

Toward Context-Aware Privacy Enhancing Technologies for Online Self-Disclosure

Tingting Du, Jiyeon Kim, Anna Squicciarini, Sarah Rajtmajer

College of Information Sciences and Technology
The Pennsylvania State University
University Park, PA, USA

arieldu812@gmail.com, jxk6167@psu.edu, acs20@psu.edu, smr48@psu.edu

Abstract

Voluntary sharing of personal information is at the heart of user engagement on social media and central to platforms' business models. From users' perspective, so-called self-disclosure is closely connected with both privacy risks and social rewards. Prior work has studied contextual influences on self-disclosure, from platform affordances and interface design to user demographics and perceived social capital. Our work takes a mixed-methods approach to understand the ways which contextual information might be integrated in the development of privacy-enhancing technologies. Through observational study of several Reddit communities, we explore the ways in which topic of discussion, group norms, peer effects, and audience size are correlated with personal information sharing. We then build and test a prototype browser extension that automatically detects instances of self-disclosure in Reddit posts at the time of posting and provides additional context to users before they post. We share this prototype with social media users, solicit their feedback, and outline a path forward for privacy-enhancing technologies in this space.

Introduction

Social networking platforms have become essential channels for connection and relationship building (Tsay-Vogel, Shanahan, and Signorielli 2018) and important sources of information and support. Central to these processes is so-called *self-disclosure* (SD), the act of revealing personal information about oneself to others (Jourard 1971; Kim and Dindia 2011). Alongside tangible personal and social rewards, SD can magnify privacy risks, as it leads to inadvertent over sharing of personal and identifying information.

A body of research has sought to understand contextual factors that may affect individuals' SD on social media (Andalibi, Ozturk, and Forte 2017; Umar, Squicciarini, and Rajtmajer 2019; Wang, Burke, and Kraut 2016; Ma, Hancock, and Naaman 2016; Stutzman, Capra, and Thompson 2011; Dwyer, Hiltz, and Passerini 2007; Krasnova et al. 2010; Bak, Kim, and Oh 2012; Zhao, Hinds, and Gao 2012). The majority of these studies rely on observational data collected from social network sites and focus on objective context, e.g., topics of conversation (Umar, Squicciarini, and Rajtmajer 2019; Bak, Kim, and Oh 2012). A smaller subset

of prior work has utilized survey data to understand users' perceptions of privacy risks and contextual influences on users' decision-making, e.g., perceived costs and benefits, audience, and content intimacy (Bazarova and Choi 2014; Stutzman, Capra, and Thompson 2011; Dwyer, Hiltz, and Passerini 2007; Krasnova et al. 2010; Cheung, Lee, and Chan 2015; Wang, Burke, and Kraut 2016; Ma, Hancock, and Naaman 2016).

Much of the research on disclosure has focused on social network sites where self-disclosure typically occurs among a select group of viewers, often including real life contacts (Shelton, Lo, and Nardi 2015). In contrast, relatively anonymous social media sites like Reddit allow users to communicate without adhering to the "profile and friend list" format of social network sites (Leavitt 2015). On Reddit, site-wide norms discourage participation with one's real name as a privacy-protecting measure, and even email verification is not required to create an account. This makes it possible for users to have multiple and disposable accounts (Proferes et al. 2021). The anonymity afforded by Reddit enables users to speak freely about controversial topics and share sensitive or personal information (Leavitt 2015; Ammari, Schoenebeck, and Romero 2019). Prior research suggests that anonymity, characterized by high awareness of self and low awareness of others, is associated with significantly higher levels of self-disclosure in online communication (Ammari, Schoenebeck, and Romero 2019; Joinson 2001; Clark-Gordon et al. 2019; Umar, Squicciarini, and Rajtmajer 2019; Ma, Hancock, and Naaman 2016). However, users are not fully protected even on anonymous platforms. Research has demonstrated that user profiles can be linked across different social media platforms despite pseudonymous identities (Backes et al. 2016).

