- USE THIS NOTEBOOK FOR YOUR PAPER

Navigating the Dilemma: Is Big Data Manipulating Us ?

TikTok, the wildly popular social media platform, has mastered the art of leveraging big data to deliver tailor-made content to its users. However, its meteoric success has also stirred unexpected controversy, with some members of Congress pushing for legislation that could ban TikTok in the U.S. The driving force behind this advocacy is the suspicion that the platform's owner, the Chinese company ByteDance, might be under the influence of the Chinese government, raising concerns about potential data control and manipulation. In *(“Pros and Cons of Banning TikTok,” 2023)*, many politicians expressed the concern that "the Chinese government could use it to control data collection on millions of users. Or control the recommendation algorithm, which could be used for influence operations if they so chose. Or to control software on millions of devices, which gives it opportunity to potentially technically compromise personal devices."

While the debate about TikTok's potential ties to the Chinese government remains inconclusive, the broader issue of big data's capacity to manipulate individuals is not a novel revelation. A decade ago, the infamous Cambridge Analytica scandal exposed the toxic potential of big data in purposefully shaping psychological mindsets. In *(Cambridge Analytica: How Did It Turn Clicks into Votes? | Big Data | The Guardian, n.d.)*, the author illustrated in detail how Cambridge Analytica meticulously targeted diverse groups with highly personalized advertising based on their personality data. This methodical categorization enabled the delivery of tailored messages designed to gradually influence minds and even alter political stances. For instance, "if you're talking to a conscientious person, you talk about the opportunity to succeed and the responsibility that a job gives you. If it's an open person, you talk about the opportunity to grow as a person. Talk to a neurotic person, and you emphasize the security that it gives to my family."*(Cambridge Analytica: How Did It Turn Clicks into Votes? | Big Data | The Guardian, n.d.)* The manipulation's impact was startlingly significant, with the company asserting its ability to identify 'Brexiteers' among voters, donors, politicians, and journalists through a 10-page document titled "Big Data Solutions for the EU Referendum."

In fact, an increasing number of people, especially politicians who are highly attuned to people's beliefs, are recognizing the potent influence of data, giving rise to what is now termed "data politics." In *(Ruppert et al., 2017)*, it's pointed

out that "data is generative of new forms of power relations and politics at different and interconnected scales." The ability to possess, distribute, access, and utilize data is seen as a novel form of power and knowledge, capable of shaping not only decisions but also perceptions of the world. Due to data's power of "capture(ing) but also colonize(ing) minds, souls, bodies, and spaces", governments are particularly vigilant about how data influences their citizens *(Ruppert et al., 2017)* This awareness "is evident in the recent controversies about how Big Data was used in the US election and UK referendum to create personalized political advertising to influence how people voted." The deployment of personalized political advertising aimed at shaping individuals' voting behaviors underscores the pervasive impact of data manipulation on the ideological level.

The influence of big data extends beyond ideology, permeating into the fabric of people's daily lives. Individuals govern their health by willingly becoming subjects of health data. Booking hotels and choosing restaurants are guided by data recommendations. Watching videos and reading articles are filtered through personalized data, shaping individual experiences. However, the potential for misuse looms large, as individuals risk being misled and manipulated into believing almost anything presented with compelling graphics and strategically placed statistics.

Even without intentional manipulation, the pursuit of profits and convenience can lead people into what is known as the "filter bubble," where information is selectively presented, creating a skewed reality. A pertinent example can be found in the Chinese version of TikTok, known as Douyin. Here, discussions have arisen around personalized video feeds and comment sections. Users have come to realize that not only are they exposed to sharply differing opinions in their video feeds, but their comment sections also reflect a distinct bias, showing only perspectives they align with, just like they are introduced to their personalized advertisements in a commodity website. For example, in a video discussing gender-related issues, male and female users saw contrasting comment sections, reinforcing their existing views. The motivation behind it is not ideological brainwash but simply the desire to increase traffic. This mirrors the market-oriented (MO) data applications in commercial corporations like Amazon and Taobao. As demonstrated in *(Zhang & Song, 2022)*, big data significantly enhances a firm's ability to target customers and boost sales. Parallel to commercial advertisements and products, content aligned with specific viewpoints is continuously sent to vulnerable target groups, gradually boosting "sales" (influencing individuals in various aspects of their lives). Users, like customers, find joy in the constant delivery of personalized content that satisfies their preferences, acquiring precise information, fulfilling spiritual needs and

