 2021.

[21] Amita Tagore (n 10).

as far as claiming that even "Cambridge Analytica…couldn't dream of this level of targeted advertising."[22]

There are multiple forms of information regarding people. Some may be considered intrinsic to people, such as their gender, sex, religion, and financial status. These kinds of information would have fallen under the definition of 'sensitive personal data' under the PDP Bill, 2019. However, even innocuous information which would not be considered 'personal data' under the 2019 Bill, are important data points for drawing inferences. In the context of both sensitive personal data and seemingly innocuous non-personal data, the PDP Bill, 2022, takes two significant steps backward. The former, sensitive personal data, no longer finds any place in the newly proposed Bill.[23] Further, Clause 2(13) of the new Bill is inadequate for addressing the dangers that even harvesting of non-personal data pose, given that it classifies personal data as only those kinds of data that can lead to identification of the data principal.

Singh elaborates on how data like electricity bills help determine the overall economic profile of different areas.[24] A household with high electricity bills would lead to an inference of high economic status with its attendant social attitudes and tastes. In this manner, evaluations can be made regarding the type of online advertising is most visible to such individuals, and what issues are of greatest importance to them. This is a stark example of exactly the kind of data that Clause 2(13) of the new PDP Bill, 2022, fails to engage with. Such forms of data, whether they can lead directly to identification of an individual or not, can still be sufficient to draw inferences about them and subject them to microtargeting.

In this manner, effective data collection teams can allocate individuals into different groups and infer the political preferences of each group. Singh uses the example of non-Yadav Other Backward Classes ('OBC') voters in Uttar Pradesh.[25] Through gathering both personal data like their age, sex, and education level, along with other innocuous bits of data such as electricity bills, a highly personalized profile can be created for each non-Yadav OBC. In this manner, a seemingly amorphous and diverse collection of

---

[22] Shivam Shankar Singh, *How to Win an Indian Election: What Political Parties Don't Want you to Know* (Penguin Books 2019) 76.

[23] Nivedita Krishna, 'Digital Personal Data Protection Bill 2022: How it has Left Both Civil Society and Industry Body Shell Shocked' *The Times of India* (29 November 2022) <https://timesofindia.indiatimes.com/blogs/niveditas-musings-on-tech-policy/digital-personal-data-protection-bill-2022-how-it-has-left-both-civil-society-and-industry-body-shell-shocked/> accessed on 3 December 2022.

[24] Shankar Singh (n 22) 64, 74.

[25] Shankar Singh (n 22) 75.

people can be allocated into a group for campaigning purposes.[26] A highly targeted message can be communicated to only these individuals, through social media and online advertising, regarding the issues that matter to them the most.

Apart from the fact that parties are able to advertise their messages this way, the other important aspect of online campaigning is that not every party can effectively use this resource. Singh elaborates on how the BJP used data collection and targeted campaigning in the Tripura Assembly elections of 2018.[27] He notes that the incumbent Communist Party of India (Marxist) ['CPI(M)'], did not have the resources to make the kind of promises the BJP was able to during the election campaign. More importantly, the CPI(M) did not possess the vast amounts of data regarding their constituents that the BJP did. This example demonstrates how even a relatively significant political party such as the CPI(M) lags behind in terms of adoption of technologically supported methods of campaigning. As significant and major parties become more attuned to the advantages of this approach, the importance of regulating their activities will increase.

The sources from which data was collected is also worth noting. The PDP Bill, 2019 created an exception to the notification requirement when the information is already public under Clause 14(2)(g)[28], which implies that such information may be harvested freely and without any accompanying disclosure or attainment of consent. An example of information that is already public is the rolls of the EC.[29] Further, the BJP developed mobile applications, independently or in conjunction with private parties, which required information such as name, sex, religion, and so on, from those who downloaded them.[30] Importantly, Singh notes that none of these actions undertaken by the BJP were illegal.[31] The lack of any legal regulation accentuates the need for a governing law that addresses these activities. Under the new 2022 Bill, various forms of data may not even qualify as personal data to begin with, given the truncated nature of Clause 2(13) of the Bill. Hence, the question of notification would likely not arise at all. As already noted above, data which does not cross the threshold of Clause 2(13) would be sufficient to create a profile of large swathes of people for targeted advertising.

---

[26] ibid; William A Gorton 'Manipulating Citizens: How Political Campaigns use of Behavioural Social Science Harms Democracy' (2016) 38(1) New Political Science 61.
[27] ibid.
[28] Personal Data Protection Bill 2019, cl 14(2)(g) ("Personal Data Bill").
[29] Pahwa (n 11).
[30] Shankar Singh (n 22) 127, 140-146.
[31] ibid.

III. Drawbacks of Data Analytics in Elections

Modern data analytics has made online political microtargeting into a form of behavioural advertising. Behavioural advertising, used in large part by commercial entities like Big Tech companies, tracks users' online activity to market specific ads.[32] This form of advertising has its own benefits and drawbacks. The question of balancing the two is difficult, however, considering the dynamic way in which microtargeting keeps evolving and because the new ways in which data is harvested from different sources opens up further possibilities of the kinds of targeted messages that could be sent to consumers.

There are certain advantage/benefits to microtargeting. It is often useful to mobilize a portion of the electorate which may not be politically active by sending them direct and targeted messages,[33] and is also advantageous for nascent and up-coming political parties, given it is a cheap and easy way to broadcast their message to compete with more established parties in the initial stages of its existence at the local level.[34] However, this requires sufficient technological prowess and efficiency, as well as a basic amount of data regarding the social and economic make-up of the constituency in question. As may be evident, when elections transition from the local level to a larger stage, the expenses related to undertaking this become more onerous. Hence, this potential benefit of microtargeting for smaller parties is, in any case, swiftly eroded. Regardless, of the possible benefits of microtargeting, the dangers of such online advertising have been demonstrated amply in recent times. The Cambridge Analytica scandal was the highest profile of these, but by no means the only one. The threats of microtargeting can be roughly allocated under two headings: a) manipulation of voters; and b) violations of privacy.

In the context of manipulation, parties can maximise the turnout of constituents who are in favour of their stand. Conversely, they can use ads to dissuade constituents who prefer their opponents through 'dark campaigning'. Dark campaigning is a form of campaigning that informs voters about the negative aspects of their opposing parties, rather than providing positive

---

[32] Joseph Turow, *The Daily You: How the New Advertising Industry is Defining Your Identity and Your Worth* (YUP 2011); Frederik Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Kluwer 2015).

[33] Holt Kristoffer and others, 'Age and the Effects of News Media Attention and Social Media Use on Political Interest and Participation: Do Social Media Function as Leveller?' (2013) European Journal of Communication 19, 19-20.

[34] European Parliamentary Research Service, *Social Media in Election Campaigning* (*Briefing*, 21 March 2014, <https://www.europarl.europa.eu/RegData/bibliotheque/briefing/2014/140709/LDM_BRI(2014)140709_REV1_EN.pdf> accessed 15 October 2021.

messages regarding policies.[35] This can also include disinformation or half-truths to distort the image of opposing parties. The Donald Trump campaign in the 2016 US Elections was accused of using this strategy among African Americans. Using targeted ads, the Trump campaign was alleged to have shown African Americans ads about Hillary Clinton where she refers to them as "super predators" and "serial sexual harassers".[36] This discouraged African American turnout at the elections. Even if these constituents would not have voted for Trump, the objective was to ensure they did not vote for Clinton either, thus reducing her vote share.

Parties could also purport to prioritise the singular issue that is most important to each individual voter. This may even be entirely contradictory issues.[37] Taking the example of tribal people in Tripura, a party may target them with messaging that claims the economic and social upliftment of them as being their most important issue. At the same time, it could target other portions of the population by stating that development of forest and tribal areas for industry is the focus of the party. Clearly, these two stances are in conflict, as using the forest for industrial purposes is against the interests of the tribal population. Considering that the ads are only being shown in a targeted manner, they are hidden from the constituency which has received the directly contradictory promise.

This also misleads the electorate regarding how important an issue is to a political party.[38] Microtargeting creates an illusion that a particular party is completely devoted to a specific issue because that class of voters receives information and advertising that is targeted. These multiple promises to multiple people create a dissonance for both the electorate and the parties themselves. As different pledges have been made to different groups, a party might then struggle to determine which issue is of greatest importance.[39] Additionally, even though the marketplace of ideas benefits in some ways from this form of campaigning, it also creates significant fragmentation in the public conversation. The marketplace becomes multiple markets where

---

[35] Gorton (n 26).

[36] Joshua Green & Sasha Issenberg, 'Inside the Trump Bunker, With Days to Go' *Bloomberg* (27 October 2016) <www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go> accessed 18 October 2021); McKenzie Funk, 'The Secret Agenda of a Facebook Quiz' *New York Times* (New York, 19 November 2016) <www.nytimes.com/2016/11/20/opinion/the-secret-agenda-of-a-facebook-quiz.html?_r=0> accessed 21 October 2021.

[37] Borgesius (n 32) 87.

[38] D Sunshine Hillygus and Todd Shields, *The Persuadable Voter: Wedge Issues in Presidential Campaigns* (YUP 2008) 14.

[39] Borgesius (n 32) 88.

individuals do not visit other markets as the information that matters to them is provided directly to them through microtargeting.[40]

In terms of the privacy violations that occur as a result of this, the actual collection of data has already been elaborated upon.[41] But the other primary issue is the passing on of data to third parties by an otherwise reputed fiduciary. The argument used by most individuals regarding not bothering with their privacy in the context of Facebook and other such companies is based on trust. People believe that giving their personal data to reputed companies is acceptable as it will not be misused. However, the indirect collection of data by other malicious third parties like Cambridge Analytica or others, is a risk that should be given greater consideration. This is especially a risk when multiple private parties are entrusted with doing different aspects of data collection.[42]

Microtargeting can, therefore, expose the political process to both good and bad. As a result, its regulation cannot merely be confined to data privacy law, though it must be of primary importance. Election laws must also take cognizance of data privacy issues and work in tandem with privacy regulations to ensure the greatest degree of protection. The next part of the paper will look into election and data privacy laws across the world, before commenting on how the Representation of Peoples Act, 1951 can incorporate privacy concerns into its ambit.

Part II

IV. Data Privacy in Elections and Indian Election Laws

The dangers related to lack of regulation of data mining in the context of political campaigns and elections has witnessed increased recognition.[43] This new found awareness had led to an acknowledgment that there must be specific guidelines in place for political parties rather than reliance on general data privacy law. As Clause 50 of the PDP Bill, 2019, allowed the DPA to frame such guidelines for different industries, this responsibility must be taken up, as will be elaborated upon later. Under the 2022 Bill, however, the power to frame such guidelines seemingly does not vest with the DPA

---

[40] ibid.
[41] Neil Richards, *Intellectual Privacy: Rethinking Civil Liberties in the Digital Age* (OUP 2015).
[42] Borgesius (n 32) 87.
[43] Normann Witzleb (n 13).

any more. Clause 20, which lists the powers of the DPA does not include the framing of guidelines[44], and instead, grants the discretion to issue Rules under the Bill to the Central Government under Clause 26.[45] Further, data privacy laws must be synthesized with election laws to ensure they complement each other in addressing the specific challenges that arise. The responsibility to protect data and electoral integrity must be apportioned between both the EC and the DPA.

The EC under Section 123(2) of the Representation of Peoples Act, 1951, permits restriction of corrupt practices, which includes multiple forms of speech. While there is, obviously, no jurisprudence whatsoever on the point of whether data mining can be considered a corrupt practice, a case could be made for its inclusion. The Supreme Court had termed "undue influence" as being anything which reduces the free will of the electorate to the extent that the individuals have almost no choice.[46] However, the standard has subsequently been expanded wherein covering up important details such as criminal antecedents of a candidate amount to unduly influencing constituents.[47] Importantly, the Supreme Court has also maintained that the threat of violence is not necessary for Section 123(2) to be violated.[48]

On this basis, microtargeting could be considered a form of undue influence. The most effective usage of this provision would be to tally it with data privacy. Thus, a breach of data privacy laws would create a presumption that there had been violations of privacy which give rise to undue influence. This is because of the difficulty associated with conclusively proving that microtargeting has had a particular threshold level of influence on constituents. If such an inference is drawn, the determination of undue influence would not be a purely subjective determination but rather be premised on whether the party has adhered to data privacy requirements. Thus, breaches of data privacy law, and the practice of microtargeting using this data, may cumulatively create a presumption that the practice of a party has been acted in a manner that has unduly influenced the voters in an upcoming election. The fact that the Supreme Court departed from the "no free choice" standard and classified certain actions or omissions as automatically amounting to undue influence, allows for this. This would form a strong deterrent against parties attempting to skirt data privacy laws.

---

[44] Personal Data Bill 2022, cl 20.
[45] ibid, cl 26.
[46] *Shiv Kirpal Singh v V.V. Giri* (1970) 2 SCC 567.
[47] *Krishnamoorthy v Sivakumar* (2015) 3 SCC 467.
[48] ibid.

What is more important, however, is to now examine the PDP Bills, both 2019 and 2022, and its relevance to elections. While election laws may play an important role at deterring microtargeting, the stage at which the manipulation of voters begins is via mining and usage of data to profile constituents. The preliminary conclusion that is evident from this, as will be elaborated upon subsequently, is that no single statute or authority will be able to adequately address the entire spectrum of issues that arise in the context of data collection and microtargeting. This paper will now turn to the analysis of this dilemma by starting with an examination of the provisions of both PDP Bills.

V. Details of the Data Protection Bill

The withdrawn PDP Bill, 2019, was based largely from the most comprehensive data privacy legislation in the world at present, the General Data Protection Regulation ("GDPR"). Thus, several provisions mirror the contents of the GDPR.[49] For example the definition of "Personal Data" provided under Clause 3(28) of the PDP Bill contains the same wording as the GDPR equivalent, with one important distinction.[50] The PDP Bill includes "inferences" under the ambit of personal information, unlike the GDPR. This is a lacuna in the GDPR which has been commented on negatively by scholars.[51] The importance of inferences is that the raw data which is accumulated is often of less importance than the inference derived from it as a result. On the face of it, classifying inferences as personal data is valuable, as it provided individuals with the full ambit of rights associated with personal data under the PDP Bill, 2019. However, this lacuna which had correctly been addressed in the 2019 Bill has now been removed in the 2022 Bill. Rather than a helpful step, this is a significant regressive move in the development of the Bill.

Clause 3(13) of the PDP Bill, 2019 defined Data Fiduciaries as individuals, the state, or juristic entities (which includes political parties) that accumulate personal data.[52] Chapter II of the 2019 Bill outlined the obligations of Data Fiduciaries. Clause 4 restricts the collection of personal data except for

---

[49] Anirudh Burman, 'Will a GDPR-Style Data Protection Law Work for India' (*Carnegie India* 15 May 2019) <https://carnegieindia.org/2019/05/15/will-gdpr-style-data-protection-law-work-for-india-pub-79113> accessed 25 September 2021.

[50] Personal Data Protection Bill, cl 3(28).

[51] Sooraj Shah, 'This Lawyer Believes GDPR is Failing to Protect You – Here's What She Would Change' *Forbes* (30 January 2019) <https://www.forbes.com/sites/soorajshah/2019/01/30/this-lawyer-believes-gdpr-is-failing-to-protect-you-heres-what-she-would-change/> accessed 24 September 2021.

[52] Personal Data Bill cl 3(13).

lawful purposes.[53] While there are no laws which restrict political parties from collecting personal data, they would still abide by the restrictions under the PDP Bill, alongside respecting the rights of Data Principals. Clause 5 restricts the usage of collected data for stated purposes. However, Clause 5(b) truncates this protection by allowing the data to also be used for which "…is incidental to or connected with such purpose, and which the data principal would reasonably expect…".[54] The party which is in power may have access to far more data than its competitors, given that it processes data for various purposes related to administration of government. Clauses such as 5(b), and Clause 12 which grants the ability to process data without consent, could be misused by a ruling party to gain an advantage.

In the 2022 Bill, there are various troubling provisions in respect of how Data Fiduciaries can now collect data. Clause 8 which deals with "deemed consent" forgoes the consent of the data principal entirely.[55] While not all the sub-clauses are unusual or potentially harmful, there exists the possibility of abuse in respect to certain instances of deemed consent, such as with sub-clause (2).[56] Further, instead of an exemption from notification, Clause 8(8)(f) now merely presumes consent in the case of publicly available data about an individual, such as information from electoral rolls that parties can easily access.[57] Further, Clause 8(9) outlines a vague notion of deemed consent for any "fair and reasonable" purpose as may be prescribed, after consideration of certain conditions within the sub-clause.[58] All these provisions, in the background of the overall watering down of the definition of "personal data" under Clause 2(13), provide significant leeway for collection of data from which to draw inferences.

With regard to the DPA which was constituted under Clause 41(1) of the PDP Bill, 2019it is entrusted with carrying out a number of functions.[59] The most important of these functions in the context of elections is the power to lay down codes of best practice for different industries under Clause 50.[60] As already mentioned, this power seemingly no longer vests with the DPA at all. Regardless, even if the 2019 Bill was taken as the benchmark, one preliminary issue associated with the DPA's functions is that it seems to clearly be aimed toward commercial entities. The penalties under Chapter X seem to

[53] ibid, cl 4.
[54] ibid, cl 5(b).
[55] Personal Data Bill, 2022, cl 8.
[56] ibid, cl 8(2).
[57] ibid, cl 8(8)(f).
[58] ibid, cl 8(9).
[59] PDP Bill 2019 cl 41(1).
[60] ibid, cl 50.

be aimed at the turnover of commercial entities, presumably, because that is the primary objective of such Data Fiduciaries.[61] While, as mentioned already, the definition of a Fiduciary could prima facie apply to parties, the reliefs provided seem to be oriented more toward commercial entities.

The 2022 Bill under Clause 19 refers to setting up the DPA.[62] While the Chapter on criminal penalties has been excluded in the 2022 Bill, the civil penalties provided under Clause 25 read with Schedule 1 remain seemingly focused on commercial entities.[63] Hence, this overall approach of the Bills, no matter which version, appears to remain fixated with the dangers of collection and processing of data by commercial entities, but not by political actors. This facet appears to have been overlooked.

Clause 7 of the 2019 Bill outlined the obligation of Data Fiduciaries to inform the Data Principals about the collection of their data, along with other forms of information collected, in the notification, that must be issued as mandated by the Clause. However, a caveat in this provision exists in the form of data collected from a source other than the Data Principal itself. If the data is accrued from a third-party source, the notification needs to only be done "…as soon as reasonably practicable…".[64] There is no indication of what this "reasonable" period might be. Also of note, is Clause 7(b) requires only the "nature and categories" of data be notified to the Principals, and not necessarily the exact content of said data.[65] Hence, a category of data may include credit or financial information regarding the Data Principal, but would not necessarily include what kind of information in this category, such as outstanding loans or scheduled payments, have been harvested The processing of personal data can be done without consent for reasons provided under Chapter III, such as under Clause 12 which exempts notification for effectuating laws or orders of a Court[66], or under Clause 13 which provides the same exemption when it comes to information required for employment purposes.[67] In Clause 14(2)(g), data which is publicly available may be collected and processed without seeking permission from the Data Principal.[68]

The 2022 Bill refers to roughly the same obligations that existed for Data Fiduciaries under the 2019 Bill. Chapter 2, from Clauses 5 to 11, outlines the

---

61 ibid, ch X.
62 Personal Data Bill 2022, cl 19.
63 ibid, cl 25; Sch 1.
64 PDP Bill 2019 cl 7.
65 ibid, cl 7(b).
66 ibid, cl 12.
67 ibid, cl 13.
68 ibid, cl 14(2)(g).

various requirements that Data Fiduciaries and Significant Data Fiduciaries must adhere to while carrying out their activities.[69] Within these provisions, various implicit exceptions to the requirement of consent from Data Principals have been incorporated. Clause 9(9) of the 2022 Bill allows Data Fiduciaries to transmit data to each other when it has been obtained via consent from the Data Principal already.[70] The notice requirements have become somewhat vaguer than they had been under the first iteration of the Bill in 2019. Clause 6(1) refers to providing an "itemised list" of data that is sought to be obtained from the Data Principal, without any indication as to how specific these "items" need to be.[71] The discretion that appears to be vested in a Data Fiduciary may be misused.

