Sceptical Optimism? Dealing with the Problems of Our Time
Vol. 1, No. 2/ 2023

DOI: 10.46586/eelp.1.2.19-29
ISSN 2940-3065

Data Monopolization by the Tech Industry:
Implications for Democracy

Shumaila Hussain Shahani
Ruhr-Universität Bochum, Germany

Seminar Paper (April/ 2023)
Seminar: Ethics and Politics in the Digital Era

How to cite this article
DOI: 10.46586/eelp.1.2.19-29
1 Introduction

In the contemporary era, the tech industry possesses vast repositories of both user-generated and non-user data (Bond et al. 2012, Garcia 2017) accumulated through their digital platforms, which are utilized by billions of people worldwide. However, concerns have been raised about how this data is being used. For instance, during the 2016 U.S. Presidential election, instances of Russian interference using targeted advertising came to light. Similar allegations were made regarding the Brexit referendum and the 2017 French elections, which further highlighted the potential for human data to be exploited for political gain. These incidents have heightened concerns over the manipulation of democratic processes through the exploitation of personal information.

This paper argues that the unparalleled access to human data enjoyed by the tech industry has led to an unprecedented concentration of power and resources, allowing them to exert significant control over the thought processes and decision-making abilities of billions of people worldwide. This phenomenon has given rise to a new economic regime known as Surveillance Capitalism (Zuboff 2019). Zuboff describes this new economic order as one that “claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales” (2019, definition, no page). Essentially, the regime of Surveillance Capitalism is built upon the expropriation and exploitation of human data, which poses a direct threat to human autonomy and individual agency – the fundamental components of a democratic society. Moreover, this paradigm creates a power asymmetry by concentrating a significant amount of power through data control in the hands of a select few, leading to far-reaching social, economic, and political consequences. Zuboff argues that this power dynamic amounts to an “overthrow of the people’s sovereignty” (2019, definition, no page), highlighting the exigency for ethical considerations and comprehensive regulatory frameworks to reassert people’s sovereignty in global politics.

The paper begins with outlining methodologies employed by companies for data collection, the nature of data collected, and the likelihood of its manipulation for political gains. This section not only expounds on the scope and magnitude of data collection but also highlights the extent to which tech industry exercises psychological control over the populace, thereby wielding significant influence on global politics. The subsequent section delves into the underlying systemic factors contributing to this issue, expounding on the inherent flaws within the data-driven business model of digital platforms. I argue that it is not feasible for tech companies to prioritize privacy without risking profitability, given their heavy reliance on data collection. The paper concludes by underscoring the concentration of global political power in the hands of a select few, and the significant implications this has for global politics. The central thesis of the paper asserts that the unregulated accumulation of data by the tech industry has led to the consolidation of power among a select few, resulting in the erosion of democratic principles.
2 Data-driven Business Model of Digital Platforms

To understand the risk of data misuse and its potential to subvert democratic processes, it is imperative to scrutinize the business model of digital platforms that rely on data acquisition as a primary aspect of their operations. Technology companies use algorithms to improve user engagement, which is essential to retain their customer base and drive revenue growth. To achieve this, they gather vast amounts of data from various sources to build comprehensive data repositories. Through big data analytics, this information is analysed to generate user profiles and identify patterns and correlations that reveal user behaviour and preferences. The analysis enables platforms to gain valuable insights into user behaviour, improve user engagement and retention, and drive revenue growth.

To acquire this data, the companies often rely on complex and interconnected sources such as social media platforms, search engines, and mobile devices. This data collection is becoming increasingly sophisticated, with the ability to capture a vast array of information about users’ habits, social relationships, preferences, attitudes, thoughts, opinions, heartbeats, sleep patterns, and even dreams. A New York Times article examined Facebook’s patent applications to determine various methods employed by the company to obtain data from its users (Chinoy 2018). Facebook has demonstrated its capability to identify specific television programs being viewed through a mobile phone’s microphone, track the duration of sleep by monitoring phone activity, and predict major life events like birth, death, weddings through credit card transactions. Moreover, Facebook has the capacity to identify pictures captured with the same device by establishing a unique camera signature and detecting minute details such as lens scratches or pixels within the images. Such details are leveraged to draw connections between users who have uploaded photos taken with the same camera and predict the strength of their relationship based on the frequency with which they use the shared device.