Our work builds on prior research exploring the effects of contextual factors on SD in relatively anonymous online communities like Reddit. In particular, we take a mixed-methods approach to study the the contextual information which might be integrated in the development of privacy-enhancing technologies. The following research questions scaffold this work:

- **RQ1:** How do contextual factors correlate with observed self-disclosure in Reddit communities?
- **RQ2:** How can privacy-enhancing technologies support improved contextual awareness of self-disclosure?

We collect, annotate, and model conversations in three subreddits selected to represent communities of users self-aligned along personal and professional identities. Using this data, we explore contextual factors that appear correlated with SD (RQ1). We complement our observational data analyses with a survey-based study to understand how users react and make self-disclosure decisions in specific scenarios. Using a persona-driven approach, we present study participants with a prototype privacy-enhancing tool, namely, a context-aware browser extension that provides SD alerts and solicits participant feedback (RQ2). Our work makes the following key contributions:

- We identify key contextual factors and quantitatively measure their correlation with self-disclosure in an anonymous online community, Reddit;
- We build and test a prototype of context-aware privacy-enhancing tool to support improved user awareness around personal information sharing online.

Our findings align with prior work highlighting the impacts of topics, audience, and norms on SD (Wang, Burke, and Kraut 2016; Umar, Squicciarini, and Rajtmajer 2019; Ma, Hancock, and Naaman 2016). We deepen this literature, taking a user-centered approach to understand how these contextual impacts can support improved SD decision-making. Our work highlights the potential benefits of prompting users about the presence of SD and relevant context and suggests that users would welcome a simple, explanatory, context-aware SD notification tool.

Related Work

Social Underpinnings of Self-Disclosure

Several existing theories speak directly to the role of personal information sharing in building relationships. Social identity theory (Attrill and Jalil 2011) explains that people engage in ingroup-favoring behavior to maintain their social identity. SD behavior is a performance of identity. Social identity theory explains that people are likely to share more personal information and build relationships when they perceive others as part of the in-group (shared identity) (Turner et al. 1987; Tajfel 1982; Attrill and Jalil 2011).

Social penetration theory (Tang and Wang 2012), in particular, has been engaged in prior work on SD to explain the function of personal information sharing in relationship building online. Specifically, users steadily reveal more personal information as they build trust and deepen relationships with others. Social exchange theory (Cropanzano and Mitchell 2005) suggests the rational assessment of benefits and costs of personal information sharing. Trust acts as a critical factor in the exchange. Communication privacy management theory (Petronio, Child, and Hall 2021) shows that SD and relationship-building are affected by various factors, including perceptions of privacy, norms, and values.

Contextual Influences on Self-Disclosure

While SD is voluntary, it is also an influenced behavior. Prior work suggests that platform affordances can impact SD (Crowley 2019). Likewise, McKenna and Bargh (2014)

and Walther (1996) suggest that context specific to social media, e.g., asynchronicity and anonymity, has impact on SD.

Peer influence on individuals' online behavior has been a subject of substantial work. Researchers have demonstrated that individuals draw guidance from observing their peers' actions. Bakshy et al. (2012) studies "social contagion", suggesting that individuals are more likely to adopt behaviors when they observe their friends engaging in the same. In the context of SD, this means users' willingness to share personal information is shaped by perceived social norms. Cheung, Chiu, and Lee (2011), Li (2011), and Zhou (2011) also highlight the role of social influence on users' behavior. Cheung, Lee, and Chan (2015) examine the relative impacts of social influence on SD behaviors on Facebook, and find that social influence emerges as the single strongest factor driving SD in their study.

Audience also plays an important role in users' decisions around SD. Audience size and directedness—whether content is shared with a small, specific group or a larger, more general audience—impacts users' sharing decisions. Acquisti and Gross (2006) find that users are more cautious when sharing sensitive information in public settings. Wang, Burke, and Kraut (2016) report a negative correlation between network size and SD, while tie strength and network density demonstrate positive associations. Collins and Miller (1994) conduct empirical studies on dyadic relationships, determining that individuals are less likely to reveal personal information to acquaintances as compared to close friends. Choi and Bazarova (2015) explore the interplay between privacy boundaries and network attributes (size and diversity of audiences within the boundary) on users' SD.

