



Hey Siri, Where's My Data?

The Use of Social Media Exhaust Data

Kevin Angotti, Ragunath Gunasekaran, & Jennifer Ruiz



Introduction

In the last decade, social media has found its way into every aspect of most people's lives. We now live in the age of the "digital human" [23]. Its reaches can be found in shopping habits, social norms, and even in relationship choices. Companies across the world have found usefulness not only for the captured data from users but their exhaust data as well.

Exhaust Data is defined as "data that are a by-product of a process whose primary purpose is something other than data capture" [11]. For example, for every image that is shared, tweeted, retweeted, or liked, a range of exhaust data is generated: who shared, who viewed, what device was used, the time of day, etc. [24]. Companies around the globe use exhaust data from these sources as a new form of currency, one that provides insights into the lives of those who generate the data [4]. Companies such as Facebook, with over 500 million users, harness the data generated to create new engagement initiatives [18]. Exhaust data provides enough data points to allow for predictive modeling of user activity, both on social media platforms and in real-world activities [15]. The bulk of the discourse on Big Data's downsides today addresses threats to personal information privacy, Loss of privacy can in turn result in other harms such as identity theft and cyberbullying or cyberstalking [2]. Our research examines several areas surrounding exhaust data from social media sources and provide observations for future usage.

Why Is This Data Science?

Data, as a whole, is currently engulfed throughout our everyday lives. Data about data or metadata, as it is known, can be used to help in decision making and automation. Metadata is all around us, and it's presented all the time; we can't shut it off. Nearly every device uses it or generates it, or both [6]. What is the cost of the amount of data we produce with the technology of today? Using online search engines, websites, and social media platforms, we create a byproduct from that data called exhaust data. This type of data can be used for many different things; instead, it is for the benefit of the user or a company, which means that it can either be used for the good or the bad. Determining that will cause us to look at different aspects of exhaust data.

How we use social media today allows for mass amounts of exhaust data to be retrieved and used for the benefit of others. Each social media platform enables users to use the platform in different ways, but they all filter the same thing, data. We use these platforms for friendships, keeping in touch, pictures, or simply just putting your thoughts out for all to see. Each of these platforms allows companies to gather data, and this collection occurs from the result of actions you the user is doing. exhaust data is information-seeking behavior, which can be used to infer people's needs, desires, or intention [3]. You provide access to this data by way of their user agreements. From what you as the user do determines how much data can be retrieved. This is the exhaust data collected, and while they are designed to be free services that help facilitate connectivity among its users, they have become more like platforms that deploy strategies to get its users to produce more and more data. The data collected will be analyzed for profits. Moreover, the results from such platforms are sold to advertisers and marketers [16].

Objectives

1. Analyze the current uses of Exhaust Data generated from social media sources
2. Discuss current ethical concerns and provide recommendations
3. Suggest future uses of Exhaust Data from Social Media



Deliverables

Our primary objective to identify current uses of exhaust data generated from social media sites. Research regarding this usage, as well as the ethics surrounding it have not been investigated properly. Providing solutions regarding the ethical and management challenges of usage for incremental improvement.

Data Exhaust can yield some important insights for businesses.

1. Improve Market Research
2. Help Shore Up Cybersecurity
3. Data Exhaust from Manufacturers
4. Data Exhaust is a Revenue Opportunity
5. Racing to Win with Data
6. Responsible (and legal) use of Data Exhaust



Conclusion

One of the primary jobs of a data scientist is to use data to solve complex problems in the real world. The data generated from information posted on Social Media sites provides a wealth of data points to solve problems. These could include everything from crime prevention, consumer behavior, targeted marketing, and so many others.

In summary, if IOT is the New internet and Data is the new oil, don't let IOT Data Exhaust go to waste. It just might contain the insights to fuel the competitive advantage that helps to win the future race in your market. But first it requires building the capabilities to efficiently capture and analyze this Data, and revamping innovation process that leverage the power of IOT-generated Data into new products, services, and business models that separate from the pack.

Acknowledgements

-Thank you to Twitter, Instagram, and all the other apps that Facebook owns for the great idea

-Dr. Shankar Parajulee for his support and knowledge

-Coffee for keeping us awake during grad school

-Countless data scientists around the world who further our knowledge and drive our curiosity each day



Literature Review

[1] Carson, A. L. 1. acarson@iastate. ed., & Ou, C. carol. ou@unlv. ed. (2019). Notes on Operations Metadata Revisited: Updating Metadata Profiles and Practices in a Vendor-Hosted Repository. Library Resources & Technical Services, 63(4), 196–205.

[2] Markus, M. L. (2015). New games, new rules, new scoreboards: the potential consequences of big data. *Journal of Information Technology (Palgrave Macmillan)*, 30(1), 58–59. <https://doi.org.ezproxy.bellevue.edu/10.1057/jit.2014.28>

[3] George, G., Haas, M., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, (2). Retrieved from <https://search.ebscohost.com.ezproxy.bellevue.edu/login.aspx?direct=true&db=edsbig&AN=edsbig.A367721667&site=eds-live>

[4] Harford, T. (2014). Big data: A big mistake?. *Significance*, 11: 14-19. Retrieved from <https://doi.org/10.1111/j.1740-9713.2014.00778.x>

[5] Gupta, M., George, J.F. (2016). Toward the development of a big data analytics capability. *Inf Manag* 53(8):1049–1064.