Ethics of Online Data Collection in Advertising

Jackie Diep

CSC 309

University of Southern Mississippi

Biloxi, Mississippi

Jackie.Diep@usm.edu

***Abstract*—This electronic document briefly analyzes some methods used to collect personal data online and the ethics of collecting user data to promote the efficiency of online advertising. Specific examples such as the Cambridge Analytica scandal and hypothetical examples will be used to help the reader better understand the possible ramifications of data collection, and a short overview of some of the advantages and disadvantages of data collection will be discussed.**

***Index Terms*—Advertising, Cambridge Analytica, Data Collection, Ethicality**

# Introduction

Many popular online services including, but not limited to, mobile applications and social networking websites collect in depth user data to be sold to advertising companies in the future. Oftentimes, the methods used to collect data are underhanded and intentionally misleading.

The most popular example of questionable online data collection was Cambridge Analytica, a political firm that collected data from over 87 million Facebook users using software that not only directly lied to over 300,000 users who consented to data collection, but also collected information from friends of the affected persons without consent [1].

This document will describe some of the popular methods used to collect user information, discuss the acquisition ethicality of data collection in advertising, and assess the advantages and disadvantages of the practice.

# Methods of Online Data Collection

Four of the most popular types of data collection are referred to as Clickstream, Search, Purchase, and Profile data [2]. Clickstream data, which tracks browser history, and Purchase data collection, which tracks shopping cart and purchase history, will directly send cookies to a personal computer to follow and record the user’s movement through the related websites to target the involved user’s specific tastes. Search data collection is a more general form of data collection used by search engines like Google to determine which terms over millions of uses will generally lead to certain products. The top search result slots will then be sold to companies involved in producing these products. Finally, Profile data collection is a more targeted form of data collection, like Clickstream and Purchase data collection, but it does not send anything to the user’s computer. Instead, it will pull data directly from social networking platforms to customize advertising for the individual user within that platform [2].

As for the previously mentioned Cambridge Analytica - Facebook issue, a “data-harvesting quiz app” acquired consent for data collection by falsely claiming to require the information for academic reasons. Consenting users would allow their data to be recorded throughout the quiz, followed by the collection of the profile data of every member of their friends list [3].

Often, online platforms asking for consent for data collection may explain half of the story. An everyday example of this is when mobile applications request for location permission to match individuals with other users. What goes unsaid is that advertisers may also purchase large amounts of a user’s location data to learn what stores the user frequents, allowing them to be targeted with highly efficient advertising [4]. Mobile social networking and dating applications often ask for far more than simple location data, and who is to say that Tinder, for example, should be allowed to change the course of peoples’ lives with access to collected contact information and mutual friends that a user may not have specifically consented to?

# Ethicality of Online Data Collection

To most people, the most important part of online data collection is consent; however, how much and how specific should the consent be? According to the last example including mobile applications, partial consent should be enough, but the senior vice president of public affairs in the U.S. for Ipsos, a market research company, complete consent must be given towards every possible usage of collected data [3].

In an ideal world, users should choose to stay completely informed about their data; however, it is also true that most users do not. Many users skip passed important explanations like Terms and Conditions due to either time or mental constraints. It is likely that they would also skip passed any in depth explanation of data collection, and how a company or group of companies may use personal information for the same reasons.

Clearly, users are not in a position to manage data collection, so the responsibility must be on the advertisers and the limitations set upon them by themselves or the government. As stated by Julia Clark, senior vice president of public affairs in the U.S. for Ipsos, a market research company, “A core tenet of market research ethics is transparency [3].”

Ethically, data collection is neutral, and the conversation revolves around the usage of the data afterwards. Great good can come from a well placed advertisement, or “matching” two people on a dating application, but great harm can come from improperly secured or utilized information. If users do not have the time or ability to manage data collection, then marketers and advertisers must be held responsible for the positive *and* negative ramifications of data collection.

# Advantages and Disadvantages of Online Data Collection

For the user, data collection creates better advertising that benefits them in ways such as overall satisfaction and shopping efficiency. For companies, more user engagement and increases in sales and revenue [5]. Online data collection in advertising is a symbiotic relationship meant to benefit both the user and the advertisers in exchange for a small loss in privacy. The conundrum comes from the actual worth of the relationship versus the loss of a basic human right of going online.

In favor of online data collection in advertising, at its best, better advertising can drastically increase the quality of an online experience. Because advertisers pay for harvested data, websites that have become a major part of everyday life, such as Youtube and Facebook are able to stay free of charge, and mobile apps such as games and instrument tuners can similarly be free of charge or lower in price than would be the case otherwise. The world would be an altogether different place if what we consider major parts of the internet like search engines were locked behind paywalls.

On the other hand, collected data is a huge security risk that could be a major factor in the downturn of a user’s life. If a database were to be breached or if any private information were to become public, malicious individuals like stalkers, hackers, or “doxers” could utilize this information to harm users. Nothing is ever truly secure, and every new database with personal information is another risk to the user.

# Conclusion

There are numerous and varied methods that personal data can be collected and used in advertising including, but not limited to, directly tracking a user’s browser and purchase history, tracking millions of searches around the world on a popular search engine, and pulling information directly from a user’s profile page.

Ethically, data collection is a relatively neutral topic that revolves more around the actual usage of the data, rather than the act of collecting the data itself.

There is a symbiotic relationship between user and advertiser in terms of data collection because better advertising brings strong benefits to both parties; however, the loss of privacy online can and has been used to great negative effect by malicious individuals.

The advertisers and marketers are the ones primarily responsible for keeping data collection ethical and collected data secure because the average user has limitations that make them unable to, but users should still understand that underlying agreement that their personal information may be sold in exchange for access to a service or application.

# References

1. P. Bump, *Everything you need to know about the Cambridge Analytica-Facebook debacle*, Washington Post, Mar. 2018. Accessed on: Sept. 26, 2019. [Online]. Available: <https://www.washingtonpost.com/news/politics/wp/2018/03/19/everything-you-need-to-know-about-the-cambridge-analytica-facebook-debacle/>
2. C. Pumphrey, *How do advertisers show me custom ads?*, HowStuffWorks, Sept. 2012. Accessed on: Sept. 26, 2019. [Online]. Available: https://computer.howstuffworks.com/advertiser-custom-ads1.htm
3. S. Steimer, *The Murky Ethics of Data Gathering in a Post-Cambridge Analytica World*, American Marketing Association, May 2018. Accessed on: Sept. 26, 2019. [Online]. Available: https://www.ama.org/marketing-news/the-murky-ethics-of-data-gathering-in-a-post-cambridge-analytica-world/
4. L. Matsakis, *The WIRED Guide to Your Personal Data (and Who Is Using It)*, WIRED, Feb. 2019. Accessed on: Sept. 26, 2019. [Online]. Available: https://www.wired.com/story/wired-guide-personal-data-collection/
5. LOTAME, *What is Behavioral Targeting?*, LOTAME, Sept. 2018. Accessed on: Sept. 26, 2019. [Online]. Available: https://www.lotame.com/what-is-behavioral-targeting/