Topics of discussion also influence users' SD. Different topics are associated with different degrees of SD (Wang, Burke, and Kraut 2016). For instance, subjects such as work and parenting may contain more personal details compared to, e.g., weather and food (Umar, Squicciarini, and Rajtmajer 2019). Relatedly, Ma, Hancock, and Naaman (2016) conduct an online experiment to study the relationship between content intimacy and SD in social media. They find people disclose less as content intimacy increases, and anonymity serves as a moderator for this relationship.

Context-Aware Privacy-Enhancing Technologies

The design of privacy-enhancing technologies (PETs) has been extensively studied, with a particular focus on contextual factors that influence their effectiveness and user adoption (O'Hagan et al. 2023; Coopamootoo 2020; Wang et al. 2014; Acquisti et al. 2017; Wang et al. 2013; Wisniewski, Knijnenburg, and Lipford 2017; Knijnenburg and Kobsa 2013). Several studies have explored how audience awareness can be incorporated into the design of PETs. For instance, privacy nudges that visually emphasize a post's reach can help users avoid regrettable disclosures (Wang et al. 2014, 2013; Ferreyra, Meis, and Heisel 2017).

Lesser considered in the design of PETs are the influences of peers and the importance of community norms, both important determinants of user sharing behavior (Garg et al. 2013; Utz and Krämer 2009). Information sharing

typically considers an imagined community and the social norms within it (Acquisti and Gross 2006; Utz and Krämer 2009). As a result, PETs have been criticized for ignoring the trust among individuals in a community (Gürses and Berendt 2010). While norms provide users with a general framework of expectations, they do not specify actions (Rashidi et al. 2020). People learn social norms by interacting with and observing others in a community, making peer behavior a vital observable signal of appropriate flow, and ultimately contextual integrity (Nissenbaum 2004).

Study 1: Observational Study of Identity-Driven Subreddits

To answer RQ1, we conduct an observational study using data collected from Reddit. We analyze conversations from various Reddit sub-communities and identify contextual factors correlated with self-disclosure behaviors.

Data Collection

We gathered posts and corresponding comments from three subreddits: *r/parenting*, *r/projectmanagement*, and *r/politicaldiscussion*. These were selected to represent communities of users self-aligned along personal and professional identities. All data was collected via a Reddit data archiving project¹. Full statistics are provided in Table 1.

Data Analysis

Data labeling. We used a BERT-based model to label all Reddit posts and comments for the presence of SD. In particular, we fine-tuned the pre-trained BERT language model (Devlin et al. 2018) on a public Reddit dataset. The OffMyChest dataset (Jaidka et al. 2020) consists of 12,860 labeled sentences sampled from comments on subreddits within *r/OffMyChest* and *r/CasualConversations*. Sentences in this dataset have 6 labels. We used 2 of these to train our model: *information disclosure*, comments revealing personal information; *emotional disclosure*, expressing the author’s positive or negative feelings.

Topic modeling. We used the unsupervised BERTopic model (Grootendorst 2022) to extract most frequent topics. There are three primary hyperparameters in this algorithm: number of nearest neighbors; number of components; and minimum cluster size. We used the hyperparameter optimization framework Optuna² to search for and select optimal hyperparameters.

False Discovery Rate (FDR) correction. We used the Benjamini-Hochberg (B-H) Procedure to control FDR to 5% when testing multiple hypotheses simultaneously.

Findings

RQ1: How do contextual factors correlate with observed SD in Reddit communities?

Topics of discussion. We identified the most frequently discussed topics in each of our three subreddits using the BERTopic model, and then calculated the corresponding SD rate for each topic. Table 2 lists the top four topics from each subreddit. The SD rate for a given topic was determined by calculating the ratio of samples containing SD to the total number of samples within that topic.