These data points are then ingeniously correlated with other data points to construct detailed psychographic profiles of users, providing valuable insights that drive revenue growth in a myriad of ways. For example, social media platforms utilize data to customize the content displayed on users’ feeds. This personalized approach ensures that users receive content that is aligned with their interests and preferences, ultimately leading to increased engagement and retention. By retaining their user base, tech companies amass more user data to enhance their algorithms, resulting in even more personalized content. This, in turn, leads to an increase in revenue generated from ad impressions and clicks. Moreover, tech companies utilize algorithms to deliver targeted advertising to users based on their interests, demographics, and behaviour, increasing the likelihood of users engaging with the advertisements on their platform. By monitoring website traffic and user activity, businesses gain a deeper understanding of their customers’ needs and preferences, allowing them to prioritize their efforts more effectively. For instance, through data analysis, businesses can identify inefficiencies in the supply chain and suggest ways to improve it, such as scrutinizing supplier delivery times, determining the most efficient delivery routes, and adjusting inventory levels to meet customer demand. This can help reduce costs, increase profitability, and improve overall business
operations. Furthermore, businesses can also explore new sources of revenue through data. For example, a firm specializing in electronic products could spot the increasing trend of home automation through data analysis and decide to expand its product line to include smart home products to capture this emerging market.

Additionally, advertising companies can tap into this data to design highly targeted and efficient advertising campaigns. The effectiveness of this approach lies in ensuring that the right message is conveyed to the right audience, thereby maximizing the impact of advertisement. For instance, if a user has been searching for a new car online, they may start to receive ads for car dealerships or car insurance companies. Using advanced data analytics techniques, advertising companies can predict and cater to the future interests of consumers with a high degree of precision. A concrete example of this could be seen in the recommendation systems utilized by e-commerce websites, such as Amazon. Based on a user’s browsing and purchase history, Amazon’s algorithms can predict the customer’s future interests and make personalized product recommendations. This helps Amazon reach niche audiences that are more likely to be interested in their products or services. By doing so, Amazon increases the likelihood that users will engage with the advertisements and ultimately make a purchase.

Another example of predictive analysis is the Google’s autocomplete algorithm, which can predict and suggest search queries to individual users through the analysis of the most commonly searched terms by other users. The primary benefit of this service is that it reduces keystrokes and provides relevant suggestions in real-time, thereby saving time and effort for the user. This approach to search queries allows users to quickly find what they are looking for and also enables Google to tailor its search results and advertising to individual user’s preferences. This ultimately increases the likelihood of user engagement and revenue growth for the company. It is evident that data acquisition is an essential tool for digital platforms to gain valuable insights, make informed decisions, increase profitability, and maintain a competitive edge in the marketplace (Qi & Tao 2018).

The practice of data acquisitions also benefits users in a multitude of ways. By allowing access to information such as browsing history, search queries, and demographic data, users enable advertisers to provide ads that align with their interests and requirements. This personalized approach to advertising can assist users in discovering products and services that may have otherwise gone unnoticed and in a more timely and efficient manner. For instance, if an individual has been scouring various websites and performing online searches for laptops, advertisers can leverage this information to showcase ads for laptops from various brands with feature specifications and pricing that correspond with the user’s preferences. As a result, user can identify laptops that may have otherwise been overlooked and make a well-informed decision when purchasing. Furthermore, personalized ads are reportedly more engaging and efficacious in capturing users’ attention. This approach can make users feel acknowledged and valued by advertisers, ultimately leading to a better overall brand experience. The convenience and personalization are reportedly major factors that motivate users to willingly provide their data to technology companies (Calvin 2017, Bass 2019). However, it is essential to assess whether this consent is truly informed and made with full knowledge of the implications of such data sharing.
It is worth noting that the data collected by tech companies represents only a small portion of the overall data collection process, which extends beyond the scope of these companies (Deibert 2019). In addition to the frontline companies that explicitly request permission for data collection, there are various other entities involved in this multifaceted process. These entities include analytics businesses that utilize the data harvested by the frontline companies to construct psychographic profiles of users for their clients, who then use these profiles to employ micro-targeting tactics. Moreover, there are companies that specialize in developing and supplying algorithms, software, techniques, and tradecraft to both frontline companies and analytics firms, augmenting their capabilities and facilitating more effective use of the collected data. The entire digital infrastructure’s sustainability is reliant on a vast network of businesses that provide essential hardware, software, and energy necessary to maintain these operations. Thus, the business model of frontline companies that users consent to provide their data to further entails engagement in business-to-business transactions and flow of data that many users remain oblivious to. Consequently, this may result in users unwittingly providing vast amounts of personal data to several other services, through an obscure process of information-sharing.

