Code of Ethics: Social Media Platforms and Data

Jeremy Cruz
u1275138@utah.edu

April, 5, 2021

1 Preamble

This ethical code and standards includes a recommended guideline for all websites that have a user base from which the entity providing a service has access to and/or collects data on their users (social media). In the interest of protecting civil rights, the principle behind all of the suggested regulations and restrictions in this code is protecting individuals' ability and capacity to consent to usage of their data.

We have reached a point in society where, in non-authoritarian governed states, technology has grown faster than the law. As such, the ability to be a participating member of society and receive equal social opportunities and fulfillment requires participation in social media. However, because technology has evolved faster than the law, existing entities are able to take advantage of their users in ways that are unethical.

The importance of valuing the individual's ability to consent to the usage of their data is based upon the idea of self agency. In states where individuals have free and equal protection under the law, and privacy is considered a right, citizens should have the self agency to protect and govern their right to privacy when using websites online.

This code of ethics is intended to guide lawmakers, legislators, and government agencies in the United States of America in order to advise them to best protect the people who they represent. While it is intended to be implemented in the United States of America, it would be applicable in any sovereign state that equally values the freedom of privacy of its citizens. With regards to this code of ethics, it is encouraged and recommended to enact laws that deploy and enforce this code. It is also the recommendation that the penalty for failing to comply with these regulations not be financial and instead restrict the website and company's ability to deploy their software in the applicable state until compliance is met.

This code of ethics outlines methodologies that do not endeavor to inhibit the current behaviors of social media companies, but rather ensure that they are being performed in an ethical manner. This mandates that users have control over their data and how it is being used, but most importantly that they are able to provide Informed Consent in this process.

2 Definitions

As used in this Code, the following terms shall be used to refer to the following definitions:

Code - used to refer to the entirety of the contents of this document.

Entity (Entities) - used to refer to any website or affiliated organization that owns that website, including all of its affiliated subsidiaries or parent companies, social media or otherwise, that collects data on individual users of their website.

Usage Data - used to refer to all data collected on individual users of a website, whether associated with a particular user or not. This includes what they do on a website, biographical information, whether voluntarily entered or collected automatically.

User(s) - used to refer to any individual using a website, whether they have created an account or not.

Service(s) - used to refer to the website provided by the Entity or any services provided by the Entity related to their website

3 Principles

The entirety of this Code rests on the concept of protecting two fundamental civil rights: that of privacy and that of autonomy (and with that autonomy, Informed Consent).

Privacy - Not only is privacy a civil right, privacy is also a necessity for people to remain autonomous. [1] With access to all your data, and enough data on a person, companies can create an in depth psychological profile on you and can influence your behaviors such as buying a product or voting a certain way. This is not inherently wrong, but it is wrong to do if the individual has not consented to being a participant or subject in such a process and is a violation of their privacy otherwise.

Informed Consent - Ownership is a moral right. [2] The data that users volunteer (name, address, phone-number, etc) still belong to the users and are not owned by the social media platform. This data can be 'lent out' to companies but only with explicit informed consent from the User. The U.S. Department of Health Human Services' Office for Human Research Protections provides regulations on Informed Consent.[3] Research performed, whether in a public or private institution, follow these guidelines. Companies that collect and use User Usage Data should be held to the same standards as research institutions. The primary reason they currently are not is because nothing requires them to disclose what

[1] "Privacy and the Threat to Self" - Michael P. Lynch

[2] "What is Ownership"

[3] See requirements here: https://www.hhs.gov/ohrp/regulations-and-policy/guidance/faq/informed-consent/index.html

they are doing, so they are able to operate without being under scrutiny, which makes the need to require a standard of Informed Consent all the more important.[4]

4 Tenants

4.a - Consent must not be required or mandated for Users to provide the Entity with any Usage Data to order to use the Service [5]

4.a.i - The User should be opted out of providing Usage Data upon creating an account as the default behavior

4.a.ii - When retroactively implementing the Code, all accounts should be opted out as a default behavior

4.a.iii - Providing compensation to the User in exchange for their Usage Data as an incentive to provide consent would not be considered unethical.

4.b - The Entity should not track behavior outside of the behavior exhibited by the User when using the Entity itself

4.b.i - Keeping track of ordinary website traffic into the Entity website[6], as well as outbound traffic regarding which links the User clicks on when exiting the Entity's website, is the only data outside of the Entity that should be able to be recorded ethically by the Entity.

4.b.ii - If the Entity has an application that is installed onto a User's device, or a tab using the Entity's website is open within a web browser, the Entity must not record any data outside of the Entity while the User is not actively using the Entity.

4.b.iii - The Entity must not use the failure of the device manufacturer, browser, etc. to limit cross-application access as grounds to collect more Usage Data than they ought to.

