

SOCIAL CONSTRUCTION OF PRIVACY: REDDIT CASE STUDY

A **Research Paper** submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia | Charlottesville, Virginia

In partial fulfillment of the requirements of the degree
Bachelor of Science, School of Engineering

By:
Rajiv Sarvepalli
Spring 2021

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signature _____ Date _____
Rajiv Sarvepalli

Approved _____ Date _____
Richard D. Jacques, Ph.D.

SOCIAL CONSTRUCTION OF PRIVACY: REDDIT CASE STUDY

INTRODUCTION

In the technological age of today, privacy becomes a more and more valuable commodity. With so many companies that live off the idea that information is money, it becomes increasingly concerning the amount of an individual's information that is public. It is public in every sense of the word, not just to a group of people, but to the whole world. Consider the constant data scandals that plague our technological world. Whether it is Facebook, Google, or governments, someone is always getting caught selling, collecting, or losing data that many consider infringes on their privacy. Therefore, as stewards of these technologies, we must develop preemptive ways of protecting the privacy of the individual in an information-based world focused on the collective. The heterogeneous nature of society, especially with respect to privacy, makes the perspective vary greatly from person to person. This study shall focus on Reddit, an anonymous social media since individuals within anonymous social media communities tend to view anonymity as some form of privacy and therefore tend to care about it. In order to understand the perspective and definitions of privacy, privacy needs to be analyzed in the context of a society.

The conundrum of today's information market is characterized by the constant complaints of lack of privacy, with no privacy-invasive companies changing their conduct. Every year, new scandals from large technology companies are reported in the news. The same indignant response from the public happens every time, yet those companies only continue to grow. This paradox illustrates the privacy of modern society is too complex to simply view as an ethical issue but must be viewed as a social product. The ambiguous and changing definitions of privacy are too complicated to be considered anything less than a sociotechnical establishment.

RESEARCH QUESTION

As we address the concerns of privacy in social media, this study attempts to approach the following questions: how do text and imagery in social media expose locality that may be considered private by the individuals sharing that information? What constitutes privacy in terms of locality in a public social media setting? What constitutes a breach of privacy in anonymous social media? Does Reddit's perspective on privacy shift the communities' perspective in one direction or another? How do the economic players that have influence over Reddit influence the communities' privacy concerns? What is the connection between government powers, economic players, and the actions of Reddit's administrators? By limiting the groups studied to only anonymous social media and specifically Reddit, the definition of privacy and its relation to other factors (socioeconomic, interest groups, etc.) can be explored in more depth.

Reddit is chosen as the case study since it is one of the most prominent anonymous social media websites in the world. Anonymous social media, in general, should have more concern for privacy and thus more involvement in a study and analysis of privacy itself. Additionally, Reddit's community tends to be concerned with privacy and worried about data being used in inappropriate manners. Finally, the style and composition of Reddit make it accessible and simple to attain information about previous posts and timelines (that occurred on Reddit).

RELATED WORKS

The problem of privacy in anonymous social media is touched upon by several publications, but few explore in-depth the societal communities and their perspectives on privacy within anonymous social media. Only one article seemed to explore privacy in the same social media but from a completely different perspective. In that publication, they analyze the

ability of machine learning techniques to classify users based on their information. The results in this paper indicate the ability to classify gender and citizenship relatively well based purely on the user's text and therefore indicates the possibility of privacy infringement in ways the user may not have perceived as possible. However, analyzing why users perceive this as a privacy violation is not something analyzed in this paper, and something our proposition is interested in (Fabian et al. 2015).