3 Inherent Risks to Democracy

As established above, the revenue model for tech companies involves collection of vast amounts of user and providing third-party developers, applications, and services with a highly effective and measurable method for targeting potential customers. In other words, these platforms have a direct incentive to collect as much information as possible, including personal information like political preferences, beliefs, and behaviours to enable more precise targeted advertising. The more precise the targeting of ads to users, the higher the engagement rates and revenue generation for the company. However, this profit model can come at the cost of user privacy, as these platforms may engage in practices that prioritize corporate gain over the protection of personal data. This creates opportunities for political manipulation tactics that exploit personal data to sway public opinion or promote specific agendas. One prominent example of this occurred in the lead-up to the 2016 U.S. presidential elections. An external researcher harvested the personal data of up to 87 million profiles through a personality profiling application on Facebook (Heawood 2018). This data was sold to Cambridge Analytica, a political consulting firm that provided analytical support to political candidates. The collected data was used to construct psychographic profiles of voters, a research method that segments population groups based on psychological variables such as personality traits, values, attitudes, interests, socio-economic status, media preferences, and behavioural data. The psychographic profiles were then used to create customized political messages for persuasive psychological targeting, which furnished the Trump campaign with a powerful tool for influencing voters (Guess, Nyhan, Reifler 2018). The Google autocomplete algorithm’s search suggestions have also come under scrutiny due to their tendency to promote sensational or controversial content over factual accuracy. Investigations into the algorithm have revealed that, in the lead up to the 2021 Capitol invasion, Google’s autocomplete algorithm suggested search terms related to civil war that did not align with actual search volumes (Chaslot 2021).
Similarly, search options related to COVID-19 and climate change have also been found to be inconsistent with actual search volumes on different occasions. These findings raise concerns about Google algorithms prioritizing sensational or controversial content over factual accuracy, possibly to increase user engagement and retention, as such content tends to generate more attention and emotional reactions from users.

Moreover, algorithms designed to display content based on users’ preferences have the potential to exacerbate political polarization and impede democratic deliberation. For instance, if someone holds the belief that climate change is a hoax, they are more likely to engage with content that supports that view, and thus the algorithm will show them more of such content. This is because humans are naturally predisposed to confirmation bias, where they seek out information that confirms their pre-existing beliefs and disregard or avoid contradictory information (Nickerson 1998). Accordingly, users are more likely to engage with content that aligns with their views. Empirical evidence supports this argument, as demonstrated by a 2016 study that examined the online interactions of 376 million Facebook users across over 900 news outlets. The study revealed that individuals tend to gravitate towards news that reinforces their existing viewpoints (Schmidt et al. 2017). Consequently, the majority of Americans who consumed false information during the 2016 U.S. Presidential elections were Trump supporters or individuals with conservative political opinions (Guess, Nyhan, Reifler 2018). However, this phenomenon creates echo chambers that serve as a feedback loop, amplifying and validating pre-existing beliefs while limiting exposure to differing perspectives and silencing alternative viewpoints. As a result, users may be less likely to consider or engage with viewpoints that differ from their own, ultimately impeding meaningful democratic discourse and deepening political biases, which lead to a hostile and divisive social and political climate.

Research further suggests that users’ perceptions and behaviour are highly susceptible to the content they consume (Neubaum & Krämer 2016). This means that exposure to content expressing a particular opinion induces users to adopt that viewpoint, thereby creating a reciprocal relationship between user’s opinion and their perception of the prevailing public opinion on the matter. A separate study on the interplay between suggestion, cognition, and behaviour indicates that both intentional and unintentional suggestions have the ability to impact an individual’s cognitions and behaviour (Michael, Gerry, Kirsch 2012). Consequently, when users encounter the suggestion “civil war is inevitable” from Google’s autocomplete algorithm, they may perceive it to be correlated with the search volume on the platform, indicating that it reflects what most people are thinking at that moment (Chaslot 2021). As a result, users may perceive the suggestion as being grounded in truth and feel compelled to take action.

