Christopher Gomes

Professor Enos

ENGW 3302

September 23, 2018

The Ownership Dilemma

Google, Facebook, Twitter, Amazon, Microsoft, Salesforce, Netflix, and Spotify all have one thing in common: they use their users' data to improve their products and create revenue. The public's lack of understanding about how these companies take advantage of their users' data has allowed these companies to continue these practices for years. Fortunately, recent privacy breaches at major companies and awareness campaigns by privacy advocates have finally brought the discussion out in the open. Society is now in a better position to determine the best way to correct company use of user data and find the best solution that is fair to both parties. The relationship between companies and their users' data can be viewed in two ways: as a transaction where the users give up some rights to the information for the use of a service or as a liability to the company where they are required to treat the data with the acknowledgement that the users still own it.

The business plan of using user data for profit goes back to the early 2000s when internet companies were first created. Since internet services were not as pervasive in society during this time, it was difficult for internet users to find value in paying for services online, which left companies to find other ways to fund the services and tools that they wanted to offer. Providing advertisements, especially tailored ones, became the go to method for companies to create revenue streams without charging their users, especially since it allowed the companies to build a much larger user base compared to if they charged a fee. In the last few years, data collection began to increase even more due to the rise of machine learning algorithms. Basic matching and filtering of user preferences grew into a race to predict a wide array of user behaviors and preferences. These algorithms require enormous amounts of data, which incentivized companies to increase the amount of information that they collect. User information is now powering extremely robust and influential algorithms and has become the foundation for the data science boom that has been occurring for the last few years. Without this data, data science and machine learning would not have the prevalence that they have today.

For the data science community, finding the correct balance for data collection and use is a crucial part of the community's survival in the future. If legislators decide that data collection between users and companies is a transaction where users get access to the service for some decreased ownership of the information they put on the platform, then companies would be free to analyze and monetize this data without much say from the user. This perspective is great in continuing and improving machine learning research. This research has already led to, among other things, the personalization of services, such as Spotify, which have revitalized industries and created jobs, and health care advances, such as the better detection of tumors in MRI scans. Unfortunately, this way of treating user data also poses the risk of abusing users' privacy. A great example of this is when "Facebook gave unfettered and unauthorized access to personally identifiable information (PII) of more than 87 million unsuspecting Facebook users to the data firm Cambridge Analytica", which included not only the people that used the third-party app on Facebook, but also the friends of those users [1]. This is only one example of how bad actors, in both the private and public sector, could be able to use user data to misinform and take advantage of unsuspecting people, which is especially dangerous due to the ability of machine learning algorithms to create extremely accurate predictions and mappings of users. This perspective is very similar to what is enforced by current legislation in the USA and in most countries, but there is a movement to rethink the relationship between companies and their users' data.

A new perspective places the protection and ownership of a user's data as a basic right that the user is entitled to and that a company cannot take away. In this light, a company would never be able to do what Facebook did, because the user would always keep full control over their data. In this way, companies that hold user information would actually be taking on a liability and would not be able to do anything without extremely detailed communication and consent from its users. The legality that would be required would significantly slow down advances in creating more efficient and useful machine learning algorithms or products but would keep companies from having the ability to misuse their users' data. It was for reasons like this one that some data science professionals were hesitant to fully support the recent GDPR legislation, which strictly regulates how companies can use user data in Europe.

In response to Facebook's Cambridge Analytica incident, Jim Isaak and Mina J. Hanna published "User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection" in IEEE's Computer Society publication *Computer*. This article contains an explanation of what occurred, a list of what the authors' feel should be included in future privacy and data protection legislation, and a description of the current legislation on the topic. The authors' try to use this article to motivate readers to want to protect the privacy of their data.

Isaak and Hanna target this article to users of technology platforms, which can include people with and without a technical background. To accomplish this goal, the authors avoid explaining detailed technical aspects and instead focus on the bigger picture. For example, in their explanation of the Cambridge Analytica incident, they do not explain how Facebook's APIs function and focus strictly on the overarching timeline of Cambridge Analytica's steps. In addition, they explain words and phrases that non-technical users may not understand. An example is when they write “a personality quiz on Amazon's Mechanical Turk platform and Qualtrics, a survey platform” [1]. Since, parts of the audience would not be familiar with Amazon’s Mechanical Turk platform and Qualtrics, they provide a very basic explanation of what it is without getting very detailed. Overall, the article has a wider potential audience compared to other written resources about the same topic.

With this audience in mind, the article’s goal is to motivate them to care about the issue of user privacy and fight for more stringent laws. The layout of the article fuels this call to action. At the start, the focus is to give validity to the argument that something should be done regarding the current state of user privacy, and it uses the Cambridge Analytica problem as the specific example to support the argument. Then, the authors outline what they feel the user should want in terms of data privacy, in a clear organized list with bullets, such as “Users must easily be able to delete personally identifiable data from any site, cloud service, or collection device” [1]. Lastly, after providing this guideline they then present the reader with the current state of proposed legislations. In these three parts or steps, the authors provide the background information that the readers need in order to understand and assess the problem, and then they provide the readers with direction on what to change to solve the problem, which motivates readers to not only want to bring change but to also know how to. In this way, these new defenders of privacy can actually make a difference, instead of just being upset with the issue.

Moreover, the authors attempt to motivate the audience by appealing to their emotional tie to the ethical standards that citizens of democratic nations expect from their government. An example being that mass data collection and analytics are “disruptive forces [that] have a tangible influence on citizens' rights such as statutory rights—due process, equal representation before the law, the right to appeal, and trial by jury—and constitutional rights like freedom of expression, voting, and non-discrimination" [1]. The authors know that the rights that citizens have in the judicial system are valuable and are something that the reader will feel very strongly about. Therefore, if the authors can convince the reader that they should have a similar feeling about data privacy, the reader will gain the same emotional tie and become more motivated to act.

Lastly, the authors present their persona by creating an urgent and blunt tone, which makes the audience feel that the issue is important and that the authors are being straight-forward and honest. Right from the beginning, the authors write that "it has never been more imperative to have an open discussion about the proliferation of technology in our lives and how it will affect our privacy rights and our security on both personal and national levels" [1]. In this sentence, the authors present the reader with the idea that this topic is important and that now is the time to act. Specifically, the use of the word "imperative" and the acknowledgement of technology's prevalence in society creates a sense of importance and urgency.

The authors effectively continue their tone throughout the article into the section outlining guidelines for what they feel should be in future legislation. Specifically, they use firm language, such as the use of the word "must" in "Each website and application must disclose any ongoing content placed on the user's device, as well as the uses of that content" [1]. Even though the authors call these bullets principles, they are written to be far more than that. The wording places these points as commands that corporations need to be following and not just suggestions to consider, which coincides with the purpose of the article. In the end, the article was not written to just suggest the exploration of this issue. It was written to demand that urgent actions be made to correct it.

Technology has presented a completely new battle ground that has made society rethink how rules and laws have been made, regarding the privacy of user data. Therefore, as technology consumes more areas of people’s lives, it is important that governments and professionals are confident that interactions over this new frontier are fair, especially since technology has the power to exponentially increase both the positive and negative resulting effects.

Works Cited

[1] J. Isaak and M. J. Hanna, “User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection,” *Computer,* vol. 51, no. 8, pp. 56–59, Aug. 2018.