

## A review of Reddit's Privacy Policy

Author: Ferdous Alam

Reddit is a social networking and social news/aggregator site. As of February 2022, the site ranks as eighth most visited site in the US and twenty second globally. Reddit is comprised primarily of user-generated content. Although the company is currently private, "Reddit could make its stock market debut in March, Bloomberg reported. The social media company could seek a valuation as high as \$15 billion."

The Reddit Privacy Policy statement articulates what data is collected and how it is used and shared. They also provide information about how they protect the information. In addition, they provide the details about the choices and rights available to the Users. In terms of the collection of data it is categorized into three broad categories:

I) The first classification is the data that 'a user would provide'. Reddit mentions that "If you create a Reddit account, we may require you to provide a username and password." This is quite ambiguous because it is not possible to create an account without providing a username or password. In addition, Reddit will collect email address, bio, profile account preferences and settings. Any content that is submitted to the service as well as any communication (private messages, chats) that takes place among other users is collected. Reddit also collects any action that is taken while using

the app or the platform. Essentially, they have a very broad definition of action in the platform stated as "We collect information about the actions you take when using the Services" which likely refers to an all-encompassing data collection within the app.

II) The next classification falls under the umbrella of what data is collected automatically. This includes "IP address, user-agent string, browser type, operating system, referral URLs, device information (e.g., device IDs), device settings, mobile carrier name, pages visited, links clicked, the requested URL, and search terms". While Reddit will delete any IP addresses collected after 100 days, which may or may not be a reasonable time, they specifically mention that they will store the IP address that was used to create the account.

III) The third category of collection of data is from (a) third party, (b) information collected from other sources and (c) information Collected from Advertisers and Potential Advertisers. One of the reasoning in terms of usage is the intention to "*combine that information with the other information we have about you*". In terms of the how the data is used Reddit makes explicit claim that "*Much of the information on the Services is public and accessible to everyone, even without an account. By using the Services, you are directing us to share this information publicly and freely.*" While Reddit takes a cautious approach in terms of Aggregated or de-identified information, they do not make any explicit claim to ensure its implementation. Per their privacy statement "Reddit *may* share information about you that has been aggregated or anonymized such that it cannot reasonably be used to identify you". The CALOPPA recommends states that the website can you ensure that authorized third-party trackers comply with your Do Not Track policy and if not, disclose how they might diverge from your policy. It does not seem Reddit's privacy policy is able to confirm the tracking practices with those responsible for their or ensure that their practices correspond to what they say in their policy.

From a control perspective, "A right to privacy can be viewed as right to control information about oneself. Transparency-and-choice appears to model control because it allows individuals to evaluate options deliberately and then decide freely whether to give or withhold consent." This seems unlikely since there isn't option to use the service in an anonymous manner. Overall, in terms of the collection of data looking at it from a multidimensional privacy analytic framework the dimension of meta-harm (collection, dissemination) seems to have a significant

impact of not being properly addressed. If we took a deeper dive and look at the FTC Report “Protecting Consumer Privacy in an Era of Rapid Change” it does not appear that Reddit has “contractually prohibits downstream recipients from trying to re-identify the data”.

The FTC report also noted “the lack of transparency about the practices of information brokers, who often buy, compile, and sell a wealth of highly personal information about consumers but never interact directly with them. Consumers are often unaware of the existence of these entities, as well as the purposes for which they collect and use data.”. In terms of data sharing Reddit mentions “With our service providers. *We may share information with vendors, consultants, and other service providers who need access to such information to carry out work for us*”. In addition, Reddit may also share data “*With our affiliates. We may share information between and among Reddit, and any of our parents, affiliates, subsidiaries, and other companies under common control and ownership*”. From a first principal standpoint it appears to not be meeting the FTC guideline have mentioned as stated above.

From a data protection perspective Reddit “use HTTPS” for data transmission as well as offer option to maintain the security of your account by configuring two-factor authentication. Although it addresses the data transmission and access, there are not enough information provided in terms of how they are protecting the data that they have collected. In addition, the duration of the data that is kept is stated “as long as it is necessary for the purpose(s) for which we originally collected it”. It would not be unfair to conclude that whether Reddit has a “comprehensive data management procedures throughout the life cycle of their products and services” is unclear from their privacy statement.