triggering dopamine release. This tailored approach fosters engagement and creates a symbiotic relationship between users and platforms.

However, the consequence is akin to losing autonomy step by step, as individuals find their lives increasingly dictated by big data. From the scheduling of daily activities to shaping beliefs, the pervasive influence of data poses a threat to individual autonomy, raising urgent questions about the responsible and ethical use of this powerful tool.

The hidden peril in this manipulation is our limited understanding of how individual-driven data are distributed, making it challenging to protect ourselves. The main obstacle stems from the formidable walls around the data and algorithms maintained by commercial giants, rendering it difficult for us to gain a comprehensive understanding. Social media data (SMD), in particular, pose challenges due to their "inconsistency, lack of structuring, and rapid growth." *(Osaulenko & Horobets, 2021)* Additionally, there is a dearth of public information about the demographics of the user base and the algorithms themselves. *(Zannettou et al., 2023)*

In this scenario, individuals may fall victim to algorithm bias, resulting in unfair treatment based on their data—ranging from exclusion from pivotal opportunities to facing elevated insurance premiums. Furthermore, the inability to discern which data are objective and which are intentionally manipulated leaves us with an obscured view of the world. This lack of transparency not only undermines our ability to protect ourselves from potential harm but also impedes our capacity to form a reasoned perspective on the realities that surround us.

In the era dominated by big data, a cautionary stance emerges—use the power of data, but don't surrender blind trust, for the very tools that enhance efficiency and accuracy can wield influence over us. This prompts a crucial question: Are we willing to sacrifice some degree of efficiency and accuracy for the sake of preserving our individual "freedom" from algorithmic biases and potential manipulation, or do we accept the trade-off and let data rule, risking the erosion of autonomy in the process? Striking a delicate balance between leveraging the benefits of big data and maintaining a critical awareness of its potential pitfalls becomes imperative in navigating the intricate terrain of data-driven decision-making.

References

Cambridge Analytica: how did it turn clicks into votes? | Big data | The Guardian. (n.d.). Retrieved October 22, 2023, from https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie

Osaulenko, O., & Horobets, O. (2021). Social Media Data in the Big Data Environment. *Scientific Bulletin of the National Academy of Statistics, Accounting and Audit*, *3–4*, 23–31. https://doi.org/10.31767/nasoa.3-4-2021.03

Pros and Cons of Banning TikTok: Data privacy vs. unnecessary worry. (2023). *Congressional Digest*, *102*(2), 30–30. https://proxy.library.upenn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=161461205&site=ehost-live

Ruppert, E., Isin, E., & Bigo, D. (2017). Data politics. *Big Data & Society*, *4*(2), 2053951717717749. https://doi.org/10.1177/2053951717717749

Zannettou, S., Nemeth, O.-N., Ayalon, O., Goetzen, A., Gummadi, K. P., Redmiles, E. M., & Roesner, F. (2023). *Leveraging Rights of Data Subjects for Social Media Analysis: Studying TikTok via Data Donations* (arXiv:2301.04945). arXiv. http://arxiv.org/abs/2301.04945

Zhang, H., & Song, M. (2022). How Big Data Analytics, AI, and Social Media Marketing Research Boost Market Orientation: Companies can use big data analytics, artificial intelligence (AI), and social media marketing research to increase market orientation. *Research Technology Management*, *65*(2), 64–70. https://doi.org/10.1080/08956308.2022.2022907

In [ ]: