

# Dumping the Closet Skeletons Online: Exploring the Guilty Information Disclosure Behavior on Social Media

Yukun Yang,<sup>1</sup> Yeman Huang<sup>2</sup>

<sup>1</sup>School of Information and Library Science, University of North Carolina-Chapel Hill, yukun@live.unc.edu

<sup>2</sup>School of Information Management, Wuhan University, huangym@whu.edu.cn

## Abstract

Privacy issues on social media are becoming an increasing area of concern. Paradoxically, some netizens are actively divulging their privacy online. Noticeably, some information is specifically guilt-related, though confession online is considered irrational. This preliminary study strives to understand this guilty information disclosure behavior through a mixed-approach. Analyzing posts and comments in a confession forum on Reddit, we find that sex-related and recreation-related topics prevail. Our qualitative investigation produces a thematic model with 71 codes, 17 concepts, 4 frames, 3 categories, and 9 relationships, capturing the intents, content, influencers of this behavior, and the interactions among users. Our contribution relies on the investigation of this peculiar behavior to better understand people's privacy behavior. Also, we render a sophisticated framework around guilt-inducing behaviors useful for future work. We also suggest it as a mixture of conformity and counter-conformity, a modern "technology of self" and a variant of Adaptive Cognitive Theory.

## Introduction

As social media permeates people's lives, its information-sharing nature has raised concerns about privacy issues. While people are more sensitive about the risk of privacy leakage (Tsay-Vogel, Shanahan, and Signorielli 2018), some studies also find the concurrent "privacy paradox" phenomenon, the self-disclosure of personal information (Barnes 2006; Hughes-Roberts 2013), which has been increasingly prominent and epidemic in recent years (Chen 2018; Gruzd and Hernandez-Garcia 2018).

Among the privacy broadcasters, some explicitly pour out their deepest dark secrets and confess guilt, or wrongdoings, to strangers online. This peculiar guilt-related revelation, that we define as guilty information sharing here, lumps with other similar behavior under the umbrella of privacy paradox and self-disclosure. It bears a relation-

ship with negative information behavior, which undermines one's likability (Forest and Wood 2012). It also echoes stigmatizing information sharing. Similar topics like mental health (Choudhury and Sushovan 2014) and weight loss (Chancellor, Hu, and Choudhury 2018) have been studied in the context of stigmatizing information sharing. Illicit information disclosure (Costello, Martin, and Brinegar 2017) is also relevant as guilt sometimes stems from illicit behaviors.

Nonetheless, we focus on the information sharing under this guilt context because of its nature as a complex of non-normative information behaviors mentioned above. The comparison between them has yet to be studied. Building on this, our study also helps further understand the privacy behavior where the privacy information is possibly more diverse and sensitive. It also offers the chance to observe an "offline-to-online" scenario since confession is commonly found in real-world religious rituals. Since it is an under-investigated area with only limited research (Levontin and Yom-Tov 2017), we try to answer the following questions:

- 1: What kinds of guilty information are shared online?
- 2: What are the purposes of this special self-disclosure?
- 3: How do people interact with each other in the online confession community?

We make three contributions in this research. First, we shed light on a comprehensive understanding of self-disclosure by filling the vacancy of guilty information disclosure study. Second, our qualitative section culminates in a structured framework with interactions between elements that can inform future works around the guilt-inducing behaviors. Third, our findings could be useful to link with psychology, sociology, and other relevant fields.

## Data Collection and Preprocessing

We scrapped data on the subreddit, r/confession, a specific online community for posting confessions and guilts. All

posts about guilty information and confessions are out of spontaneity without intervention (Choudhury, Counts, and Horvitz 2013). Posts (“submission” in Reddit’s language milieu) and comments disparate information sharing patterns are marked separately. Data are collected from April to October 2018. All texts are cleaned through punctuation, number, stop words, and rare words removal. Only verbs, nouns, and adjectives are retained after lemmatization for topic modeling.

### Preliminary Data Understanding

Our dataset contains 229,968 records in total (submissions, 7.64%; comments, 92.36%). The submission-comment ratio is similar to that in other studies (Choi et al. 2015), indicating the normal-level dynamic of r/confession. Submissions in our dataset have 153 words per sentence, which is approximately 3 times more than the comments’. In terms of diction, words in submission have 5.2 characters averagely. While submissions have fewer unique tokens in absolute value, the average unique token per submission is relatively high, showing the different vocabulary size of two corpora. 46k unique words are calculated in only 17k submissions while 225k unique words are tallied in 212k comments. The lexical richness of submissions might be considerably higher than that of the comments.