Socioeconomic Construction of Privacy

Privacy negotiation is discussed in an STS-related journal with heuristics designed to enable better methodologies of resolving individuals conflicting opinions and perspectives (Such and Rovatsos 2016). Another STS-related journal discusses how anonymity is perceived by the users in anonymous social media, and the impossible nature of anonymous networks (to be connected in some manner, one cannot be completely anonymous). This paradox affects how well this model for communication will survive (Sharon and John 2018). Additionally, an STS-related journal analyzes the socioeconomic factors' connections to online privacy literacy, perceptions, and definitions (Epstein and Quinn 2020). An STS-related journal analyzes the demographics with connections to social media behavior, specifically the connection with demographics to privacy settings on social media accounts (Madden 2012). Finally, an STS-related journal identifies the demographics of teens and analyzes the understanding of privacy teens have. Additionally, it studies the ways privacy is achieved and attempts to determine how privacy can be achieved in public networks (Marwick and Boyd 2014). Overall, STS-related journals explore the multitude of ways demographics, societies, and individuals interact with privacy in technology, analyzing the different perspectives, characterizations, and literacy of privacy itself.

Technical Shaping of Privacy

One publication indicates the inability of private users to remain private in social media with a mix of public and private users. The public user's data of social media can be effectively used to predict the private user's behavior indicating an inability to remain truly private in most of the world's current social media (Zheleva and Getoor 2009). One paper studied the effect of Big Data on the average user's privacy and the parts they should be concerned about. They propose a concept that will enable users to be effectively informed about the relevant privacy concerns within Big Data (Smith et al. 2012). Whether anonymous social media is truly anonymous is analyzed by finding out if the information on anonymous social media is sufficient to track or identify users (Chatzistefanou and Limniotis 2019). Another publication surveys the work done so far in user privacy protection and analyzes different techniques and algorithms for privacy prevention and anonymization to determine future research directions and issues. (Beigi and Liu 2018).

Overall, the literature indicates some investigations into the problems of privacy inside social media, the failure of anonymous social media to be truly anonymous, and the different demographic perspectives on privacy. However, the current literature fails to investigate the different relationships with interest groups (communities created around a shared interest) and their definitions, perspectives, and literacy on internet privacy. The current literature investigates the technical or social impacts of privacy, but taking a sociotechnical lens to privacy is missing. Privacy is mutually shaped by both technological advancements and social changes, so to analyze only one aspect you lose the interconnectedness between the dimensions. This paper considers both aspects and attempts to explore the dynamic product that privacy is. By analyzing Reddit, a dominant anonymous social media/news platform, these relationships can begin to be explored.

ANALYSIS

To consider the relevant social groups concerning privacy in Reddit, first, we must consider the heterogeneous opinions on the matter of privacy. The obvious groups are the separation of the community and company. However, there are more social groups that should be investigated here. The investors and governmental powers that influence those investors are important players in defining privacy. Additionally, the entire community is not just one individual and there is a wide variety of opinions regarding privacy. Since Reddit is anonymous social media, define interest groups to be the defining categories of the community which has many connections to numerous socioeconomic groups. The economic players influence Reddit's policies concerning privacy and governmental entities have control over these players. For example, Tencent, a technological investor based in China, caused much controversy by investing in Reddit since the community, which is an American majority, feared Chinese governmental restrictions and influence, especially concerning privacy.

All of these social groups designate their definitions for privacy and, in doing so, attempt to enforce these definitions through whatever means available to them. Reddit as a company decides what policies to allow the entire platform to have: what types of speech are allowed and what they do with the collected user data. Users enforce their concepts through being moderators, which set community rules within their specific community (subreddit), and through market economics. Reddit attempts to please users, so indirectly, users can influence the overall companies' policies and privacy perspectives. The economic investors of Reddit influence the company's decision behind the scenes as do the governmental players. Market economics plays a role here, but government concerns, laws, or desires also heavily influence the decisions of these players, but most of this is not done openly. As a result of this, the

community's reactions also feedback into this loop refining and reconstructing the rules and privacy definitions that govern Reddit as a whole.