In the *r/parenting* subreddit, the top four topics are: support-seeking and gratitude; family member relationships; children’s birthday party; and, children’s sleeping problems. Compared to the other two subreddits, topics in the *r/parenting* subreddit generally have higher SD rates as we might expect. Discussions in the *r/projectmanagement* subreddit revolved around work-related themes. Prominent topics here include artificial intelligence, particularly ChatGPT, on the practices of project management and work-related challenges for individuals with ADHD. In *r/politicaldiscussion*, we identified a wider range of topics including gun violence, the 2020 US election, expressions of opinions and gratitude, and transgender athletes. Users tended to share their political views rather than divulge personal information. Thus, the SD rates here are comparatively lower than those observed in the other two subreddits.

We further disaggregate SD into emotional and informational categories. In the *r/parenting* subreddit, conversations revolving around support-seeking and gratitude, as well as family member relationships, predominantly feature emotional self-disclosure. Conversely, topics related to children’s birthday parties and sleeping problems mainly involve informational self-disclosures.

Audience. Data collected through our observational study supports analyses to explore the potential impact of audience on SD. We construct a network based on users’ reply interactions to explore the potential impact of user-audience ties on SD. In the network, nodes are unique users and weighted, directed edges represent pairwise interactions between users in the form of a reply to a post or comment. Using centrality as a proxy for the influence of a node in the network, we tested the relationship between a user’s SD rate and their indegree, outdegree, and closeness centrality.

We find a significant positive correlation between a user’s emotional SD and their indegree centrality ($r = 0.013, p = .02$) within the *r/parenting* subreddit. Based on the edge direction in our network, users exhibiting higher indegree attract a considerable number of replies from other users, signifying a substantial audience. Observed positive correlations suggest that a higher rate of emotional disclosure contributes to a larger audience size. In the *r/parenting* subreddit, **we find a significantly negative correlation between informational SD rate and outdegree centrality** ($r = -0.036, p < .001$). This suggests that individuals actively engaging in discussions within this subreddit tend to exhibit lower rates of informational SD.

In general, users with high closeness centrality can influence the entire network quickly. Our analysis revealed a **positive correlation between a user’s SD rate and their closeness centrality in the network** ($r = 0.027, p < .001$). Upon further examination, when we distinguished between

¹<https://arctic-shift.photon-reddit.com/>

²<https://optuna.readthedocs.io/en/stable/index.html>

Subreddit	Posts	Comments	Users	Conversations	Thread length
parenting	2917	127435	31801	55917	2.83±1.55
projectmanagement	173	3084	990	1373	3.01±1.44
politicaldiscussion	122	29908	5202	9996	4.88±3.54

Table 1: Statistics of the observational data collected from three subreddits.

subreddit	keywords	Emo SD	Info SD	SD
parenting	thank, people, good, think, comment, advice, know, right	0.360	0.066	0.381
	daughter, time, daycare, husband, relationship, need, tell	0.311	0.266	0.439
	toys, party, birthday, gift, invite, play, year, son, friends, fun	0.188	0.531	0.590
	sleep, bed, night, wake, nap, bedtime, room, routine, schedule	0.198	0.567	0.612
projectmanagement	work, like, time, need, team, people, experience, job, company	0.130	0.254	0.336
	chatgpt, questions, ask, provide, write, ai, generated, information	0.031	0.219	0.219
	ai, work, data, think, humans, change, tools, automate, impact	0.102	0.136	0.203
	adhd, work, day, struggle, energy, condition, mental, career	0.311	0.489	0.600
politicaldiscussion	guns, mass, shootings, firearms, weapons, violence, crime, laws	0.041	0.065	0.087
	biden, trump, win, vote, 2020, election, president, candidate, voter	0.101	0.091	0.158
	thanks, point, comment, wrong, good, discussion, source, agree	0.140	0.029	0.158
	trans, sports, women, gender, athletes, cis, sex, compete, biological	0.096	0.078	0.142

Table 2: Self-disclosure rates and keywords for hot topics in each subreddit. Emo SD represents the ratio of samples containing emotional disclosure, Info SD represents the ratio of samples containing information disclosure, and SD represents the ratio of samples containing either emotional disclosure, information disclosure, or both.

emotional and informational categories of SD, we observed that the correlation predominantly stemmed from the informational category ($r = 0.036, p < .001$). This finding suggests that, within the r/parenting subreddit, users with higher closeness centrality tend to share more personal information compared to other users. Users in pivotal positions within a network face higher privacy risks when sharing above-average amounts of personal information if they do not have an accurate measure of their actual audience. However, no significant patterns are identified within the r/projectmanagement and r/politicaldiscussion subreddits.