Type	Samples	Avg. Sentence Length	Avg. Word Length	Unique Tokens
Submissions	17,563	153	5.2	46,343
Comments	212,405	30	4.7	225,422

Table 1: Descriptive Statistics of the Crawled Dataset

### Topic Modeling of Submissions’ Text

Topic models are utilized to glimpse into the content of guilty information. Only submissions are included for their information richness and algorithms compatibility. To find the optimal numbers of topics for LDA modeling, we compare the coherence and perplexity of different k as Figure 1 shown. Based on domain knowledge and the best practices to avoid overfitting, we choose 27 for its relatively high coherence while perplexity value is not very informative.

Figure 2 shows the topics induced by their keywords, also the number of documents in every topic. Documents are plotted using t-SNE based on their topic probability. The guilt disclosed resembles “an encyclopedia of life”, covering almost every aspect of life. Main topics reside on “recreation”, “reddit”, “food”, “sexuality”, and “race”.

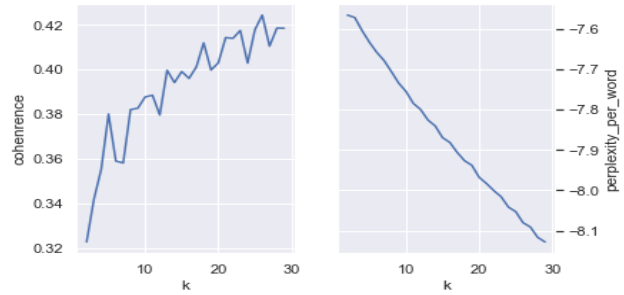


Figure 1: Model Performance of Different Number of Topics(k)

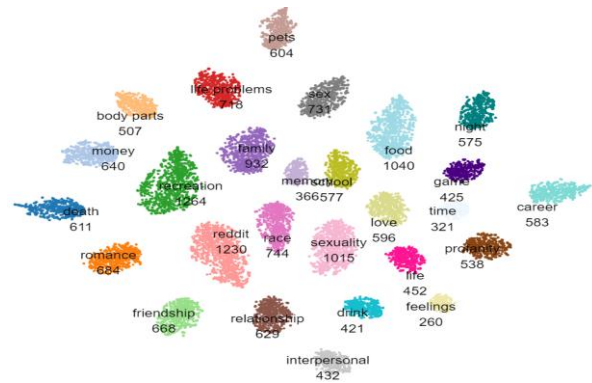


Figure 2: Topics(k) in the Submissions

### Thematic Analysis of Corpus

To capture the nuances in this disclosure behavior, we use the thematic analysis. After randomly sampling 800 submissions and 600 comments, we inductively encode the elements in the corpus individually for 5 rounds. Following a peer-debriefing session, we discuss the codes and refine our codebook iteratively until meeting mutual agreements. This collective quality-control approach is interpretive convergence (Costello, Martin, and Brinegar 2017). 71 codes (words in the parentheses) consisting of the thematic model are shown in Figure 3. The inter-rater reliability of codes reaches 83% (Miles and Huberman 1994). Our model identifies 17 concepts (captions of blocks in the picture), 4 frames (italic blue words), 3 categories (Submission, Community Interaction, Comments), and 9 relationships (words around arrows), which are all exhaustive and exclusive.

**Submissions** are the voluntary posts of guilty information. *The driving force* prompts users to confess online mainly including *eager for venting and mental relief*, *lack of social support*, and *following up other posts*. The first two concepts are endogenous and the last one is exogenous. *Eager for venting and mental relief* is the main motivation of online confession, with many users explicitly stated that “I

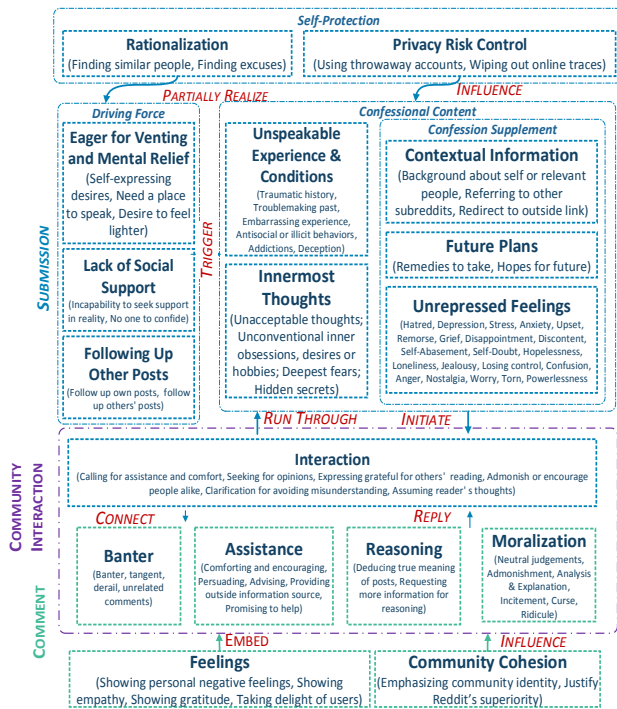


Figure 3: Thematic Model of Guilty Information Disclosure

just want to get it out my chest”. Usually, they need to relieve stress rather than expecting support or responses.