[4]Further reading exploring both sides of the argument can be found in this analysis: https://www.econstor.eu/bitstream/10419/180107/1/f-21375-full-text-Bergemann-2018-Consent$Paradox-v2.pdf$

[5]Traditionally, compliance with submitting Usage Data is included in the Terms of Service or Terms and Conditions for a website and furthermore accepting the terms is a requirement in order to create an account. Ethically, the User is given no choice in their compliance with the Entity's policy and it robs them of their agency to choose and protect their data. It is not even disclosed to them what exactly is being used and for what purposes, but the User must agree to allow their Usage Data to be used by the Entity in order use the Service at all. The most ethical course of action would be to have opting in to Usage Data be completely separate from the terms when registering an account, and not required.

[6]This is called referral traffic and it is customary to record. Outbound traffic is also customary to record.

4.b.iv - The Entity should not monitor, interpret, record, or use voice or speech picked up by the User's device unless they are specifically using a speech feature within the Entity. Services that are responsive to voice commands ought to only begin when prompted by specific keywords, and anything listened to outside of the context of that prompt out to be immediately discarded.

4.c - Users should be able to explicitly opt in or out of each and every purpose their data is being used for

4.c.i - Minors can not consent to their data being used for any purpose per the age specified in the origin country of the User. If the Entity uses the data for purposes and studies that would ordinarily require the minor's consent for research, the Entity ought to be liable for instances in which the User, if a minor, has falsified their age due to lack of proper identity verification being performed by the Entity.

4.c.ii - Each individual use that the Entity may perform with the Usage Data must be listed in such a way that the User can read through all of them clearly and opt in or out of each one specifically. Examples include but are not limited to: A/B testing, predictive modeling for specific purposes, facial recognition, research, etc.[7]

4.d - The Entity must provide detailed information about each individual use of the User's Usage Data in a way that is clearly legible and can easily be understood.

4.d.i - Each individual study, project, etc. must be detailed for the User to consent into or opt out of. When appropriate, the Entity can link to another page to provide more details so that the User has a comprehensive understanding of what they might consent to by opting into that option.

4.d.ii - The details of the study or project cannot be abstracted or hidden in any way. The Entity ought to use clear language, without the use of technological jargon, and the purpose and goals must be clear as well as any risks that might exist by consenting to being part of the research.

4.d.iii - The Entity cannot make broad, vague, or blanket statements to encompass a wide variety of issues and bundle them into one. For example, saying "to understand behavioral patterns" is not sufficient. An example of what would be sufficient is, "I consent to my data being used in an algorithm to determine the success, likelihood, or probability that a relationship might succeed or fail based on the data

[7] An example of how specific these projects need to be listed by is this: Facebook has released information claiming that they can predict with high accuracy whether or not an individual's relationship will succeed or not based on the content shared by the User. Similarly, Facebook also can predict with high accuracy when an individual will leave their job. These are examples of two specific, distinct, separate research projects that should require individual consent for each usage.

available provided through the User's activity on the Service."

4.e - The User must be able to request the removal of all of their data at any time, and the Entity must comply with this request within a short period of time (30 days as a maximum).

4.e.i - The Entity must design their algorithms such that data obtained from one user can be identified with that user and subsequently removed. If there is no way to design the algorithm in this way, this **must** be communicated to the User as a risk when they are opting into this usage purpose.

4.e.ii - The capacity and regulations for removal ought to, at a minimum, meet the bar for GDPR standards for removal.[8]

4.f. - The User must be able to see what of their Usage Data has been collected, what it is been used for in the past, and what it is being used for presently at any given time

4.f.i - The User must be able to log into their account and view a specific record of purposes of which they have been a part of. Examples include particular projects, A/B testing, etc. They must also be able to download a copy of all of their data that has been used in a digestible format.

4.f.ii - This data must be digestible, interpretable, and decipherable for the user. The Entity must not hide this behind a wall of industry specific jargon that the layperson would not be able to understand what their Usage Data is being used for.

4.f.iii - If the User has previously requested that the Entity remove their Usage Data, the User must be able to verify through this tenant of being able to see what their data is and has been used for, that the deletion occurred successfully.

4.f.iv - If for any reason the data cannot be provided, the Entity must provide a list of what data could not be given and why. This risk must be communicated in a clearly visible way at the time the User opts in to allowing their data to be used for this purpose.

4.g - The Entity must obtain explicit consent from the User to share the Usage Data with any third parties, including governments, subsidiaries or partners

4.g.i - Each individual study, project, purpose, etc. that the third party will use the Usage Data for must be detailed for the User to, following all of the same requirements set forth in Tenant 4c and 4d. All standards that apply to the Entity also apply

[8]GDPR is the General Data Protection Regulation and is an EU law on data protection and privacy that went into place in 2018. This regulation requires that corporate entities delete all data pertaining to a EU citizen, if that citizen requests it.

to third parties that might obtain the Usage Data through the Entity

4.g.ii - The details of the study, project, or purpose, as well as the details of which third party organization will receive this data cannot be abstracted or hidden in any way.