The rights of the Data Principals are located under Chapter V. These rights include the Right to Access data collected by the Fiduciary[72], the Right to Correction[73] and the Right to be Forgotten.[74] Clause 21 directs the Data Principal to request a Data Fiduciary to comply with any of the rights provided under Chapter V, in case they find that there is a breach of the rights provided. However, Clause 21(4) allows a Data Fiduciary to refuse compliance, with reasons provided in writing. The Data Principal may then appeal to the DPA.[75] Several of the same rights are transposed onto the PDP Bill, 2022, however, certain anomalous additions have also been made in Clause 16, which deals with duties of Data Principals. These inclusions abrogate the rights under Chapter 3. Specifically, Clause 16(2) states that a Data Principal "shall not register a false or frivolous claim" against a Data Fiduciary.[76] It is entirely unclear what counts as a "false or frivolous" claim, and the need to include such wording in the Bill itself is questionable. Undoubtedly, had the claim been false or frivolous, it would have been dismissed via the judicial process. Instead, the inclusion of an explicit duty under Clause 16 is likely to have a chilling effect on Data Principals exercising their data rights. In the alternative, it could even lead to counter claims made by the Data Fiduciary against the Data Principal on the ground that the objections raised by the latter, come under the category of "false and frivolous".

There are other exemptions which are of importance in the context of data collection by political parties or their agents. An example of such an

---

[69] Personal Data Bill 2022, ch 2.
[70] ibid, cl 9(9).
[71] ibid, cl 6(1).
[72] PDP Bill 2019, cl 17.
[73] ibid, cl 18.
[74] ibid, cl 20.
[75] ibid, cl 21(4).
[76] Personal Data Bill 2022, cl 16(2).

exemption that could be exploited is Clause 38 of the PDP Bill, 2019, which allowed processing of Personal Data for statistical and/or research purposes. An entity may process this data, claiming that it falls under Clause 38, but then use the research that it does to engage in micro-targeting.[77] This problem is accentuated by the fact that there is no requirement for these actors to specify what type of statistical work or research they are undertaking. A similar provision, Clause 18(2)(b), has been retained in the 2022 Bill.[78] The Central Government also has power to exempt Data Fiduciaries from the entire ambit of Chapter 2 of the PDP Bill, 2022, based on "volume and nature of personal data processed". What this implies is not clarified and the threshold in terms of "volume" and genus of personal data in terms of "nature", which could lead to such wide-ranging exemptions, is not even hinted at. Many of these provisions, both in the 2019 and 2022 versions of the PDP Bill, are relevant for the question that will be addressed now which is how effectively each of the Bills deals with inferences. It will be argued that even under the more favourable regime of the 2019 Bill, categorizing inferences as personal data does little to address the specific issues that such inferences pose for data privacy in an electoral context.

VI. Inferences, Elections, and the Pdp Bill

The regulations of inferences have increasingly been seen as crucial for ensuring adequate protection of data. Inferences as the subject of data protection presents a unique set of issues which are different to other types of data. The reasons for this, and the consequences that this has for guaranteeing data rights, will now be elaborated upon, in the context of multiple provisions of the PDP Bill, 2019, as it was. Considering the similarities between the GDPR and PDP Bill, 2019, the jurisprudence and experiences under the former will be used to analyse the potential problems that may arise in India. This will also be helpful, to an extent, in the context of the 2022 Bill given some of its provisions remain similar to those that existed under the 2019 Bill.

Status of Inferences under the GDPR and PDP Bill

One of the acknowledged flaws in the GDPR was its failure to include inferences under the ambit of personal data.[79] In Europe, there has been an

---

[77] PDP Bill 2019, cl 38.

[78] Personal Data Bill 2022, cl 18(2)(b).

[79] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' 20192 Columbia Business Law Review 494; Article 29 Data Protection Working Party, Opinion 4/2007 on the Concept of Personal Data, 01248/07/EN WP136 (20 June 2007).

increasing recognition in legal policy circles that inferences should also be addressed in the GDPR. For example, the Article 29 Working Party, an advisory body comprised of representatives from the data protection authorities of each EU Member State, the European Data Protection Supervisor, and the European Commission, has noted that inferential analytics, the process by which inferences are drawn from data, do the actual harm to individuals in the context of privacy breaches.[80] The focus of data privacy law is usually at the stage of inputs or when the data is collected. Of equal importance, is the output generated as a result of those inputs. These outputs are the basis for actions taken in the real world regarding that person. The outputs, which are inferences, require greater regulation under the law.[81]

The PDP Bill, 2019, prima facie addressed this concern by including inferences under personal data in Clause 3(28). This is similar to the California Consumer Privacy Act, which also classifies inferences as part of "personal information" under Title 1798.140, and is then accorded protection of various forms under the Act.[82] However, in the context of the PDP Bill, 2019, in actuality, the inclusion of inferences under personal data in Clause 3(28) did little to address the issues associated with inferences. The difference between inferences and other data points is the subjectiveness and non-verifiable nature of the former.[83] What this means is that inferences, although based on objective facts about a Data Principal, are subjective to the extent that an evaluator draws their own subjective conclusions based on those facts. Inferences, by their nature, are derived from objective facts or data that is accumulated about the Data Principal. This inference is, essentially, an opinion that must be subjective and based on the evaluator's personal metrics of judgment.[84] Such inferences are non-verifiable in nature given this subjectivity. This presents problems in terms of the ambit of rights that can be enforced regarding non-verifiable inferences.[85] Such non-verifiable inferences are inherently based on the judgment of the individual or entity making the evaluation of the Data Principal. It is not possible to test such subjective inferences or establish their "correctness" given that the discretion of the evaluator is built into the inference and final decision that is made.

The act of making an inference from raw data presents two separate issues for data privacy law. The first is the metric by which the inference is made.

---

80 Art 29 Data Protection Working Party (n 79).
81 Wachter and Mittelstadt (n 79) 4-6.
82 Title 1798.40, Title 1.81.5 California Consumer Privacy Act of 2018.
83 Art 29 Data Protection Working Party (n 79).
84 ibid.
85 ibid 8.

The process, criteria, or algorithm in the case of automated decision making, is the metric used by the Data Fiduciary to make an evaluation. This process is what leads to the conclusion. For example, banks use multiple criteria to determine to determine if an individual has a satisfactory credit score and is eligible to receive a loan.[86] The raw data about the applicant is put into the system to reach the final inference.

This metric has an important effect on the final inference. The Article 29 Working Group has advocated the inclusion of both inferences and the underlying metric used to reach the inference under the ambit of "personal data" under the GDPR. European case law shows that this interpretation has been partially taken into account. Even though the GDPR does not categorize inferences as personal data[87], the ECJ has included it regardless. It noted that the assessments or evaluations of individual will lead to a decision being made that affects him/her. Therefore, it is appropriate to include this under "personal data" and afford it the protections in the GDPR.[88]

However, the second issue that arises is that for an inference to be considered personal data, it must be a "verifiable" inference. While India had already crossed the initial threshold of defining inferences as personal data by including it in the PDP Bill, 2019, the verifiability question has further implications. An inference that is based primarily on a subjective metric or criteria would be difficult to "correct" under Clause 18.[89] A subjective opinion regarding a person's credit score, for example, cannot logically be open to "correction". Ultimately, it is within the bank's own subjective determination whether an applicant has demonstrated reliability in terms of paying back a prospective loan.

Thus, there is essentially no difference between the position in Europe and under the PDP Bill, 2019. The 2019 Bill included inferences under personal data, and the ECJ has also extended the definition of personal data to include inferences, depending on verifiability of the inference. While verifiability did not matter under the 2019 Bill, other attendant issues with inferences as personal data will still apply, such as accuracy of the decision-making process and the fact that certain rights, such as the aforementioned Right to Correction under Clause 18, may not even be available for inferences. These issues will now be expanded upon.

---

[86] ibid.

[87] Joined Cases C–141 and 372/12 *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I- 2081 para 40 (Cases preceding the GDPR also follow this principle).

[88] Case C-434/16 *Peter Nowak v Data Protection Commissioner* [2017] ECR I-994 para 60.

[89] ibid 34, 42-44.

In this context, while there were improvements that may have been brought to the treatment of inferences under the 2019 Bill, the fact that they were explicitly mentioned as "personal data" was a positive step. This entire discussion is now negated by the latest 2022 Bill, which ignores inferences entirely. The removal of inference from the umbrella of "personal data" constitutes a regrettable omission and the 2019 Bill remains a better starting point on this aspect than the latest version which is being considered.

Application of Data Principals' Data Rights to Inferences

The reasons or the criteria on which a subjective inference is drawn regarding a Data Principal is also not included as personal data under the PDP Bill, 2019, nor under the 2022 Bill given its shrunken definition of personal data. After objective facts and data are collected about a Data Principal, there will often be a set of guidelines or criteria which will be applied to those facts, and which will give rise to final inference made about the Data Principal. Under the GDPR, there have been arguments made to include these criteria/guidelines under the ambit of "Personal Data". However, there are indications from the interpretation of the GDPR that it may be left out.[90] For this exercise, we may take the assistance of how the GDPR, in the same context, has been interpreted. This is, evidently, because the GDPR and PDP Bill, 2019 had significant similarities and there is no jurisprudence under the latter. While this does not, of course, mean the PDP Bills, 2019 or 2022, would have been interpreted by Indian Courts in the same manner, this similarity may assist in providing us with important indicators in this regard. Further, as already mentioned, the lack of any interpretation of the provisions of the Bill, due to it having not come into force, would make any consideration of alternatives purely speculative. For carrying out this comparison and estimation of how data privacy rights may develop, the provisions of the 2019 Bill will be utilized.