Furthermore, the use of micro-targeting mechanisms is inherently manipulative in nature, rendering the data collected susceptible to political manipulation. Apart from collecting personal data, micro-targeting shapes user behaviour and opinions through tailored content and messaging, thus creating a breeding ground for political actors to exert influence over public opinion and behaviour by means of targeted disinformation campaigns. A study conducted on
the 2010 U.S. congressional elections examined the impact of political micro-targeting directed at a vast cohort of 61 million Facebook users on their subsequent voting behaviour (Bond et al. 2012, Garcia 2017). The findings showed that these messages had a quantifiable and direct impact on the political self-expression, information-seeking behaviour, and real-world voting behaviour of millions of individuals. Further, the influence of these messages went beyond their direct recipients and had a cascading effect on their friends, and even friends of friends. Additionally, these platforms serve as a conduit for advertisers to micro-target users with information that may be deceitful or biased. One such method employed to achieve this objective is the use of dark ads, a technique that permits the distribution of content to a select audience without it being publicly visible. Dark ads have the potential to present a skewed or one-sided view of a particular issue, in order to influence users’ opinions or behaviours without the scrutiny or accountability that comes with public advertising.

Micro-targeting can also be used to engage in more insidious forms of manipulation, such as spreading fake news through bots1 (Vosoughi, Roy, Aral 2018). In the 2016 U.S. Presidential elections, Russian bot accounts played a significant role in micro-targeting users with false election-related news. As a result, around 25% of Americans were exposed to misleading and false information related to the elections (Guess, Nyhan, Reifler 2018, Badaway, Ferrara, Lerman 2018). Further investigation exposed a market for reusable political disinformation bots, which can be utilized across multiple campaigns (Ferrara 2017, Nied et al. 2017). Ferrara’s study specifically identified bots that propagated narratives associated with far-right ideology during the 2016 U.S. Presidential election campaign (Ferrara 2017). These bots were found to have become inactive after the U.S. Presidential elections, only to resurface during the lead-up to the 2017 French Presidential election. A parallel pattern was noted in the context of Brexit, whereby a sizable number of bot accounts, estimated to be 13,493 in number, were found to engage in amplification of false news that favoured the Leave EU campaign (Bastos, Mercea 2017).

The aforementioned events elicit a legitimate cause for concern regarding the ethical use of personal data. The tech industry’s data collection practices carry significant implications for democracy, a system premised on the principle of popular sovereignty. Although democracy is a complex concept with diverse forms and various elements, certain fundamental conditions must be satisfied for a political entity to be regarded as democratic.2 Among these conditions is the idea of popular sovereignty, which maintains that the primary source of political power in a democratic society rests with the people. In a democratic system, the people are the ultimate authorities or rulers, and the government’s legitimacy derives from their consent. The idea of popular sovereignty thus serves as a critical safeguard against the abuse of power and the erosion of democratic principles. However, when psychological manipulation

---

1 In July 2018, Twitter made a public announcement stating that they were deleting approximately one million bogus accounts daily. Surprisingly, following this announcement, the company’s stock price plummeted. This suggests that there may have been business incentives for tech companies not delving too deeply into their own platform to eliminate bots. (The Guardian 2018)

2 All of these sources define democracy differently with some variations. However, one common element among all of them is the sovereignty of the people. See e.g., European Commission (2020), Council of Europe (2022), Schmitter & Karl (1991), Schumpeter (1943).
interferes with political preferences of people, it distorts the democratic process and compromises the legitimacy of the electoral outcomes (Persily 2017).

4 National Security Concerns

The discourse surrounding the potential threat that certain practices pose to democracy has also met with dissenting voices. While some argue that such practices are not necessarily detrimental to democratic ideals, (Kefford et al. 2022) others remain sceptical (Kokas 2022, Calzada 2023). Recent developments, however, suggest that political leaders have recognized the political significance of human data.

A 2017 report from Freedom House, assessing the state of political rights and civil liberties worldwide, shed light on the pervasive practice of social media manipulation by governments globally. The study, encompassing 65 countries, identified that 30 of them engaged in various forms of manipulative practices such as paid commentators, trolls, bots, false news sites, and propaganda outlets during the period from June 2016 to May 2017. Of particular concern was the use of such tactics in election campaigns in at least 18 of the surveyed countries, which severely impeded citizens’ capacity to access and engage with factual discourse, essential to making informed decisions when selecting their leaders (Kelly et al. 2017). Moreover, the cross-border flow of data has emerged as a crucial area of concern for both state actors and academics (Kokas 2022, Calzada 2023, Maheshwari & Holpuch 2023, Che 2023). This concern stems from the potential for data to be manipulated for political purposes, and risks associated with the sharing of sensitive data across borders.