4.g.iii - The details regarding the third party company cannot be obscured, hidden, deceiving, or intentionally hard to find. The most notable name that would be recognized by the layperson must be used. If publicly traded, the symbol must be provided as well. If the most recognizeable name associated with the third party cannot be easily determined, the Entity must provide the User with a hierarchy of company names such that they can properly recognize and be informed with whom they are sharing their Usage Data with.[9]

4.g.iv - A User consenting to allow third party usage of the Usage Data does not imply or guarantee that the user would consent to the third party giving the Usage Data to yet another party. Each time the Usage Data is transferred to another organization, consent must be re-obtained from the User.

4.h - The Entity must not profit upon, sell for money or other transactional benefits, or monetize in any way the Usage Data collected without specific, separate, explicit consent for the usage to be profited upon. [10]

4.h.i - If a specific project will have monetary benefits for the Entity, the user should be given a separate option in conjunction with consenting to that purpose to separately consent to the Entity being able to receive any monetary benefit for that usage.

4.h.ii - This tenant applies to all third parties that might receive the Usage Data as well. If something was consented to and at the time did not have a potential monetary benefit, but one is being added retroactively, consent to allow the Entity to monetize or profit off that data cannot be assumed.

4.h.ii - Consent to allowing the Usage Data to be monetized or sold can never be assumed based off previous choices of the User.

[9] The intention behind this Tenant is to prevent Entities from being disingenuous with who the data is going towards by using the name of a lesser-known subsidiary or company so as to not be obvious to the User

[10] In an ideal world, it would be required that all data obtained from users for free from the user must be made publicly accessible. In such a universe, a possible alternative would be for the Entity to provide compensation to the User as an incentive to consent, and thus not be required to publicly disclose the information as it was purchased and acquired ethically with consent. However, that is not practical as it would put the User's personally identifying information at risk and open them up to be targeted. So, instead of making data obtained for free be available for free, it is more appropriate to put some restrictions on what can be monetized and how by the Entity–specifically, by requiring consent for such a purpose.

5 Discussions and Concerns

This section discusses some of the practical specifics of implementation as well as possible limitations.

A huge reason why social media sites such as Facebook and Twitter have such a high monetary value is because of the wealth of unique information on an enormous breadth of people that only these organizations have access to. This information pool gives them an advantage in industry to be able to perform cutting edge studies and tests on how to best capitalize and optimize specific experiences for their user base. There is nothing inherently wrong with the fact that these types of organizations have a monopoly on user data; however, it is wrong that they have this data without the consent of the users.

Should an academic institution want to recreate the same results for the projects that are being performed by these corporations, they would have to receive institutional approval for the safety and security of the study. They would need to receive informed consent from all participants. Risks must be disclosed to participants and any financial incentives or purposes must be disclosed as well. In the United States, something as simple as a voluntary survey is considered research on human subjects, must have IRB approval, and disclose information regarding intent and purpose to the participant. The premise behind this Code is that we endeavor to create a standard that doesn't limit what information large corporations can collect on their users, but merely informs them and allows them the ability to consent to that usage or not.

In an ideal world, government lawmakers would pass a law forcing software companies to adhere to this Code. It would likely significantly interrupt their revenue flow and inhibit their power greatly, and thus receive strong opposition and push-back. However, while these companies would take significant hits in terms of their power and revenue, it is ethically justified because these companies did not receive their power and revenue in an ethical way in the first place.

The most complicated and least practical part of the Code to implement is the detailed listing of each individual purpose that corporations might use user usage data for simply because there are so many uses. The likelihood of a user reading all of them and opting into them is very small, which would limit corporations enormously in practicality while it would not necessarily do so in theory. Additionally, it would limit the ability of corporations to work on trade secrets and other secret projects using user data, because it would have to be disclosed. While a strict interpretation and implementation of this Code would be the most honest approach that respects the privacy and autonomy of each user, a proper peer review and discussion might narrow down a list of what specifically must be consented to into something that would be a little more generous while still maintaining these principles.

While this Code was designed specifically to guide social media websites, it would also be applicable to organizations such as Google or Amazon, which also collect a large amount of Usage Data on their users. The general principle behind this Code is not that

these organization should not have or own this data, but rather that the people they collect it from should have Informed Consent when providing this data. Theoretically, if a gaming website were to collect user data to better optimize what features of the gaming website to prompt the user to try, and the entire experience occurred within the website and only promoted features owned by the website, this would be a standard optimization technique and not misappropriation of user data. When advertisements, third parties, research on human subjects, and monetary benefits for the organization are concerned, then it is no longer an internal, operational usage and would need user consent for the data to be used for the aforementioned purposes.