The ECJ while interpreting the GDPR has made clear that, in its view, the purpose of data privacy law is not to provide transparency in the decision-making process by which an inference is made.[91] The end result i.e., the inference may be personal data under certain circumstances but not the rationale or analysis behind the inference. The drawback of this is that it removes an avenue of inquiry for a Data Principal. In the context of an unverifiable or subjective inference, one of the means of challenging it or seeking a correction could have been that the criteria of evaluation is flawed.

---

[90] ibid 57.

[91] Joined Cases C-141/12 & 372/12 (n 86) paras 45–47; Case C-28/08 *P European Comm'n v Bavarian Lager* [2010] ECR I-06055 para 49.

A similar problem is faced in the context of the Right to Access under Clause 17 of the PDP Bill, 2019. If the ECJ's interpretation of the ambit of personal data is taken as a proxy for what it could be under the PDP Bill, the process/criteria for making the inference cannot be accessed by the Data Principal. The Article 29 Working Group has disagreed with this approach and included the method of evaluation under the ambit of the GDPR.

What is evident from this is that the rights under the PDP Bill, 2019, would not have been applied equally across all scenarios. Wachter & Mittelstadt note that the *telos* (end term of a goal-directed process) of different spheres of activity will shape the way data privacy rights are applied.[92] In certain scenarios, such as the bank's credit score, certain rights may not be available at all, such as the Right of Correction.[93] This is because, in the context of giving an individual a credit score, the individual is asking to be evaluated. Therefore, the individual would not have the right to correct that evaluation, unless it is shown that there was a mistake in recording the data points or inputs. This also falls in the category of a "non-verifiable" inference, as already alluded to earlier.

In regard to a non-verifiable inference, the 2022 Bill takes yet another turn for the worse. Clause 16(4), which lists out the mandatory duties of a Data Principal, explicitly precludes a Principal from even asking for Correction or Erasure of their data under Clause 13, unless the data is "verifiably authentic".[94] The ambit of this phrase is nebulous and leaves much to imagination. A "non-verifiable" inference in this context, may arguably have been excluded entirely from the scope of the rights of a Data Principal under Clause 13 of the PDP Bill, 2022. As the analysis earlier shows, this is particularly dangerous in the context of inferences given that many are, by nature, "unverifiable". However, under the PDP Bill, 2019, there had not been any explicit bar on seeking a correction or erasure of an inference, as difficult as actually doing so may have been in practical terms. This was further enabled by the fact that inferences were recognized out rightly as Personal Data under Claus 3(28). However, even that limited scope for interference has been completely closed off by the PDP Bill, 2022.

Regardless, even if we were to take the 2019 Bill as the benchmark, certain issues would persist. When looked at in this teleological context, the way by which rights under the PDP Bill, 2019, may have been applied to

---

[92] *Peter Nowak v Data Prot. Comm'r* [2017] ECR I-994, Case C-434/16, Opinion of Advocate General Kokott, paras 35, 53.

[93] Wachter and Mittelstadt (n 79).

[94] Personal Data Bill 2022, cl 16(4).

micro-targeting and election campaigning would have been curbed. To begin with, the Right to Correction may not be available as parties are making their own determination of what they need to say to a particular electorate. Looking back to the example of non-Yadav OBC's, the party in question determines what is most effective to build support for itself. Considering this touches on the freedom of a party to conduct its election campaign and decide strategy for itself, a Data Principal may not have the Right to Correct an inference made about himself.[95]

Further, the inferences made may be unflattering or embarrassing to a Data Principal, but may still in fact be the best way to win their vote. A Data Principal may consider himself to be liberal or left-leaning, but his data may indicate policy leanings which are more associated with conservative positions. Positions taken by any individual are complex and often overlap onto both sides of the political spectrum in terms of both Right- and Left-Wing parties. The inferential analytics of a political party may lead to the conclusion that appealing to the individual's conservative positions is more effective. This might contradict the way the Data Principal self-identifies but self-identification often depends on social pressure and peer groups. For getting the Principal to vote for it, a party is unconcerned with such technicalities.

This goes back to the problem with inferences and the reason why a simplistic inclusion of it under Clause 3(28) would not have solved the problem. The PDP Bill, 2019 was ill-suited to determine whether the inference reached about a Data Principal is accurate or not. One cannot use the standard of what the Data Principal itself determines to be accurate. The consequence of this would be that all individuals would always flag any unflattering inference about them as "inaccurate", and demand its correction. In the context of powerful individuals in society, this problem is especially pronounced. It would be unbecoming of data privacy law to allow individuals to alter all negative inferences made about them.

Inferences and Rights of the Data Fiduciary

The Data Fiduciary may have itself have rights which conflict with the rights provided under the PDP Bills, both 2019 and 2022. The method or criteria of assessment may be part of the Data Fiduciary's own right to privacy as the metrics may be unique. In different contexts, the privacy rights of companies and juristic entities has been acknowledged in India, though the full ambit

[95] Opinion of Advocate General Kokott (n 92) para 56.

of rights under *Puttaswamy* do not seem to have been extended to them.[96] The Data Fiduciary, understandably, would not want to reveal the criteria which was used to make the decision as it would then be open to scrutiny by competitors. Another possible restriction may be other laws in the country itself. It has been increasingly understood that information and data form a commodity in and of itself. A private company which is entrusted with doing the collection for a party, may use Intellectual Property Laws to protect against needing to make disclosures.[97] Trade Secrets under Intellectual Property could be a ground taken by a private company to claim that revealing data would prejudice their competitive position vis-à-vis other companies offering the same services.[98] While data has not been explicitly recognized as a commodity, capable of protection under Intellectual property or competition law, other jurisdictions have recognized this possibility and included data protection under Trade Secrets within the ambit of IP.[99]

This applies in an electoral context, and not just in a commercial setting between two business competitors. The BJP's methods of categorizing different people into groups which is the basis of inferences made about them, may be of great importance for campaigning. The competitive advantage of any party in this sphere would, undoubtedly, be something they wish to maintain. The Right of Correction is similarly impacted. It can be argued that a party has a right to make its own evaluation regarding the tastes and preferences of voters. This could especially be the case if, as the ECJ stated, the purpose of data protection law is not to evaluate the "correctness" of decisions except in limited scenarios.

Thus, the balancing of the rights of Data Principals with the rights of Data Fiduciaries is necessary. The former is detailed in both the PDP Bills,

---

[96] Lomesh Nidumuri and Tejas Shetty, 'India: Right to Privacy of Companies Vis-à-Vis the Powers of the Central Government under Section 206(5) of the Companies Act, 2013 – Has the Balance Been Lost?' (*Mondaq*, 15 May 2020) <https://www.mondaq.com/india/privilege/934460/right-to-privacy-of-companies-vis-a-vis-the-powers-of-the-central-government-under-section-2065-of-the-companies-act-2013--has-the-balance-been-lost-#:~:text=INDIAN%20LAW%20ON%20THE%20INTER,AND%20DOCUMENTS%20OF%20A%20COMPANY&text=The%20recent%20judgment%20of%20the,of%20the%20Constitution%20of%20India> accessed 7 October 2021.

[97] Wachter and Mittelstadt (n 79) 55.

[98] *John Richard Brady v Chemical Process Equipments (P) Ltd* 1987 SCC OnLine Del 236: AIR 1987 Del 372; *Ambiance India (P) Ltd v Naveen Jain* 2005 SCC OnLine Del 367: (2005) 122 DLT 421.

[99] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' [2016(6)] Int'l Data Privacy L. 102, 115; 'Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the Protection of Undisclosed Know-How and Business Information (Trade Secrets) Against Their Unlawful Acquisition, Use and Disclosure', 2016 OJ (L 157) 1.

but the possible defences that can be raised by the latter seem to have been overlooked in the 2019 and 2022 frameworks. Undoubtedly, such conflict between rights would inevitably arise as the law develops and would likely have to be dealt with on a case-to-case basis. The jurisdiction of the DPA to determine whether an inference by a political party is legitimate or not is also unclear.

Considering the DPA is not conditioned for this purpose, its own *telos* is not shaped to be able to address the specificities of election advertising.[100] Therefore, a clear demarcation between the roles to be played by both the DPA and the EC will need to be created, as will be elaborated upon further. At present, it is sufficient to say that the DPA should, in principle, examine the entire process of data collection and generation of inferences through which categorization of voters takes place, while the EC should evaluate the exact effects of microtargeting itself and whether it amounts to a breach of Section 123(2) of the RPA. In the latter evaluation, the findings of the DPA regarding legality of the data collection, should be a relevant factor.

Redundancy of the Notification Requirement for Inferences

The PDP Bills largely try to protect data by providing transparency regarding its processing. The notification requirements under the respective Clauses of the 2019 and 2022 Bills, as already detailed before, strengthen the ability of Data Principals to track what kind of data is being collected from them. However, in the context of inferences, these provisions are lacking. For instance, Clauses 7 and 8 of the 2019 Bill cannot assist with Data Principals being aware of inferences about them, as the notification requirements are for personal data that is directly taken from them. By their very nature, inferences are not taken from the Data Principals but are derived from the data which is collected. Thus, the raw data collected is subject to the requirement of explicit consent from the Data Principal. However, there is little incentive to inform Data Principals about the inferences based on this data. This problem persists in the 2022 Bill as well, over and above the various additional problems in it which had not existed under the 2019 Bill.