Peer effects. To investigate the influences of peers on individuals’ SD, we thread together the Reddit posts and comments we collected to recreate conversations and examine the relationship between SD in parent post and child replies. We are interested in peer effects of SD at two levels. First, we test for peer effects in direct replies (dyadic conversation). Then we test the proportion of SD in full-length conversations. Chi-squared tests for the relationship between SD in parent post and child replies are shown in Table 3.

We observe pronounced peer effects in the r/parenting subreddit for both dyadic ties and full conversations. The effects seem negatively connected with the SD attribute in the parent posts, that is, non-SD posts attracted more SD replies in the r/parenting subreddit. We further tested relationships between emotional and informational categories of SD in parent posts and SD in child replies. Results indicated that the negative peer effects mainly stemmed from the informational category, and we even identified an opposite trend for the emotional kind, where emotional SD posts were more likely to receive SD replies.

Observations. We also manually checked post content and found that most non-SD posts were one of a few types: asking parenting-related questions, e.g., “U.S. moms - how long was your total maternity leave?”; seeking advice and suggestions, e.g., “Looking for advice - anyone else here have a teen that gets migraines?”; or, weekly threads created by moderators, e.g., “This weekly thread is a good landing place for those who have questions about parenting, but aren’t yet parents/legal guardians and can’t create new posts in the sub.” These posts did not include self-disclosure, but redditors who responded to them were highly likely to include personal information in their responses. For example, in response to the maternity leave question, one redditor shared their experience, stating, “I live in NY so was able to take 6 weeks disability and 10 weeks of paid family leave. Disability pays very little (175 a week) and paid family leave was 65% of my salary.”

We observe another interesting norm in the r/parenting subreddit. Most posts include the child’s age (and sometimes gender) when describing their situations or questions, usually in forms like “5yo” or “14f”, where “14” indicates age and “f” stands for gender. This practice makes their questions clearer and more specific, helping them receive more relevant advice from other parents, as solutions to parenting issues often vary depending on the child’s age.

Study 2: User Study of Context-Aware PETs

In Study 1, we conducted quantitative analyses to explore the effects of contextual factors on users’ SD within a community. Our findings indicated that topics, audience, and peer effects influence users’ SD but these effects vary across communities. This variation aligns with observations in pre-

variables	parenting		projectmanagement		politicaldiscussion	
	dyadic	full	dyadic	full	dyadic	full
SD parent & SD reply	280.6(-, Y)	177.93(-, Y)	0.40(N)	4.09(N)	0.19(N)	0.74(N)
Emo SD parent & SD reply	99.31(+, Y)	181.15(+, Y)	0.17(N)	0.25(N)	2.40(N)	9.53(-, Y)
Info SD parent & SD reply	535.91(-, Y)	370.29(-, Y)	0.06(N)	5.67(N)	0.19(N)	0.74(N)

Table 3: Chi-squared tests for relationship between SD in parent post and child comments in dyadic and full length conversations. "Y" represents a statistically significant result after correction, and "+" indicates SD posts receive more respective replies than non-SD posts, while "-" indicates that non-SD posts receive more respective replies than SD posts.

vious work (Proferes et al. 2021) suggesting that subreddits have their own norms, cultures, and moderation practices, so that insights in one subreddit may not translate across contexts. In Study 2, we extend our investigation of contextual factors to the development of PETs and conduct a user study to assess the efficacy and utility of such a tool.

Participant Recruitment

Our survey was created with Qualtrics³ and distributed through Amazon MTurk⁴. All human subjects research was approved and exempted by IRB⁵. On MTurk, we hired master workers to ensure high-quality survey responses. Additionally, we required participants: (1) be Reddit account holders; (2) use a laptop as their primary Internet Device (our prototype is designed for Chrome).