However, people suffering from the incapability also seek support in reality. *Lack of social support* is another outstanding impetus. In our corpus, these people usually fear seeking help in real life, since face-to-face communication could be embarrassing, the result is unpredictable and precarious, or they have “nobody to talk to”. Also, some people turn to social media as an alternative since their ongoing counseling or therapy has little effects. In addition, some submissions only serve as following-ups to other posts. Users usually strike an echo with the others and want to share their own story, or just further explain the existing confession threads.

**Confessional content**, as what confessors reveal on social media, shows noticeable story-telling hallmarks. We found *unspeakable experience or conditions* and *innermost thoughts* are two concepts generalizing guilts to confess. *Unspeakable experience and conditions* involve their traumatic history, troublesome past, embarrassing experience, antisocial or criminal behavior, deception, etc. These submissions usually are stories with concrete plots. *Innermost thoughts* submissions are more random and fragmented, and usually, have nothing to do with actions. They include unacceptable thoughts, unconventional inner habits, desires or hobbies, deepest fears, hidden secrets, and more.

**Confession supplement** is a special frame inside the **confession content**. It lubricates the story-telling or thoughts-revealing. People usually mention specific time, actors, places and other important settings as *contextual*

*information*. It is common to see people refer to other sub-reddit for readers’ better understanding. *Future plans* are usually at the end of the story. Remedies to make are mentioned if users are responsible for wrongdoings. In the confession content, *unrepressed feelings* are oozed from lines to lines. Most of them are negative, such as hatred, depression, anxiety, and hopelessness.

**Self-Protection** is a conscious act of risk minimization. Due to the sensitivity of the guilty information, people would likely to be judged morally by others. It is understandable that submission writers would resort *rationalization*. By finding excuses, for example, one comfort himself by guessing his post is “too generic for anyone to read and care about”. Due to guilty information’s privacy nature, users would take actions to prevent the guilt posting yielding impact in reality. Confessors would state using “throw-away accounts” or would “delete [the post] if needed”.

**Community interaction** is composed of interaction initiated in submissions and responses in comments.

**Interactions from submissions** suggest that people would most likely to be demanders who call for assistance, comfort, or just opinions as a mean to fulfill their lack of social support. Assuming readers’ thought is also frequent as many people “knowing [the post] sounds bad”. In fewer cases, people admonish or encourage the reader, as they may “feel” or “experience the same”.

**Interactions from comments** are more diverse and complicated. *Banter* is prevalent on the comments as witty repartee. However, serious responses also exist. Some would proffer *assistance* in the comments, namely comforting, encouraging, persuading, advising, giving promises to help, or providing direct outside links. *Reasoning* shows the commenters mental activities after viewing the submission. While some people try to understand the posts in their own way, others may request additional information. *Moralization* also appears in comments in multifarious forms, such as neutral judgment, cursing, ridicule, etc.

**Comments** are the responses of submissions. Some commenters would emphasize the *community’s atmosphere* by saying “this is r/confession” to mediate arguments or welcome the newcomers. Off-topic discussions like how Reddit is superior to other social media platforms also exist.

Much like unrepressed feelings in submissions, commenters would also reveal users’ *feelings*. However, their emotions are not all negative, with some exception that people gloating over the users.

**Relationships** between concepts and frames show the community dynamics. On the top level, submissions and comments are connected via community interactions. This process is bilateral and iterative. Submissions signal the intention of interaction, and then active comments bridge two types of users. In the submission’s category, intents trigger users to create threads. Posting confessional content becomes a way to realize users’ intents. At the same time,

rationalization also plays a role in intent realization. Privacy risk control exerts influence on the confessional content, constraining the content to avoid undesirable aftermath in reality. About comments, varied forms of comments make a reply to interaction initiated by submissions. The content of comments also embeds commenters' feelings and are influenced by the community cohesion and norms.

## Conclusions

Based on the descriptive analysis, we find similar characters like using throwaway accounts and story-telling narratives mentioned in similar research. Barring this, some interdisciplinary insights could also be highlighted:

### Counter-conformity Coexists with Conformity

As privacy paradox is partially attributed to the conformity and mutual imitation (Barth and De Jong 2017), our research supports this interpretation since the following-ups surface as a main intent of disclosure. However, guilty information is also a part of counter-conformity that related to deviant behaviors. Mass of counter-conformists promotes the conformity of the information sharing.