Clause 7 of the 2019 Bill contains an additional problem in that it allows for data to be used for purposes that are reasonably related to the purpose that the Data Principal consents to. Even if it could have been argued that

[100] For further reading, Helen Nissenbaum, 'Privacy as Contextual Integrity' (2004) 79 Wash LRev 119; Jeroen Van Den Hoven, *Information Technology, Privacy and the Protection of Personal Data* (CUP 2008).

the consent requirement mandated a clear indication regarding the possible inference to be made from the data, this is undermined by the words "reasonably related". A Data Fiduciary may simply state that the inference falls within this safe harbour without providing details about its specifics. They would not have to include this under the notice of consent at all. There is no direct mirror provision to this in the PDP Bill, 2022, which on the face of it is one of the few aspects on which the newer Bill improves on the 2019 iteration.

Alternatively, it could be that a separate entity from the political party processes the data and creates the inference.[101] Singh notes that multiple private parties assisted the BJP in their data accumulation exercise.[102] Thus, if the inferences are made by this entity and transferred to the BJP, a notification would have to be provided. Even this requirement was hollowed out by Clause 7 of the 2019 Bill stating that the transferee only needs to notify the Data Principal "as soon as reasonably practicable". There is no indication of what this entails and could be abused by the Data Fiduciary to delay revealing the inference that it received from the third party. This is particularly problematic, given the corresponding provision in the GDPR, Clause 14, mandates that the notification be provided within one month. Thus, a clear intention can be discerned in the shelved PDP Bill, 2019, to make the notification requirement more lenient. This aspect has been left unaddressed completely in the 2022 Bill. Whether a notice is required to be provided to the Data Principal when their data is transferred by a Data Fiduciary to a third-party entity, is not answered. Clause 9(9) allows for such transfers to take place provided a valid contract exists between two entities, but does not mention notice being sent to the Data Principal. This leaves open the possibility that the data of Principals may be shipped around from one party to another, after the initial grant of consent to a specific Data Fiduciary to process personal data.

Finally, the notification requirement in the 2019 Bill only mandated that categories of data be provided, and not the actual data itself.[103] This creates an unnecessary layer of confusion for the Data Principal. The category does little to inform Principals about whether the data that has been collected is potentially harmful or particularly invasive of their privacy. Another question

---

[101] Wachter and Mittelstadt (n 79) 52.

[102] Shankar Singh (n 22) 140-150.

[103] Similar to the position in the EU, *EU General Data Protection Regulation (GDPR)*: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1, art 13.

that is left unaddressed is whether the notification is necessary if the original Data Fiduciary had informed the Data Principals of the possible sharing of their data with a list of third parties. Substantively, the Data Fiduciaries could argue that the pre-emptive notification is sufficient, in substance, to satisfy the requirements of Clauses 7 and 8. The PDP Bill, 2022, makes the notification requirement even less definitive, by merely requiring an "itemised" list of data that is collected to be provided to the Data Principal. How detailed, or broad, this list can be is not mentioned in the Bill.

The lacunae in the notification requirements mean that Data Principals cannot rely on it to remain up to date regarding inferences. The possible solution to this is the Right to Access. Principals may seek information regarding the inferences made by the Data Fiduciary at any point in time. However, this right also cannot be deemed to be absolute. The Right to Access provides the option to the Data Fiduciary to restrict the disclosure to merely the categories of data that are collected. This gives significant leeway to the Data Fiduciary to keep the crux of the data collection a secret.

Furthermore, as already alluded to above, the Right to Access can become hollow in several other scenarios.[104] Data Fiduciaries may raise their own rights under other laws as a defence against complying with a data request by a Data Principal. This reduces the oversight and transparency that is possible for inferences. The Right to Access the data of Data Principals by the Principals themselves is primarily meant to provide this transparency, in conjunction with consent requirements and notification. However, as described above, there are multiple ways to keep this information opaque due to loopholes in the PDP Bills, both 2019 and 2022, as well as potential rights that Data Fiduciaries may have that guard against a need for full disclosure.

In the electoral context, this is problematic given the sources of information for parties may not be apparent. Further, the most important piece of data which is the inference made about a voter will be difficult to identify. Having access to this information would make a Data Principal more aware of the kind of targeting he/she might be exposed to. Political messaging and advertising might be easier to identify if the individual knows that he/she is being targeted in a certain way. Without this transparency, it is possible for Data Principals to be implicitly influenced without even realising that it is occurring.

Further, the specific provision to object to decisions made via automated or algorithmic processing in the GDPR had not been reproduced in the PDP

[104] Text to n 92.

Bill, 2019, and predictably finds no place in the 2022 Bill given its significantly reduced scope. Thus, keeping in mind these issues, the inclusion of inferences under the definition of "personal data" had been made somewhat redundant even within the 2019 Bill by the many barriers to the exercise of the rights provided for Data Principals.

Sensitive Personal Data

In the 2019 iteration of the PDP Bill, the category of "sensitive personal data" does not include inferences under the PDP Bill and the standards provided under Clause 15 for the inclusion of new forms of sensitive personal data would have meant that their subsequent incorporation would have been unlikely. Sensitive personal data had been subject to certain additional protections under the PDP Bill, 2019 such as being exempt from processing under Clause 13, and a need for Significant Data Fiduciaries to undertake Data Protection Impact Assessments under Clause 27 of the 2019 Bill which mandates the Assessment to happen when using data that has a risk of harm to the Data Principal.[105] Sensitive personal data is no longer protected under the 2022 Bill, but the requirement for Significant Data Fiduciaries to conduct Data Protection Impact Assessments has been retained under Clause 11(2).[106]

Under the 2019 Bill, it would have been very difficult to include data under this category given the high procedural and substantive barriers to doing so under Clause 15. Wachter & Mittelstadt have designated a category of inferences as "high-risk inferences".[107] This refers to inferences which are accumulated through data collection, and which are the basis for decisions made regarding the Data Principal. In their definition, such inferences include the following characteristics: a) privacy invasive; b) damaging to reputation; c) have low verifiability.[108]

While the GDPR contains provisions for protecting sensitive personal data under Clause 9, the PDP Bill, 2019, contained a poor mirror provision. Clause 15 allowed the Central Government to designate certain categories of personal data as "sensitive". Thus, the designation of new forms of "sensitive personal data", outside of those already included under Clause 2(36) was at the discretion of the Central Government. Importantly, under Clause 15 the DPA and authorities in the relevant sector do not make the categorization

[105] PDP Bill 2019, cl 27.
[106] Personal Data Bill 2022, cl 11(2).
[107] Wachter and Mittelstadt (n 79) 10-17,
[108] ibid; Sandra Wachter, 'Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR' (2017) 34(3) Computer Law & Security Review 436.

but are merely part of the consultation process. From the wording of the clause, it would seem the Government is not bound by the advice received from the DPA and other relevant authorities. It is often in the interests of the Government to define this category of data in a restrictive way. A sitting government would wish to keep "sensitive personal data" narrow and confined to Clause 2(36) as provided, to prevent additional barriers to collecting data. One further point to note is that the rationale behind including or excluding a certain form of data under the "sensitive" category did not need to be provided as per the PDP Bill, 2019.

Apart from this procedural hurdle, there were substantive hurdles in place as well. Clause 15(a) & (c) both use the term "significant harm" as the threshold for data to be classified as "sensitive". This standard seems unnecessarily high. A clear intention exists, therefore, to make the categorization of "sensitive personal data" as limited as possible. Clause 15(d) gives the Central Government the final word on determining whether existing privacy laws are sufficient for ensuring data protection. Finally, Clause 15(b) refers to the "expectation of confidentiality" among Data Principals, but the ability of the Government to make this determination undermines this. No consultation procedure for the Data Principals is mentioned, and as already discussed, no third-party opinion is binding on the Government's final decision.

This reduces the utility of inferences being included under Clause 3(28). An inference in the context of an election is unlikely to fall under Clause 15. The most detrimental effect of micro-targeting is the manipulation of voters and their opinions. Whether this qualifies as "significant harm" is questionable. Regardless, the overarching issue with this Clause remains the Central Government's prominent role in classification, and the minor importance of the DPA and other stakeholders.

This is problematic considering the indirect, but genuine harms caused by data analytics and micro-targeting. However, it is not always simple to demonstrate these harms and to show they fulfil the standard of "significant harm". One of the important advantages of including inferences under personal data would have been the ability to classify inferences used for micro-targeting as "sensitive". Under the GDPR, such inferences could have been afforded a much greater degree of protection. However, this possibility was remote under the PDP Bill given the wording of Clause 15.

The GDPR which has a broad definition of such data had been largely ignored under the PDP Bill, 2019. For example, data regarding political opinions, of paramount importance in an election context, falls under "Sensitive"

data under the GDPR.[109] In the 2022 PDP Bill, the concept of sensitive personal data itself has been removed, thus, completely negating the potential of greater protection for such data along the lines of the safeguards under the GDPR. This is emblematic of the fact that while the 2019 Bill required improvements, it was still preferable to the 2022 Bill that has subsequently been put forward.

Fairness

While these legal loopholes may be found in the rights under the PDP Bill, one could argue that Clause 5 of the 2019 Bill still imposed an obligation upon Data Fiduciaries to process personal data in a "fair" manner. However, the ambit of this requirement is unclear and no guidance was provided under the PDP Bill, 2019 itself. Once again, we may refer to the GDPR's experience with "fairness" requirements in this regard as a similar provision exists within it which has been interpreted by scholars. Eskens has stated that "fairness" as under the GDPR equates to the Data Fiduciary being transparent in its activities. She interprets fairness as being synonymous with concepts such as lawful and transparent. In this iteration, fairness has a fairly limited utility.[110]

An opposing interpretation has been put forward by the European Data Protection Board. The Board's approach to fairness is as a purpose limitation to data collection. Data Fiduciaries would, in this school of thought, be restricted from expanding the scope of what they could use collected data for. Additionally, fairness must also take into account the expectations of Data Principals and the actions of Data Fiduciaries must be in consonance with these expectations.[111] Wachter & Mittelstadt note that one use of the fairness provision could have been to prevent Data Fiduciaries from providing vague and overtly broad purpose uses in their terms and conditions. This practice tends to encapsulate any and all possible uses, which erodes the consent protections that Data Principals are meant to have. However, the prevailing view at present is that the GDPR's "fairness" requirement is nothing more than a transparency tool.[112] Thus, it contains procedural obli-

---

[109] GDPR 2016, art 9.