The economic investors and Reddit as a company (employees) work together internally to resolve disagreements over the direction of the company. The governmental players have indirect influence over the economic players through laws and monetary influence. The community of Reddit itself communicates directly to the company with complaints, but its main form of influence comes through market economics. Reddit desires to please the community since they are their user base. Their monetary model requires customers (users) to be financially successful. Therefore, they work to ensure that their users do not leave for other social media sites in most cases. It is important to note the distinction between purely pleasing and pleasing to ensure users stay. Reddit does have a monetary incentive in the short term to do anything more than keeping users on their websites. As such, this can lead to pervasive incentives such as working to ensure a relative monopoly of the popular anonymous social media websites.

The loop of interactions that involves all the social groups does not have a clear resolver, but there are sort of checks and balances in place that disables any player complete control. Reddit as a company can make its own decisions for some time, but economic investors will use their control to ensure that profit is being made. Just as the community cannot make direct decisions but can influence both the economic players and the company to make decisions that align with their beliefs. Finally, governmental powers must play all their parts concerning their citizens, laws, and leaders. No group or individual has complete resolve power, but the checks and balances tend towards the common goals, aggregating when there are disagreements. The concerns of privacy follow this pattern as do most decisions concerning the use of the technology of Reddit.

CASE STUDY: REDDIT

Reddit is one of the most popular discussion websites that are typically anonymous. Users rarely have profile pictures and often use names that are unrepresentative of their own names. As such, it can be considered under the larger category of anonymous social media. The first place to understand privacy within Reddit is to understand what Reddit's privacy policy is. Reddit's own privacy policy helps define what data they collect on users and how they use the data. However, the ways they can use an individual's data are often broadly classified such as "Communicate with you about products, services, offers, promotions, and events, and provide other news and information we think will be of interest to you" or "Measure the effectiveness of ads shown on our Services" ("Reddit Privacy Policy", 2020). These broad definitions make it difficult to understand Reddit's perspective on privacy. However, it appears Reddit has a focus on consent being an essential part of privacy. As they state, "We may share information about you with your consent or at your direction." However, what is consent in this context? It appears that consent can be quite easily given with default options and things such as terms of service. Additionally, you can give consent to third-party applications to link your Reddit account to other social media such as Facebook and Twitter. In this context, Reddit views you as giving consent to the privacy policies of Facebook and Twitter.

Governments have their own perspectives on privacy and what constitutes an individual's right. For multinational companies such as Reddit and many internet companies complying with the laws of different nations can be a problematic issue. The manner in which the government can use Reddit's posts is difficult to conclude. Additionally, it is difficult to

know how anonymous Reddit remains to governments. However, to get some understanding, Reddit’s transparency report released data on the subpoenas from different governments.