Accordingly, TikTok has been facing significant backlash from various governments and organizations globally, leading to multiple bans. The United States, India, the United Kingdom, France (AP News 2023), Denmark (AP News 2023), New Zealand (Craymer 2023), and Taiwan (Chung 2022) have either implemented or are considering a ban on the platform due to concerns about the platform’s ownership by the Chinese company, ByteDance (Fung & Ziady 2023, AP News 2023, Chee 2023). Governments fear that the app could serve as a surveillance tool for the Chinese government, posing a serious threat to user data privacy and national security (Maheshwari & Holpuch 2023, Che 2023). Although TikTok’s data collection practices are similar to those of other tech platforms (Fung 2023), the possibility of the Chinese government accessing this data has prompted several countries to restrict the app’s use within their government departments. For instance, in the United States, TikTok is not allowed to operate within its jurisdiction or collect data from American citizens unless it is sold to an American-based entity (Maheshwari & Holpuch 2023, Che 2023). Nevertheless, China has strongly opposed the ‘forced sale’ of the app (Che 2023, Kokas 2022). This ongoing altercation over the ownership and use of data collected by Tiktok, underscores the power that data holds in shaping public opinion and political discourse.

5 Conclusion

The tech industry’s business model heavily relies on the collection and manipulation of personal data, including that of non-users (Che 2023, Kokas 2022, Lewandowsky & Pomerantsev 2022). However, any form of manipulation carried out by these platforms undermines human
agency and autonomy, both essential components of a democratic society. The assertion that users provide consent for data collection by tech platforms remains highly contentious. As discussed above, the vast majority of users are uninformed about the exact nature of data usage by these entities, the identities of third parties with whom this information is shared, and the purposes for which it is shared. This lack of transparency renders their consent ill-informed, and the credibility of the supposed consent remains questionable. The manipulation of Facebook data during the U.S. elections exemplifies the significant political consequences of these practices. Additionally, Google’s algorithm promoting a civil war narrative illustrates how these platforms have an incentive to manipulate user data to serve their corporate interests, thus highlighting the potential risks associated with granting tech companies access to and control over vast amounts of data.

The foregoing analysis also established that online platforms put users in echo chambers, which creates a polarized social and political environment. These environments are conducive to dissemination of targeted misinformation campaigns, which not only undermine democratic values such as openness, transparency, and free exchange of ideas, but also pose a threat to the legitimacy of the democratic process. Furthermore, the content users consume online significantly influences their attitudes and perceptions, leading them to adopt opinions that align with the prevailing discourse. Thus, the psychological manipulation inherent in these practices raises doubts about the existence of popular sovereignty and generates scepticism and uncertainty regarding the extent to which a political system that relies on such practices can be deemed democratic.

It is evident that in the current economic order humans have been reduced to the status of mere data points, with personal data emerging as a highly prized commodity. The vast troves of data possessed by the tech industry have placed in their hands a formidable weapon that can be deployed to manipulate people’s thought processes and further their own interests. As a result, the tech industry has unequivocally ascended to the position of dominant player, leading to a shift in power from people to the tech industry. This development can be viewed as a coup, where the peoples’ sovereignty is undermined without the need for a complete overthrow of the state (Zuboff 2019).

To safeguard democratic principles and restore people’s sovereignty, it is crucial to bring about a shift in public opinion towards the current data collection processes (Couldry & Mejias 2018). Further, the ongoing practice of continuous data collection must be constrained and limited solely to what is essential for achieving specific purposes such as advancements in the education or health sectors. Most often, the collected data is superfluous and remains dormant within the archives of the tech industry, awaiting potential future exploitation. Moreover, it is crucial to ensure that data utilization is carried out in a transparent, accountable, and human-rights-respecting manner. This entails recognizing the significance of community participation and engagement in data collection processes and promoting data democratization to ensure its equitable benefits across all societal strata. In addition, it is essential for experts to collaborate and create alternative ecosystems that can fulfil the original promise of the digital age: to democratize knowledge and empower individuals (Zuboff 2021).
Thus, to achieve a nuanced and balanced approach towards data collection and utilization practices, it is necessary to safeguard individual rights and dignity while ensuring that data serves the collective good.

References

AP news (2023) Belgium bans TikTok from government phones after US, EU, retrieved April 25, 2023, https://apnews.com/article/tiktok-belgium-china-cybersecurity-b976fe2a56c58996e84c1040dd7f1ad

AP news (2023) Danish defense ministry bans TikTok on employee work phones, retrieved April 25, 2023, from https://apnews.com/article/denmark-tiktok-ban-defense-ministry-c3f434fa46401ea93329e1f5cb132432.

AP news (2023) France bans TikTok, Twitter from government staff phones, retrieved April 25, 2023, from https://apnews.com/article/tiktok-france-ban-cybersecurity-china-4c48564f8be7b86bf44c30969902c293