[110] Sarah Johanna Eskens, 'Profiling the European Consumer in the Internet of Things: How Will the General Data Protection Regulation Apply to This Form of Personal Data Processing, and How Should it?'(*Thesis Research Master Information Law* 2016) <https://www.saraheskens.eu/publications/Eskens-2016-iot.pdf> accessed 14 September 2022; Lee Bygrave, 'Minding the Machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019).

[111] *The European Data Protection Board*, Endorsement 1/2018 (25 May 2018) 5, 9.

[112] Wachter and Mittelstadt (n 79) 93-94.

gations only, and not substantive ones which could help with the question of inferences and their correctness. In light of this, it is unclear how substantive obligations could be derived from the requirement under Clause 5 of the PDP Bill, 2019, especially in the context of inferences given the earlier discussion on the various difficulties associated with applying data rights to them. The point on non-verifiable inferences once again poses a challenge in this regard.

As with several other regressive steps, the requirement for "fairness" in processing of data has been done away with in the PDP Bill, 2022. It is important to recognize the drawbacks that undoubtedly underlined the 2019 Bill in order to improve them. However, the solution was not to truncate the Bill even further as has been done in the 2022 version. The critique provided in this section regarding the various shortcomings in the PDP Bill, 2019, is by no means an implicit approval of the subsequent version of the law that has been proposed. Rather, the subsequent legislation undoes many of the beneficial, albeit flawed, steps that the PDP Bill, 2019 had attempted to take toward a robust protection of online data. What we must contend with now is a 2022 Bill that sacrifices several of even those basic safeguards.

Privacy as a Balancing Act

What should be clear from the discussion outlined above, is that the determination of rights under the PDP Bills is a balancing act. The degree to which Data Fiduciaries may exercise their own rights to keep data collection and the metrics used for creating inferences confidential will have to be set off against the rights of Data Principals. However, this balancing could not be appropriately done unless certain additional regulations are put in place to deal with the gaps in both the PDP Bills that have been pointed out above. These deficiencies will require the DPA to step in and lay down codes of practice, or to be addressed directly in the wording of the revamped 2022 Bill. Given the state of the Bill at present, it seems unlikely that such substantive guidelines will be included before it is finalized.

Additionally, a progressive legal approach toward privacy will be necessary. Given the very nature of data analytics which operates at a massive scale, exemplified by Singh's own accounts, simply looking at privacy as an individual right is insufficient. Moreover, group privacy cannot apply to the kinds of categories of individuals created through inferential analytics. Therefore, a completely different notion of privacy will be proposed in the next part of this paper, which is better equipped to deal with the realities of inferential analytics.

Part III

VII. The Need for Collective Privacy

Having noted the deficiencies in both the 2019 and 2022 PDP Bills vis-à-vis proper regulation of inferences, we can now turn to what the more appropriate means of analysing the competing interests under the Bill would have been. Undoubtedly, whenever a Data Principal raises a grievance regarding a particular practice, some weighing of the different stakes involved will take place to determine whether the data collection practice and usage of said data are permissible. The correct means of doing so should not merely be to look at the rights of each individual Data Principal, but rather the entire collective or category of Data Principals who are placed similarly. Looking at data rights in only an individual way does little to curb the effects of data analytics and micro-targeting. A single Data Principal may bring a claim to protect their rights, but in the context of entire populations or groups of people exposed to such data harvesting, this is merely a drop in the bucket. To effectively place fetters on this practice, the collective requires protection, not just single individuals. Before elaborating on the notion of Collective Privacy, it is important to highlight the deficiencies in two alternative approaches that are generally invoked for dealing with the issue of data mining and drawing of inferences regarding groups of people.

Alternative Approaches to Protecting Privacy of Groups

The general tack followed for addressing such an issue is to refine the already existing provisions in the statute. In the context of the PDP Bills, this may have involved strengthening the provisions on inferences through any number of ways. This is especially the case for the PDP Bill, 2019, given that it accorded explicit recognition to inferences, unlike the 2022 Bill. Thus, our starting point for this examination remains the 2019 Bill, given that the objective is to demonstrate how it could have been improved and protection of inferences made more robust, from the basic platform that the earlier Bill provided. The 2022 Bill, as has been made clear, is a significantly poorer position from which to begin.

The two primary approaches that may be adopted will be addressed to demonstrate why a collective privacy approach is necessary. The first would be to attack the issue at its source i.e., greater restrictions on the collection and processing of data. The second is to make changes to the provisions on

inferences, such as the suggestion provided by Wachter & Mittelstadt to incorporate a "Right to Reasonable Inferences".[113]

With regard to the first issue, a focus on the actual collection of data overlooks the manner in which collectives of people are aggregated using data analytics. The generic practice of classifying groups of people together is on the basis of sensitive personal data. This is because sensitive personal data is, by definition, information that individuals wish to keep private, such as their religious affiliation, ethnicity, and various other immutable characteristics. Aggregations of people via such criteria are used often by law enforcement to identify areas and communities that require special attention. Enhanced policing, surveillance and specialized tactics are then employed in these regions to ensure the suppression of threats. Similarly in an electoral scenario, one would consider that some voters would want such information to remain secret.

Even in a hypothetical world where provisions are introduced which perfectly protect one's sensitive data, such provisions would be unable to address the fact that even completely innocuous and non-personal data which is publicly available will be sufficient for the purpose of creating agglomerations of people for the purposes of micro-targeting. It may take a draconian level of data processing restrictions to prevent even the minutest forms of non-personal data from being accumulated. No legislation in any surveyed jurisdiction contains restrictions of that degree upon non-sensitive personal data. As Singh shows, electricity bills, which do not fall anywhere remotely within the definition of sensitive data, are enough to begin creating a profile of an individual.[114] Combining similarly non-personal data points will be sufficient to determine the exact manner in which to micro-target any collective of individuals.

The second approach could be to make provisions on inferences more stringent. Thus, even if it is difficult to prevent the collection and categorisation of people based on their data, one can try to restrict the drawing of inferences and consequently the actions taken on the basis of those inferences qua that person. However, there are practical difficulties associated with tracking and detecting when an inference has been drawn about individuals. The Snowden disclosures in 2013 revealed that most people are oblivious to the number of ways in which they are being influenced on a daily basis and

[113] Wachter & Mittelstadt (n 79).
[114] Shankar Singh (n 22).

how decisions are being taken based on inferences drawn about them.[115] The tracking en-masse of peoples' behaviour online, as well as programs such as X-Keyscore and trackers within online applications, led to a wealth of information being mined about millions of people, making it possible to then carry out surveillance. It is significant that until the initial story surrounding the existence of such programs was published, there was almost no knowledge in the public domain about these illicit activities and no recognition of the consequences. Thus, this presents an enforcement problem whereby a restriction on inferences cannot be properly implemented.

Regardless of these shortcomings, the primary issue with using the tactic of tweaking provisions of the PDP Bill, 2019, is that it remains focused on an individual-centric approach toward data privacy. The objective in big data analytics, and for political parties, is to target entire collectives of people. In a hypothetical scenario, if a sharp and technologically aware individual is able to discern that they are being microtargeted, they can only raise an objection with the relevant authorities in terms of himself/herself. The difficulty of this has already been commented upon above in the context of the Snowden revelations, and has been elaborated upon at length by scholars.[116] In any case, in the hypothetical scenario, though this individual is able to avoid being targeted it does not in any way stop a party from achieving its primary purpose which is to target the larger collective. When the objective of microtargeting is to influence the collective, individual-centric data privacy rights are insufficient.

Addressing this issue from the standpoint of the privacy of entire collectives of people is in consonance with how parties in elections view their target audience. Logically, the objective is not to influence specific individuals but to provide a broad message that can appeal to the largest number. Hence, the manner in which the data privacy rights of people are protected must necessarily reflect that reality rather than remain focused on purely individual rights.

---

[115] European Parliament Resolution of 12 March 2014 on the US NSA Surveillance Programme, Surveillance Bodies in Various Member States and their Impact on EU Citizens' Fundamental Rights and on Transatlantic Cooperation in Justice and Home Affairs (2014) <http://www.europarl.europa.eu/sides/getDoc.do?type=TA&language=EN&reference=P7-TA-2014-0230> accessed 5 May 2022.

[116] Lee Bygrave, *Data Protection Law: Approaching its Rationale, Logic and Limits* (Kluwer Law International 2002).

Group Privacy

The Right to Privacy largely follows an individual rights model. The implication of this is that the right attaches primarily to an individual rather than a collective or group entity. Bygrave notes that this holds true for both European and American law, two places where privacy law has developed more than others.[117] There has been a slow evolution of this position with the Article 29 Working Group recognizing the concept of a collective interest in privacy rights.[118] This has largely been seen as an agglomeration of multiple individual grievances and not as a separate and independent legal right in and of itself. Mantelero notes that this interpretation has meant that relief granted is usually premised on protection of individuals within the agglomerations, rather than collective relief.[119] Therefore, the development of data privacy law has only come to the stage of group privacy, rather than collective privacy.

Contemporaneous data privacy scholars have identified two primary concepts of group privacy. The first, concerns the privacy of individuals within the group's settings.[120] The second concerns an entity which is recognized as a body of individuals in law with that entity itself having privacy rights.[121] Neither of these concepts are useful in the context of elections. The peculiar situation in political microtargeting is that the individuals so targeted do not necessarily identify as a group. They are disparate and non-aggregated individuals whose only commonality is the inferences that political parties draw about them. This restrictive idea of group privacy has been recognized and Bygrave has posited a more suitable alternative. He proposes that individual privacy rights in statutes be transposed to such collectives which do not self-identify as such.[122]

---

[117] Lee Bygrave, 'Privacy Protection in a Global Context—A Comparative Overview' (2004) 47 Scandinavian Studies in Law 319; Article 29 Data Protection Working Party 'Letter to Mr. Larry Page, Chief Executive Officer' (2013) <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2013/20130618_letter_to_google_glass_en.pdf> accessed on 18 October 2021.