Experimental Protocol

Our 6-part survey was adapted from (Chang et al. 2018) and (Bazarova and Choi 2014):

- Part 1: Basic demographic questions, e.g., age, gender, income, race/ethnicity, education, and occupation.
- Part 2: Participants asked what types of information they consider "personal" or "sensitive".
- Part 3: Participants asked about perceived group norms, i.e., frequency of SD for themselves and other community members.
- Part 4: We adopted the motivation question from Bazarova and Choi (2014) to investigate the motivation of participants' SD on Reddit, but separating the category "information storage and entertainment" into "information storage" and "entertainment" for finer granularity.
- Part 5: Participants asked to what extent they were concerned that their personal information could be misused or exposed by others. These were adopted from two verified questionnaire items from Chang et al. (2018).
- Part 6: Participants were asked to take on a given persona and suppose they were using our privacy-enhancing tool before submitting a post to a given subreddit. The Chrome extension detects if there is any personal information in the post title and content, and then notify users the detection result along with audience size and the rate

³<https://www.qualtrics.com>

⁴<https://www.mturk.com>

⁵The study (STUDY00020561) was reviewed and approved as exempt by the Institutional Review Board (IRB) at Pennsylvania State University for human subjects research.

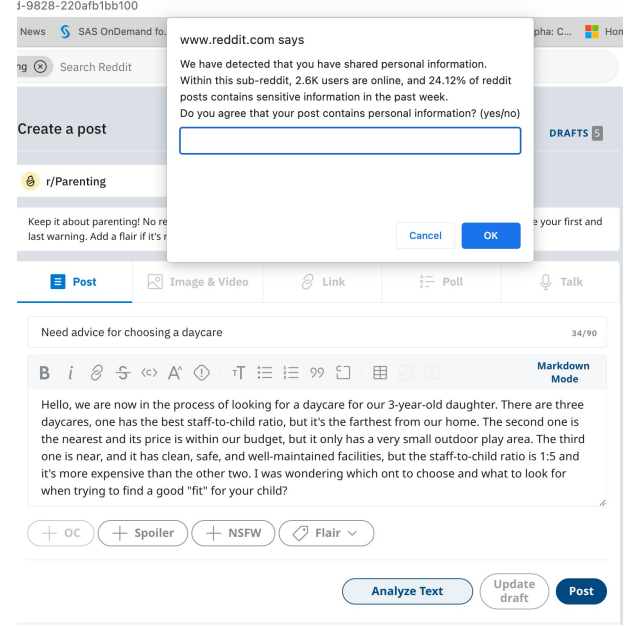


Figure 1: Screenshot from our prototype privacy-enhancing tool. The Chrome extension detects if personal information is present in the post title or content. A pop-up notifies users if SD is detected and provides approximate audience size and average rate of SD in that subreddit over the past week.

of SD over the past week in the targeted subreddit. It also collect users' feedback about whether they agree with the detection result. A screenshot is provided in Figure 1.

We designed three such scenarios to understand how participants embedded in different contexts would engage with our prototype privacy-enhancing tool.

- Scenario 1: The user is posting to the r/parenting subreddit seeking advice on choosing a daycare for their 3-year-old daughter, comparing three options based on staff-to-child ratio, proximity, outdoor play area, and cost.
- Scenario 2: The user is posting to the r/ProductManagement subreddit expressing concern over industry budget cuts and potential layoffs. They detail their experience at a mid-sized tech company in NY, mentioned their \$3k/-month mortgage and a young daughter, and seek advice on staying competitive and reducing the risk of job loss.
- Scenario 3: The user is posting to the r/politicaldiscussion subreddit expressing unwavering support for Pres-

Scenarios	Peer Disclosure Rate		
	25.0% $\pm$ 3.0%	75.0% $\pm$ 3.0%	total
scenario1	44	47	91
scenario2	47	45	92
scenario3	50	47	97
total	141	139	280

Table 4: Number of responses for each scenario and peer disclosure rate (*pdr*) setting.

ident Smith. They outline reasons for their support, citing President Smith’s focus on rebuilding the economy and handling the pandemic. They acknowledge differing opinions and invite others to share their views.