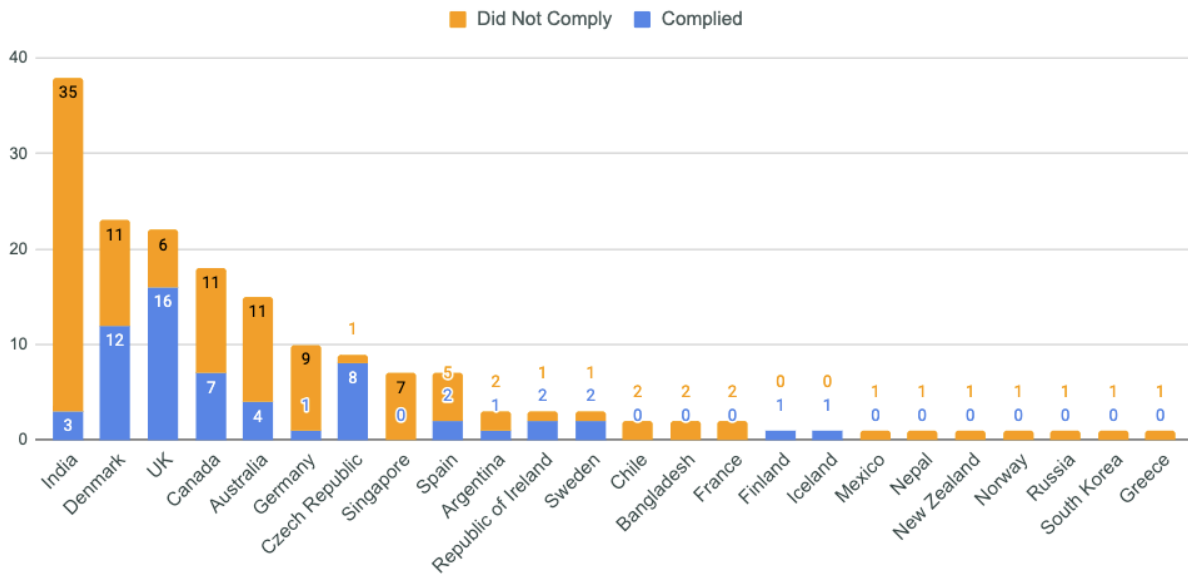


Figure 1. This image illustrates the “Compliance rate for International User Information Requests, by country” (“Transparency Report 2020”, 2020).

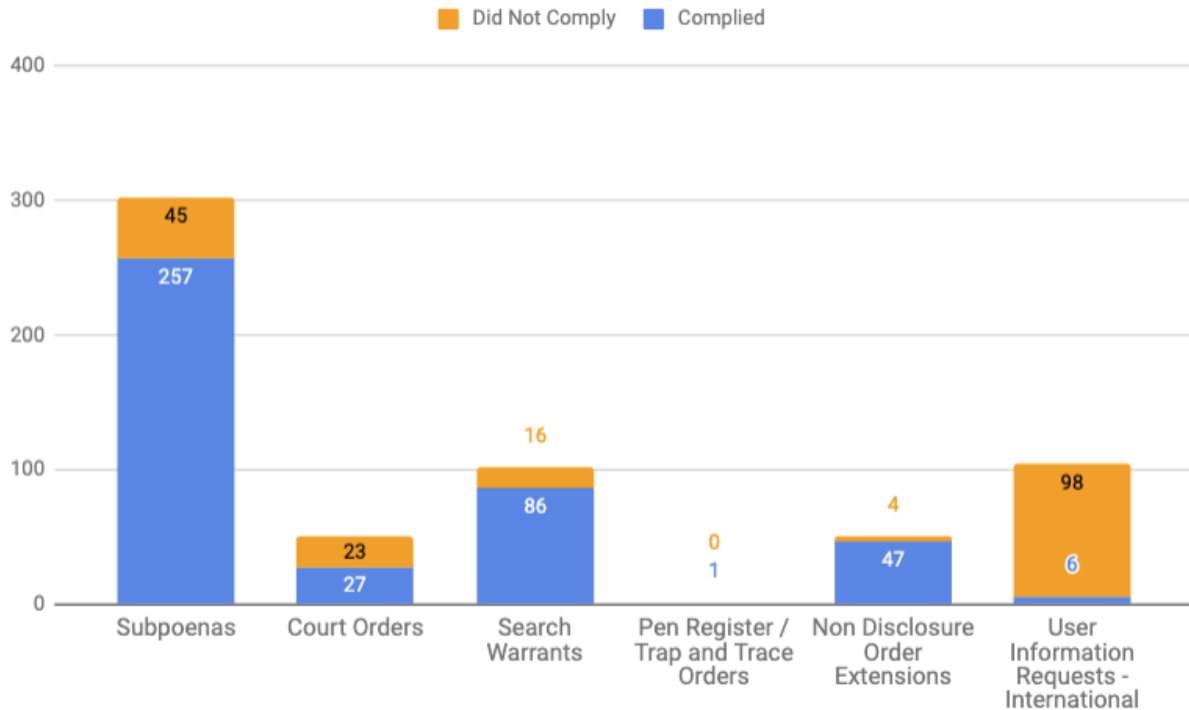


Figure 2. This image illustrates the “Chart 17. Compliance rate for standard non-emergency LE/Gov requests” (“Transparency Report 2020”, 2020).