### Social Media as a “Technology of Self”

The guilty information sharing functions similarly to the traditional religious confession that prevents “internalizing negative feelings” (Meek, Albright, and Mcminn 1995). When this scene moves from offline to online, it also corresponds to Foucault’s theory of “technology of self”, the practices that are undertaken by the self and directed toward it via self-transformation. The medieval confessional practices, as an example of the “technology of self”, is projected to the digital age. Our finding expands the connotation of social media as an interpellation tool (Friesen 2017); it could also act as a direct confessional tool.

### Collision of Social Network Participation Phases

Previous study proposed the Adaptive Cognitive Theory to address the privacy paradox: user’s privacy concerns on SNS are different and they progress from Initial Use (IU), then Exploratory Use (EU) to Managed Use (MU) (Hu and Ma 2010). Our observation does not follow this model. Users could have the IU and EU overlapped. Also, people in different phases may have similar privacy risk controls. This implies the complication of user participation on SNS as they may follow a more abrupt and random pattern.

## References

Barnes, S. B. 2006. A Privacy Paradox: Social Networking in the United States. *First Monday* 11(9).

Barth, S., and De Jong M. D. T. 2017. The Privacy Paradox – Investigating Discrepancies between Expressed Privacy Concerns

and Actual Online Behavior—A Systematic Literature Review. *Telematics and Informatics* 34(7): 1038–58.

Chancellor, S.; Hu, A.; and Choudhury, M. D. 2018. Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. New York, NY: ACM Press.

Chen, H. T. 2018. Revisiting the Privacy Paradox on Social Media With an Extended Privacy Calculus Model: The Effect of Privacy Concerns, Privacy Self-Efficacy, and Social Capital on Privacy Management. *American Behavioral Scientist* 62(10): 1392–1412.

Choi, D.; Han J.; Chung, T.; Ahn, Y. Y.; Chun, B. G.; and Kwon, T. T. 2015. Characterizing Conversation Patterns in Reddit: From the Perspectives of Content Properties and User Participation Behaviors. In *Proceedings of the 2015 ACM on Conference on Online Social Networks*, 233–43. Palo Alto, Calif.: ACM Press.

Choudhury, M. D., and Sushovan D. 2014. Mental Health Disclosure on Reddit: Self-Disclosure, Social Support, and Anonymity. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 71–80. Menlo Park, Calif.: AAAI Press.

Costello, K. L.; Martin, J. D.; and Brinegar, A. D. 2017. Online Disclosure of Illicit Information: Information Behaviors in Two Drug Forums. *Journal of the Association for Information Science and Technology* 68(10): 2439–48.

Choudhury, M. D.; Counts, S.; and Horvitz E. 2013. Social Media as a Measurement Tool of Depression in Populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, 47–56. Paris, France: ACM Press.

Forest, A. L., and Wood, J. V. 2012. When Social Networking Is Not Working: Individuals with Low Self-Esteem Recognize but Do Not Reap the Benefits of Self-Disclosure on Facebook. *Psychological Science* 23(3): 295–302.

Friesen, N. 2017. Confessional Technologies of the Self: From Seneca to Social Media. *First Monday* 22(6).

Gruzd, A., and Hernandez-Garcia, A. 2018. Privacy Concerns and Self-Disclosure in Private and Public Uses of Social Media. *Cyberpsychology Behavior and Social Networking* 21(7): 418–28.

Hu, Q., and Ma, S. 2010. Does Privacy Still Matter in the Era of Web 2.0? A Qualitative Study of User Behavior towards Online Social Networking Activities. In *Proceedings of Pacific Asia Conference on Information Systems*, 591–602.

Hughes-Roberts, T. 2013. Privacy and Social Networks: Is Concern a Valid Indicator of Intention and Behaviors?. In *Proceedings of 2013 International Conference on Social Computing*, 909–912. Alexandria, VA: IEEE Press.

Levontin, L., and Yom-Tov, E. 2017. Negative Self-Disclosure on the Web: The Role of Guilt Relief. *Frontiers in Psychology*.

Meek, K. R.; Albright, J. S.; and Mcminn, M. R. 1995. Religious Orientation, Guilt, Confession, and Forgiveness. *Journal of Psychology and Theology* 23(3): 190–97.

Miles, M. B., Huberman, A. M. eds. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, Calif.: SAGE.

Tsay-Vogel, M.; Shanahan, J.; and Signorielli, N. 2018. Social Media Cultivating Perceptions of Privacy: A 5-Year Analysis of Privacy Attitudes and Self-Disclosure Behaviors among Facebook Users. *New Media & Society* 20 (1): 141–61.