Full text of the instructions and scenarios has been provided in Supplemental Materials ⁶.

To understand how peer effects might influence participants’ willingness to self-disclose, we further introduced an independent variable (*pdr*) displaying a dummy average rate of SD in the given subreddit over the past week in each scenario. Participants were divided into two groups: one group was shown a relatively low average SD rate, with a random percentage generated in the range of 25.0% $\pm$ 3.0%, while the other group was shown a high average SD rate, with a percentage in the range of 75.0% $\pm$ 3.0%. The number of participants for each scenario and peer disclosure rate (*pdr*) setting are shown in Table 4.

Data Analysis

Data cleaning. After an initial round of surveys, we encountered bot attacks. In response, we introduced two logical validation questions within the survey to assess participants’ attentiveness. Furthermore, we manually checked the responses’ attitude, longitude, and IP addresses, excluding those with identical or near-identical values across these three attributes. A manual review process was also employed for answers to two open-ended questions, eliminating those with nonsensical content. Out of the 301 responses received, 18 were filtered out based on these criteria. An additional 3 participants requested their data points be removed from our analyses, resulting in 280 valid responses. Participants’ demographics are provided in Supplemental Materials.

Affinity diagramming. To analyze survey responses, we used an affinity diagram (Lucero 2015). Affinity diagram is a qualitative analysis method enabling designers and researchers to identify key aspects for improvement, prioritize user feedback, and improve the usability of a prototype tool. We initiated the process by constructing an affinity diagram on the MIRO whiteboard⁷. The lead annotator began by creating the diagram and grouping similar participant responses into categories using sticky notes. The second annotator reviewed all the codes, and whenever there was disagreement with the lead annotator (< 5% of instances), the two authors discussed to arrive at consensus. Subsequently, both the lead

and second annotators iteratively refined and reorganized the diagram through multiple rounds.

Findings

RQ2: How can privacy-enhancing technologies support improved contextual awareness of self-disclosure?

Association between deviating from group norms in personal disclosure and concern about online identification.

Privacy concerns are generally viewed as a major reason why users choose *not* to share personal information online. However, SD does not always pose a privacy risk. For example, users may believe their SD is less risky when others are also sharing similar information. To investigate this further, we asked participants to what extent they are concerned about being identified through disclosed personal information and the potential misuse of their personal information. We measured the difference between participants’ self-reported SD frequency and their perceptions of their peers’ SD frequency, naming this metric a user’s *frequency difference*. We then calculated correlations between each participant’s frequency difference and their level of concerns: (1) concern about being identified online ($r = 0.17, p = .004$) and (2) concern about personal information being misused ($r = 0.02, p = .64$). Our findings indicate that participants with a higher frequency difference in personal information disclosure, meaning they perceive themselves to disclose more than their peers, tend to be more concerned about being identified online.

Discrepancy between perceived sensitivity and willingness to share information.

Taking into account participants’ perceptions of sensitive information, we proceeded to investigate whether these perceptions aligned with their willingness to share such information. Among participants’ who consider a respective category as sensitive, the number of participants choosing not to submit a post containing that information was approximately the same as those choosing to submit, except for the first scenario (age as sensitive information) where more participants opted to submit. We asked participants the reason for their decision if they chose not to submit the post. Most frequently selected reasons were “because of the % of Reddit posts containing sensitive information in the past week” and “I didn’t realize there was personal information in my post, but now I have realized it after the tool notified me.”

We asked additional questions of participants who decided to submit a given post despite SD notification, to understand their reasoning. Through affinity diagramming, we identified four primary reasons: (1) Belief that sharing the data would not result in harm. For example, one participant stated, “*I don’t see any potential threat in posting this particular message. I don’t feel as though any of the information disclosed is sensitive.*”; (2) Perceived benefits from sharing the information. For example, one participant mentioned, “*for me i think the benefit oversee the risk;*”; (3) Observation of other users sharing similar information. One participant commented, “*Most of the posts have personal information, so I don’t feel alone.*”; (4) Trust in the community/plat-

⁶<https://shorturl.at/4CO2c>

⁷<https://miro.com>

form. One participant stated, “*It is very safe and secure. It has many options to keep our data’s privacy.*”