These charts provide an understanding of the requests made by different countries to subpoena or receive information from Reddit and how often Reddit complies. However, it appears that Reddit can give away your information to the governments, it does not choose to do so all the time. In this sense, there is some degree of privacy, but it is impossible to know what constitutes the reasoning and therefore definition of this privacy.

Finally, we come to the user perspective on privacy. Given the wide variety of opinions when considering such a large group such as the entire userbase of a well-established social media, the aggregate opinion will be analyzed. It appears the majority of users are not especially concerned about privacy, since they think the anonymous nature of Reddit conceals the

connection between their Reddit identity and their identity. However, many users are concerned about governmental influence (especially from China) in using the data available on Reddit and censoring certain content. A Time's magazine covered this concern stating that many users were concerned with the potential censorship issues (Law 2019). This stemmed from a company called Tencent investing in Reddit. The Peoples Bank of China licensed the same company to try a social credit system; a system that concerns many for its privacy implications.

CONCLUSION

Additionally, privacy literacy most likely differs from interest group to group, but many will probably be surprised at the lack of anonymity in anonymous social media. The younger demographics probably view their privacy as somewhat compromised already and are less concerned. The company's and users' expectations on privacy and consent often clash, but reconciliation is not made through a lack of privacy literacy. Many do not consider what Reddit does with their data and when privacy scandals, people are outraged but for a short time. The lasting consequences are not recognized so companies that fail to establish the bounds on privacy continue to foster and grow.

REFERENCES

- Beigi, G., & Liu, H. (2018). Privacy in social media: Identification, mitigation and applications. *CoRR*, abs/1808.02191. <http://arxiv.org/abs/1808.02191>
- Chatzistefanou, V., & Limniotis, K. (2019). Anonymity in social networks: The case of anonymous social media. *International Journal of Electronic Governance*, 11(3-4), 361–385.
- Epstein, D., & Quinn, K. (2020). Markers of online privacy marginalization: Empirical examination of socioeconomic disparities in social media privacy attitudes, literacy, and

behavior. *Social Media + Society*,6(2), 2056305120916853.

<https://doi.org/10.1177/2056305120916853>

Fabian, B., Baumann, A., & Keil, M. (2015). Privacy on reddit? towards large-scale user classification.

Madden, M. (2012). Privacy management on social media sites. *Pew Internet Report*, 1–20.

Marwick, A. E., & Boyd, Danah. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*,16(7), 1051–1067.

<https://doi.org/10.1177/1461444814543995>

Reddit Privacy Policy. (2020, October 15). Retrieved from

<https://www.redditinc.com/policies/privacy-policy>.

Sharon, T., & John, N. A. (2018). Unpacking (the) secret: Anonymous social media and the impossibility of networked anonymity. *New Media & Society*,20(11), 4177–4194.

<https://doi.org/10.1177/1461444818768547>

Smith, M., Szongott, C., Henne, B., & von Voigt, G. (2012). Big data privacy issues in public social media.2012 6th IEEE International Conference on Digital Ecosystems and Technologies(DEST), 1–6. <https://doi.org/10.1109/DEST.2012.6227909>

Such, J. M., & Rovatsos, M. (2016). Privacy policy negotiation in social media. *ACM Trans.*

Auton. Adapt. Syst.,11(1). <https://doi.org/10.1145/2821512>

Law, T. (2019, February 11). Chinese Company's Reddit Investment Angers Users Who Fear Censorship. *Time*. https://time.com/5526128/china-reddit-tencent-censorship/?utm_source=reddit.com.

Transparency Report 2020. (2020). Retrieved from

<https://www.redditinc.com/policies/transparency-report-2020-1>.

Zheleva, E., & Getoor, L. (2009). To join or not to join: The illusion of privacy in social

networks with mixed public and private user profiles. Proceedings of the 18th International Conference on World Wide Web, 531–540.

<https://doi.org/10.1145/1526709.1526781>