Cases Analysis on Data

The industry around data grosses billions of dollars annually. Data on people around the world is used in numerous ways such as research, marketing, surveillance and more. In 2018 the General Data Protection Regulation or GDPR for short was put into effect across the European Nations. This was done to help protect individuals' privacy and protect their personal data from misuse and improper handling. This gave EU citizens more control over data such as names, pictures, biometric data as well as other data that can be used to identify an individual. Till then users did not have much control over how their data was used once it became public. The new regulations also set forth standards on how controllers and processors of data were responsible for any misuse or poor handling. Controllers and processors of data are help to fairness and transparency standards. This was a monumental achievement for the citizens of the EU, but over seas in the United States such federal regulations do not exist or are not as protective. In this Case Analysis I will argue that contractarianism shows us that the United States should follow Europe's lead because users in the United States may be unaware of how their data is being used or that it has been compromised.

In a research paper written by Michael Zimmer titled "But the Data is Already Public" (Zimmer) we can see the ethical conflicts when data is either used unintentionally or used to partially reidentify users of a data set. In this paper it explains how datasets, collected off Facebook accounts from a specific college were publicly released. The team that released these datasets made attempts to conceal the identities of the users to protect privacy. However, despite these efforts, the users were reidentified with the anonymous data provided. Part of the efforts made by the research team who released the data created a TOU, or "Terms of Use" agreement that specifically prohibited the use of this data to reidentify the people. Specific details of these datasets such as a list of college majors were used to identify the specific college these records came from. Granted PII was not part of these datasets, the

students were still partially identified. The research team removed the data from public view shortly after reidentification took place. The team did not face any repercussions other than scrutiny. Under the GDPR, this would be considered a breach and the research team that released the data would have been help responsible for its misuse. Under the GDPR it is the responsibility of the controllers and processors to maintain ensure that data is used properly and that the users who provided the data fully understand and are properly notified how the data will be used with no vagueness.

"In July of 2021, European regulators in Luxembourg fined Amazon Europe a whopping $877m fine for data breaches and failing to comply with general data processing principles under GDPR." (Clark) Fines like the one Amazon received in July of 2021 should give controllers and processors of data a larger incentive to better protect the privacy and security of user data. Not all companies face this level of repercussions for data that is improperly handled, but if an instance occurs, the companies responsible will be held accountable for their actions or lack there of. Not only does this apply to data in Europe, but any business that handles data of Europe citizens even outside its borders. "This means the reach of the legislation extends further than the borders of Europe itself, as international organizations based outside the region but with activity on 'European soil' will still need to comply." (Palmer)

Using the ethics of Contractarianism, specifically under a "Veil of Ignorance" a concept proposed by John Rawls can show us how the United States should adopt these same practices because it is only fair for all of society, regardless of where you stand. Would the research team make a better "Best effort" at concealing identifiable information in the datasets? Perhaps the data would have never been made public in the first place. Not knowing what side you're on, the side of the controllers and processors, or the side of the end user could help us adopt rules that are fair for everyone. Its only right that users fully understand what data will be collected, how it will be used and if a breach occurs, that they are notified immediately to take proper actions. This also falls under "justice as fairness" (Rawls) when proper responsibility and actions are taken against the misuse of data to create stronger

incentives for protecting data fully. Users provide data, and that data is indeed public, but that does not mean that it can just be used in any way people want. It's the responsibility of controllers and processors to make better efforts in informing the public exactly how their data is used. This would create a strong balance between consent and use of data.

These rules do not only apply to commercial business, educational institutions, and research teams. Elizabeth Buchanan writes in her paper titled "Considering the ethics of big data research: A case of Twitter and ISIS/ISIL" about how a model called "Iterative Vertex Clustering and Classification" or IVCC for short was used to identify ISIS and ISIL sympathizers and community members on twitter. "This method enables greater detection of specific individuals and groups in large data sets, with its enhanced capabilities to identify and represent following, mention, and hashtag ties." (Buchanan) The ethical concern behind this model is that even though the data is public and produced by the individuals themselves it makes it difficult to "protect individual liberties" (Buchanan). A good example of this is when whistleblower Edward Snowden reveled that the National Security Agency had a history of unconventional surveillance methods that created many ethical concerns and how they can affect the constitutional rights.