[118] Article 29 Data Protection Working Party, 'Letter to Mr. Larry Page, Chief Executive Officer' (2013) <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2013/20130618_letter_to_google_glass_en.pdf> Accessed 18 October 2021.

[119] Alessandro Mantelero, 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era' in Linnet Taylor, Luciano Floridi and Bart Van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies* (Springer Link 2017).

[120] Edward Bloustein, 'Group Privacy: The Right to Huddle' in Edward J Bloustein, Nathanial J Pallone (eds), *Individual and Group Privacy* (Routledge 2018).

[121] Alan Westin, *Privacy and Freedom* (Ig Publishing 1970).

[122] Bygrave (n 110).

Collective Privacy

Mantelero has built on this to elaborate how the subject of data privacy law must shift away from individuals/recognized group entities, toward clusters of individuals who do not identify as a collective but are grouped as collectives on the basis of the inferences made about them.[123] Individual privacy rights must be suitably modified in such situations to provide adequate relief to such disaggregated collectives.[124] This conception of "collective" is consistent with the manner in which micro-targeting occurs. Singh's elaboration of how data gathering targeted tribals in Tripura and non-Yadav OBC's in Uttar Pradesh, demonstrates how these individuals are implicitly classified together without them knowing that they are part of a collective at all. They are not recognized as a group or juristic entity in law and their grouping occurs only in the context of inferential data analytics.[125] Recall how certain people were categorized together on the basis of their electricity bills. These people would not even think of themselves as a collective, given most of them would not even know of each other.

These categories are created based on data analytics and recognizing certain clusters of information from the data collected. All individuals with high electricity bills are formed into a cluster, based on their bills. While these individuals remain oblivious to what is happening, for the analyst they are an autonomous category, primed for a particular form of election advertising.[126] This form of categorization has become increasingly common, especially in the context of elections and for law enforcement.[127] However, these collectives do not currently have recognition in data privacy law in any jurisdiction.

What is important about such collectives is that the way the individuals are affected is not based on their individual data, but rather on the collective data regarding their respective clusters.[128] Thus, the decisions made regarding micro-targeting are based on the overall inferences drawn about the group to which they belong. This issue is further complicated by the fact

---

[123] Mantelero (n 119) 143.
[124] Mantelero (n 119) 144.
[125] Shivam Shankar Singh (n 22) 75.
[126] ibid 64, 75.
[127] Oskar Josef Gstrein, Gerard Jan Ritsema van Eck, 'Mobile Devices as Stigmatizing Security Sensors: The GDPR and a Future of Crowdsourced "Broken Windows"', (2018) 8(1) International Data Privacy Law 69.
[128] David Bollier, *The Promise and Perils of Big Data* (*Aspen Institute, Communications and Society Program, Washington, DC*, 2010) <https://www.aspeninstitute.org/wp-content/uploads/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf> accessed 8 September 2022.

that individuals will belong to multiple such clusters in big data analytics. Specific individuals will be subject to different types of advertising or treatment, depending on the different clusters they are grouped into. Due to the fact that they do not identify as a "group", and thus, group rights cannot be transposed onto them[129] Mantelero enunciates a "collective privacy" principle, which looks at clusters of individuals who do not identify as groups.[130]

This approach to privacy is far more effective than looking at individual interests. The reason for this is that through data analytics a group norm is culled out, rather than the individual norms.[131] For instance, a particular part of a city may be deemed to have one primary concern or a particular political leaning due to demographic factors such as predominant religion and average income. Based on data collected about the majority of individuals in that area, only certain kinds of advertising and information will be provided to them. Those individuals who may not fall into the majority category may be exposed to certain kinds of advertising based on the collective into which they have been placed due to their residence in that part of the city. These individuals may have different opinions regarding this kind of collective electoral profiling i.e., not all of them may find it to be problematic.[132] Thus, if an individual-centric approach is taken, some of these individuals who are targeted may not, for various reasons, care for the collective interest of the electorate to not be targeted in such a manner. One such cause could be that the individual supports the targeted messaging as the party undertaking it conforms to the specific political outlook and ideology of that voter. Hence, the individual believes that society would benefit from these messages being sent directly to them without a thought for how this may be part of a strategy of manipulation and implicit coercion.

Parties are unconcerned with providing the electorate with all the necessary information regarding any particular issue. Their focus is to only elaborate on those points that they determine will encourage people to vote in their favour. The ability of parties to do this in a sophisticated and precise manner is accentuated by data analytics. The question that then arises is how best to protect the collective from undue influence. Newman proposes that the collective interest should be what governs the decision regarding whether such collective profiling is detrimental or not.[133] Depending on the

[129] Luciano Floridi, *The Ethics of Information* (OUP 2013).
[130] Mantelero (n 119).
[131] Pam Dixon and Robert Gellman, 'The Scoring of America: How Secret Consumer Scores Threaten your Privacy and your Future' (World Privacy Forum, 2014).
[132] Mantelero (n 119) 148.
[133] Dwight G Newman, 'Collective Interests and Collective Rights' (2004) 49(1) American Journal of Jurisprudence 6.

context, different collectives may arise. For elections, the freedom of the vote and collective interest against disinformation or half-truths is of relevance.[134] This interest would not be dependent upon the individual voter, as highlighted above, but rather centre around a general principle that underpins the objective of free, fair, and legal elections. Thus, in line with Newman's suggestion, inferences and microtargeting which go against the collective interest of having free and fair elections must be ceased.

This has also been recognized by Koss and Perry who advocate a shift away from purely individual rights and toward rights against the inferences made as a result of clustering people in this manner.[135] Disaggregated people being made into a collective have an interest in their privacy and ownership over personal information at an individual level.[136] However, over and above those individual concerns, the issues at a collective level would involve possible negative consequences of inferences made regarding them as a collective rather than as individuals.[137] This is the crucial distinction that the recognition of collective privacy allows us to address.

Thus, the special concern of data privacy law in the context of elections needs to be inferences made about disaggregated collectives of people.[138] Both of these aspects were insufficiently addressed in the PDP Bill, 2019, which followed the typical and standard individual-centric approach. In any case, inferences are inadequately protected under the PDP Bill and the rights provided under Chapter V had several barriers to their applicability. The question of collective privacy is completely missing from the PDP Bill, 2019, and there is no indication of how it is supposed to be dealt with. The 2022 Bill, as has been elaborately detailed, is even more deficient in significant and profound ways on this point. Thus, the legal framework was lacking in terms of its ability to deal with inferences in the context of collective privacy and

---

[134] Snigdha Poonam and Samarth Bansal, 'Misinformation is Endangering India's Election' *The Atlantic* (1 April 2019) <https://www.theatlantic.com/international/archive/2019/04/india-misinformation-election-fake-news/586123/> accessed 25 September 2021.

[135] Walter Perry and others, 'Predictive policing: The Role of Crime Forecasting in Law Enforcement Operations', (Rand Corporation 2013) <http://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf> accessed 2 October 2021; Kelly K. Koss, 'Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post-Wardlow World' (2015) 90 Chicago-Kent Law Review 301.

[136] Fred Cate and Viktor Mayer-Schönberger, *Data Sue and Impact* (*The Centre for Information Policy Research and The Centre for Applied Cybersecurity Research, Indiana University* 2013) <http://cacr.iu.edu/sites/cacr.iu.edu/files/Use_Workshop_Report.pdf> accessed 25 September 2021.

[137] John Finnish, 'The Authority of Law in the Predicament of Contemporary Social Theory' (1984) 1(1) Notre Dame Journal of Law, Ethics & Public Policy 115.

[138] Edward Bloustein, *Individual and Group Privacy* (Routledge 1978).

is required to be comprehensively addressed in the revamped Bill that is currently being contemplated. The current rendition of the 2022 Bill, however, leaves even more to be desired than the PDP Bill, 2019.

Enforcement of Collective Privacy Rights

Now that the specific purpose and importance of collective privacy have been detailed, the issue of how to enforce such a collective privacy right must be addressed. Mantelero proposes that the respective Data Privacy Authorities of different countries make the risk assessment of these inferences.[139] However, the issue with this has already been elaborated upon where there is a specialized agency such as the EC already in place to address election-related concerns. An intriguing solution is put forward by Kammourieh who relates the Right to Privacy of a group/collective to several human rights principles.[140]

The Right to Privacy of a group may be related to the Right to Dignity, specifically the right to Self-Determination. The right is rooted in Article 1 of the International Covenant on Civil and Political Rights ("ICCPR").[141] The original interpretation of Article 1 was confined to external self-determination which legally allowed colonized peoples to declare independence.[142] The evolution of the right has now encapsulated an internal right to self-determination. This allows groups within groups to demand constant social and political rights. In this way, the notion of self-determination has expanded to include rights such as "informational self-determination" which may be directly relevant to data analytics and collective privacy rights.[143] However, Paton recognizes that the ICCPR contemplates the exercise of these rights by recognized entities or groups. In this context, there is no recognized group but merely passive clusters that are created due to them being clubbed together by data analytics. Such passive clusters or collectives cannot exercise rights under the ICCPR.[144] Due to this, Kammourieh falls back on policy recommendations and changes to help such passive collectives better regulate the collection and usage of their data.

---

[139] Mantelero (n 119) 150.

[140] Lannah Kammourieh, 'Group Privacy in the Age of Big Data' in Linnet Taylor, Luciano Floridi and Bart Van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies* (Springer Link 2017) 54.

[141] International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR) art 1; Antonio Cassese, *Self-Determination of Peoples: A Legal Reappraisal* (CUP 1995) 11.