When we look at the rights of the User, per Reddit’s user agreement, blanket consent from the User is implied. “When Your Content is created with or submitted to the Services, *you grant us a worldwide, royalty-free, perpetual, irrevocable, nonexclusive, transferable, and sublicensable license to use, copy, modify, adapt, prepare derivative works of, distribute, store, perform, and display Your Content and any name, username, voice, or likeness provided in connection with Your Content in all media formats and channels now known or later developed anywhere in the world. This license includes the right for us to make Your Content available for syndication, broadcast, distribution, or publication by other companies, organizations, or individuals who partner with Reddit. You also agree that we may remove metadata associated with Your Content, and you irrevocably waive any claims and assertions of moral rights or attribution with respect to Your Content.*”

From Solove’s taxonomy of privacy, the information collection would clearly create problems with what is referred to as Surveillance. Essentially, they have a very broad definition of action in the platform stated as “*We collect information about the actions you take when using the Services?*” which likely refers to an all-encompassing data collection within the app and thereby may not have a justification how there may be overarching surveillance.

The other broader problem is the Identification issue where it appears that once there is a transaction made through the app, Reddit will collect information that impede on the user’s ability to be anonymous or pseudonymous. Per the guideline “If you purchase products or services from us (e.g., Reddit Premium or Reddit Coins), we will collect certain information from you, including your *name, address, email address, and information about the product or service you are purchasing.*” While more sophisticated users would understand the lack of being anonymity in the platform, however this is not clarified in Reddit’s guideline. One can however create a throwaway account which will allow individuals to engage in some pseudo anonymous form. However, Reddit mentions that “Much of the information on the Services is public and accessible to everyone, even without an account. By using the Services, you are directing us to share this information publicly and freely.” Therefore, it is unclear whether the service can be used with the intention of being anonymous.

Another potential issue with deleting the account seems to not encompass or provide option to erase one's footprint. Per Reddit's privacy policy "*when you delete your account, your profile is no longer visible to other users and disassociated from content you posted under that account. Please note, however, that the posts, comments, and messages you submitted prior to deleting your account will still be visible to others unless you first delete the specific content.*" This seems quite complex that associated history will not be deleted, when a general review of the privacy guideline seems to indicate they have the resources and are collecting all associated data with the account.

Reddit mentions that the service can only be used by "Children under the age of 13 are not allowed to create an account or otherwise use the Services. Additionally, if you are in the EEA, you must be over the age required by the laws of your country to create an account or otherwise use the Services, or we need to have obtained verifiable consent from your parent or legal guardian." Per COPPA rule "if the company is covered by COPPA, you need to have certain information in your privacy policy and get parental consent before collecting some types of information from kids under 13." In the absence of required verification of age Reddit, may be collecting information of its user base from user's who do not meet the age threshold.

In terms of Access Participation FTC guideline mentions "Access is the third core principle. It refers to an individual's ability both to access data about him or herself — i.e., to view the data in an entity's files — and to contest that data's accuracy and completeness. Both are essential to ensuring that data are accurate and complete. To

be meaningful, *access must encompass timely and inexpensive access* to data, a simple means for contesting inaccurate or incomplete data, a mechanism by which the data collector can verify the information, and how corrections and/or consumer objections can be added to the data file and sent to all data recipients" To review Reddit's response I had requested the data that is being collected under my account. Unfortunately, the first attempt did not result in a successful attempt in retrieving the data. In addition, what is more concerning is that there is a thirty- day moratorium in terms of how often you can request the data that Reddit have collected. Attempting to send a follow up request showed the following message: "It looks like we've already received this kind of request from you in the last 30 days. For an updated copy of your data, please come back after 30 days have passed since your last request. "

In summary the while Reddit's privacy policy addresses with a fair amount of detail in explaining their process of data collection, usage as well as some of the choices and rights of the users in their platform, they seem to be lacking some of essential features like anonymous browsing, choice of content ownership, right to delete; furthermore, more transparency is warranted from stored data protection as well as clarification regarding the dissemination of user data to third parties and affiliates.

### Reference:

Reddit's Veil of Anonymity: Predictors of engagement and participation in media environments with hostile reputations, 2018, Danielle K. Kilgo, Yee Man Margaret Ng, Martin J. Riedl, and Ivan Lacasa-Mas

Faceless Bodies: Negotiating Technological and Cultural Codes on reddit gonewild, 2013.  
Emily van der Nagel

Privacy Policy <https://www.redditinc.com/policies/privacy-policy-september-12-2021>