When using social media platforms, users must agree to a Terms of Service agreement. Terms of Service agreements set forth by platforms like Twitter, Facebook, Google etc. have a history of being lengthy, hard to understand and are generally not read before accepted. Generally, a TOS or TOU agreement covers use of content, ownership of content and the platforms access to personal information. There are many factors in why people do not read or fully understand TOS agreements before accepting them. Generally, the length and complexity are an issue. Some assume safety in the fact that millions of other users also use the platform with no concern. Sometimes people find that the benefit of the platform's services outweigh their privacy concerns. The problem in this lies in how data is collected and used after agreed to and the constant changing of these agreements. "One may implicitly

agree to one's data sources being used for marketing purposes while that same person would not want their data used in intelligence gathering. But big data research does not necessarily provide us with the opportunity to consent to either use, regardless of the intent." (Buchanan) This creates an unfair ethical dilemma for the users of these platforms. Companies like Twitter are less than transparent about how data of its users is collected and used in general. Though the GDPR does not specifically mention Terms of use, service, or conditions in its regulations, it does require that data is used and collected for purposes specified and that it is used only in the manner made transparent to the public and the end user. TOS agreements may not be mentioned, but privacy policies are a must, and these can overlap one another. "A Privacy Policy is required by the GDPR and other privacy laws in order to protect users and ensure proper business practices by website owners and app developers." (Bass) Furthermore, if that data is used in a way not specifically intended, it is the responsibility of the company to notify its users of the "breach" in a timely manner.

How does this relate to the theory of contractarianism and the "Veil of Ignorance". It all relates back to fairness for all. Users of these platforms should understand that they have no reasonable expectation to privacy, and many are aware that their data can be mined and used in ways not intended, but they are not given much choice in the matter of how the data is used. It is important and only fair that users understand how their data is being collected and used. The GDPR helps EU citizens in this aspect and adopting similar regulations in the United States could not only benefit the end user of these platforms but also protect companies like twitter by establishing trust with its users through its transparency by creating baselines for user privacy.

Using contractarianism and Rawls "Veil of ignorance" I find it fair not only to citizens of the United States but also companies to benefit from regulations like the GDPR. Users will benefit from a fair level of transparency and not need to rely only on an ambiguous "Best effort" to protecting their personal data. Additionally, this can benefit businesses in many ways as well. Businesses can earn the

trust of their users by providing exactly how data will be used and process and when misused or breached can take the proper actions to notify users in a timely manner. It may be difficult for companies to adopt principles like this, but many have already made conscious decisions to shift most of their focus on end user privacy. A set of federal regulations in the United States can protect companies by giving them a clear understanding of the guidelines they need to meet so they are not subjected to hefty fines and repercussions from the poor handling of data. It will give business a better idea of how to be compliant and avoid massive financial loss, as well as the loss of users on its platform brought on by user mistrust. This will not be an easy effort. Many businesses may face challenges constructing and implementing new guidelines to comply with privacy, but the benefits of this outweigh the cost. There will be users who may leave platforms and withdraw consent. This may seem problematic if users of these platforms simply left, as many companies rely on the mass number of users and data to generate revenue. It will not be an easy effort to address all these concerns, but it could improve the relationship between end users and businesses. With a certain level of mandatory transparency, many might feel safer in using these platforms who previously left.

Works Cited

Bass, Ross. “Will the GDPR Affect Your Terms and Conditions Agreement?” *TermsFeed*, 15 June 2018, www.termsfeed.com/blog/gdpr-terms-conditions/. Accessed 12 Feb. 2023.

Buchanan, Elizabeth. “Considering the Ethics of Big Data Research: A Case of Twitter and ISIS/ISIL.” *PLOS ONE*, vol. 12, no. 12, 1 Dec. 2017, p. e0187155, https://doi.org/10.1371/journal.pone.0187155.

Clark, Kendra. “Google’s $400m Penalty and Impact of the 5 Heftiest Data Privacy Fines on 2023 Ad Plans.” *The Drum*, 15 Nov. 2022, www.thedrum.com/news/2022/11/15/googles-400m-penalty-the-impact-the-5-heftiest-data-privacy-fines-2023-ad-plans.

Palmer, Danny. “What Is GDPR? Everything You Need to Know about the New General Data Protection Regulations.” *ZDNet*, ZDNet, 17 May 2019, www.zdnet.com/article/gdpr-an-executive-guide-to-what-you-need-to-know/.

Zimmer, Michael. ““But the Data Is Already Public”: On the Ethics of Research in Facebook.” *Ethics and Information Technology*, vol. 12, no. 4, 4 June 2010, pp. 313–325, link.springer.com/content/pdf/10.1007%2Fs10676-010-9227-5.pdf, https://doi.org/10.1007/s10676-010-9227-5.