[142] Jonathan Charney, 'Self-Determination: Chechnya, Kosovo, and East Timor' (2001) 34 Vanderbilt J of Tech L 455; UNGA Res 1803 (1962) GOAR 17th Session Supp 17, 15.

[143] Kammourieh (n 140) 54.

[144] George Whitecross Paton, *A Textbook of Jurisprudence* (4th edn, Clarendon Press 1972).

Additionally, Mantelero correctly notes that one cannot protect collective rights the same way one would protect corporate group rights. Corporate rights can be enforced by the central authority of the organization or entity. Collectives of the kind described here, do not have any central authority at all and are atomistic in nature.[145] For collective privacy to fulfil its purpose, a method to enforce it is necessary.[146] To achieve this outcome, it is possible to look at collective or general interests that are protected by centralized authorities. Examples of this include financial regulators and consumer redressal bodies which seek to ensure the basic and general protection of consumers and investors.[147] Similarly, the DPA could focus on the general interest of the collective in ensuring that it is not manipulated and its right to a free vote is not hindered.

However, a return to the DPA would once again mean dependence on centralized authority. If this were seen as the only viable solution, it would remain a problematic one as a centralized authority remains more susceptible to political influence and bias. Mantelero's suggestion of having such a centralized authority act as gatekeeper seems feasible at face value but he acknowledges one of the greatest concerns with this would be the impartiality of the authority.[148] In India, this role could be fulfilled by the combination of both the DPA and the EC. The scrutinization of data collection, especially that of inferences and the metrics on which such inferences are reached, and their utilisation for microtargeting can be apportioned between both authorities depending on various factors.

The determination of the potential detriments of such collection from the perspective of its consequences for political disinformation, classification of constituents and microtargeting, and undue influence of voters, requires the expertise of both the DPA and EC. Such considerations must be looked at in terms of the consequences for the individual as well as the disaggregated collective to which that individual belongs.[149] The interests of this collective,[150] which is created by the Data Fiduciary through inferential analytics, must be analysed both from the perspective of unfair means of collecting the data and the criteria provided for the inferences drawn about voters as collectives.

---

[145] Bygrave (n 110).
[146] ibid; Mantelero (n 119) 150.
[147] Mantelero (n 119) 150.
[148] ibid 152.
[149] Danielle Keats Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions' (2014) 89(1) Washington Law Review 1.
[150] Alessandro Mantelero, 'The Future of Consumer Data Protection in the EU Rethinking the "Notice and Consent" Paradigm in the New Era of Predictive Analytics' (2014) 30 Computer Law & Security Review 643.

This would, hypothetically, be the province of the DPA. Simultaneously, the EC may investigate the potentially detrimental effects such a practice may have on elections and whether it amounts to a violation under Section 123(2) of the RPA.

Another possible method would be for the DPA, in conjunction with the EC, to carry out Data Protection Impact Assessments.[151] Such Assessments have been proposed in various jurisdictions across the world in different contexts and can provide a benchmark for determining what activities may pose a risk to collective privacy. DPAs are often best placed to identify technological innovations and practices which could pose a risk to collective privacy. Thus, it may examine the various methods by which parties collect constituents' data and act accordingly. This is a form of predictive analytics which makes a presumption about potential dangers from specific activities of data collection and inference drawing. As Cohen and Floridi point out, such predictive analytics requires a proper understanding of the social or political consequences of the specific data practice in question.[152] This cannot be done by the DPA in a vacuum and must involve the EC and its expertise and knowledge of the way that elections function and how voters can be influenced. This goes back once again to Mantelero's suggestion that a multi-department collaboration is necessary to effectively enforce collective privacy rights.

The specifics of how these two institutions may work together to frame rules for the conduct of parties must necessarily be context specific. Wright correctly notes that general guidance in this realm is difficult to lay down.[153] Rather, criteria for governing the mining of data, drawing of inferences, and then using such inferences to aggregate people for the purpose of microtargeting, will necessarily depend on the specific legal framework and social context. This would involve a far more in-depth analysis of how to balance different interests and the specifics are not the subject of this paper.

While this may be a rough demarcation in terms of the responsibilities undertaken by the DPA and EC respectively, it is unclear whether there can be a guarantee of their independence from the Central Government in power

---

[151] Mantelero (n 119).

[152] Luciano Floridi, *The Fourth Revolution: How the Infosphere is Reshaping Human Reality* (Oxford University Press 2014); Julie Cohen, 'What Privacy is for' (2013) 126 Harvard Law Review 1933.

[153] David Wright, 'A Framework for the Ethical Impact Assessment of Information Technology' (2011) 13(3) Ethics and Information Technology199; Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford University Press 2010) 231; Mantelero (n 119) 150.

at the relevant time. The DPA has not, of course, even been formed yet. The EC is routinely accused of bias and failure to ensure completely free and fair elections.[154] The appointment of the DPA is to be determined by the Cabinet of the Central Government, which could create issues of independence as well as individuals sympathetic to whichever ruling government exists could be appointed to the DPA.[155] Regardless, it is necessary to recognize the concept of Collective privacy and to take whatever steps are feasible to ensure the greatest amount of impartiality and rigour on the part of regulators.

In the current scenario of data protection and elections in India, a collaboration between the EC and DPA seems to be the most appropriate means by which to enforce Collective Privacy. Various scholars have noted that this solution lies in the realm of policy where balancing must be done between different interests and components within society.[156] Data Protection Authorities do not normally concern themselves with the impact of the use of data (influence and manipulation of voters), but rather, focus on the collection of data itself. However, as exhaustively elaborated upon above, it is necessary to incorporate enforcement that looks at the negative effects of data mining and drawing inferences on a collective scale. This is the stage where the EC, which focuses on the actual impact of inferences drawn about collectives of voters, would be needed.

To this extent, Mantelero's proposed solution involves having multiple agencies or departments involved in order to ensure that different stakeholders and interests are adequately addressed.[157] Examples of this have been seen in consumer protection and labour law.[158] In the former, it is in the collective interest of all consumers to ensure product security and prevent unfair commercial practices, however, consumers are atomized and not connected to each other. In order to enforce these general interests, institutions must take the lead in setting policy. These steps would not be taken in isolation by the DPA. In the electoral context, the implementation of safeguards on data protection and inferences would be of minimal assistance without the expertise of the EC in addressing the actual manner in which such data

---

[154] 'Election Commission Biased, Poll Schedule Made to Benefit PM Narendra Modi: Rahul Gandhi' *India Today* (17 May 2019) <https://www.indiatoday.in/elections/lok-sabha-2019/story/election-commission-partial-poll-schedule-made-for-pm-modi-s-campaigns-rahul-gandhi-1527525-2019-05-17> accessed 15 October 2021.

[155] Personal Data Bill, cl 42.

[156] Wright (n 156).

[157] Mantelero (n 119).

[158] European Commission, *Second stage consultation of social partners on the protection of workers* personal data' 7, 10, 16–17 <http://ec.europa.eu/social/main.jsp?catId=708> accessed 3 March 2022.

and inferences are used to influence elections. Thus, a collaborative effort between both institutions would be required in order to appropriately protect the collective privacy of large groupings of people who are ill-equipped and poorly placed to ensure their rights are not violated.

Part IV

VIII. Conclusion

For data privacy law, the collective matters as much as the individual. If the subject of the law was merely the individual, it could serve to protect him/her but do little to deter the drawing of inferences about, and profiling of, the entire collective. In modern data analytics, the profiling of a single individual is of relatively little importance. The targets of data collectors and those interested in targeted advertising are entire collectives of people.

The logic is a fairly straightforward one i.e., through data analytics about large factions of people, the greatest number of them may be implicitly manipulated. For a commercial entity, that can mean persuading sufficient people to choose their product over those of competitors. For social media and other websites, it is to market that data to advertisers for revenue. And for States and political parties, it is to persuade voters during elections or to convince them regarding the wisdom of their respective policies while in office. Privacy law, therefore, must keep pace with the ways in which data collection takes place and how it is used. As the Edward Snowden disclosures revealed, at the best of times, individuals are mostly unaware of the number of ways in which their activities are being monitored and how their behaviour is being shaped implicitly.

One of the greatest threats that arise from these practices is the possibility for entities to manipulate elections. While this is usually thought of in the context of external manipulation by foreign enemies, internal actors could just as easily influence politics and society. The news that the PDP Bill, 2019, was withdrawn, after 3 years of waiting, was disappointing and shows that recognition of data privacy is still at its most nascent stage in India. However, it also provides an opportunity, in that the gaps identified in the PDP Bill can be addressed more holistically in a reworked legislation. This hope has not been crystallized in the PDP Bill, 2022. The revised Bill has gone in the opposite direction in many ways, including removing the explicit recognition accorded to inferences as being personal data, the removal of sensitive personal data and the obligations that are imposed on Data Principals under Clause 16 which have a chilling effect on the enforcement of rights under

Chapter 2 of the 2022 Bill. Thus, the only appropriate course of action is an about turn on various steps taken in the 2022 Bill that aggravate the issues associated with inferences and microtargeting. The PDPD Bill, 2019, is a better starting point from which to gauge the improvements that must be brought about in our data privacy laws.

Thus, a restoration of the recognition of inferences as personal data is a sine qua non for bringing the PDP Bill in tune with modern methods of microtargeting. The PDP Bill, 2019, which started from the position that inferences fall under the ambit of personal data, is more attuned to this reality than the 2022 version. Further, the corpus of laws as contained in the 2019 Bill must be complemented by a recognition of collective data privacy and provisions for conducting Data Protection Impact Assessments, regardless of the type of Fiduciary. Any Fiduciary, regardless of whether it is significant or not, should be required to conduct such Assessments given the significant impact that data privacy in this area has on politics and social engineering via microtargeting. In order to ensure enforcement of collective privacy in this context, the DPA and EC must coordinate with each other and implement an improved PDP Bill which includes a recognition of this approach to data privacy, in tandem with electoral laws. This a modern reality of data privacy that the law must equip itself to address.