Supportive privacy-enhancing technologies. SD is often framed as self-privacy violation that takes place intentionally or unintentionally to achieve specific goals including relationship building, social connectivity, identity clarity, and social control (Abramova et al. 2017; Bazarova and Choi 2014; De Choudhury and De 2014). PETs have helped users become aware of their SD behavior and protect their data in an online environment (Ioannou et al. 2021). One study conducted a field trial using two types of privacy nudges: one that informed users about the audience of their posts and another that delayed the time of posting. They found that both nudges reduced SD; users revised their posts to avoid potential regret (Wang et al. 2014).

Prior work has shown that visual cues of social norms can also influence user SD. In other words, when people observe others sharing personal information, they are more likely to do the same (Spottswood and Hancock 2017). Likewise, perceived social norms can influence SD (Chang et al. 2016).

Our analyses of Reddit posts and comments suggest that context (e.g., peer’s SD behavior, topics of discussion) can serve as cues for users’ SD decisions. Building upon these insights and prior work (Wang et al. 2014; Kaushal et al. 2017; Schaub et al. 2015), our prototype tool exposes key contextual information, namely group norms and audience size, to support better informed SD. *Our design acknowledges that less sharing is not always better for a user and that, in as much as contextual elements may impact social rewards, they are critical to balanced decision making.*

Our survey asked participants to share their perceptions of our prototype and suggestions for improvement. Of 280 respondents, 239 (85.4%) either “strongly agree” or “somewhat agree” that the tool could enhance their privacy risk awareness. Additionally, 220 (78.6%) either “strongly agree” or “somewhat agree” that the tool could help them manage potential privacy risks. Of the 280, 119 participants responded that there was nothing they would modify about the tool. 11 participants’ answers were not interpretable; 13 responded that the tool was not helpful for them; 10 left responses blank; and, 28 simply offered compliments on the tool⁸. Using an affinity diagram (Lucero 2015), we identified two main themes amongst responses: (1) enhancing tool transparency; and (2) user-centered design.

Enhancing tool transparency for improved user trust and understanding. Many participants reflected a desire for greater transparency in our tool, including precise indications of what personal information is contained in their posts. They asked for detailed explanations of how the tool works internally, e.g., how our algorithm decides whether posts contain personal information, and requested guidance on how to avoid privacy risks.

Specify personal information in the post. 52 participants (of 119) wished the tool would have clearly identified which

pieces of personal information was contained in their post and determine what could be identifiable. P15 noted,

It would be more helpful if the tool could specify exactly what the personal information is [and] highlight personal information to make it easier and more convenient for me to modify the post.

P32 wanted not only to be informed about what contents are personal but also what contents can be identifiable: “*I would have liked to see what it flagged as identifiable information.*”

Explain the tool’s functionality. Four participants wanted to know how the tool made its assessment, specifically *why* and *how* it concluded that what the participant was sharing was personal or sensitive information. For example, P227 asked “*how [the algorithm] categorizes this information, and how it decides what is considered personal.*”

Highlight potential consequences. Participants wished our tool could explain the potential consequences of posting sensitive information. For example, participants suggested the tool “*highlight what the sensitive information was on the page and tell the person how badly that can be to share that.*” (P137) and another participant similarly stated,

[The tool] can increase my privacy risk awareness by giving me the chance to think through potential risks I may face when sharing information online. (P38)

Offer guidance on navigating risk tradeoffs. Thirteen participants recommended that the tool guide users on how to avoid potential privacy risks while preserving the intended meaning of their posts. For example, participants noted,

The tool can [prompt] me to consider the consequences of my actions and leading me to take appropriate steps to protect my privacy. I would like to see the tool include more specific advice on how to protect my privacy. (P38)

Likewise, P268 wanted “*to have an option to reword it so it has less personal information but makes the same point.*”

User-centered design. Participants’ recommendations for improving the tool included enhancing user agency, improving the user interface (UI), and expanding the tool’s functionality and features.

User agency. Four participants requested a greater sense of agency from our privacy-enhancing tool. For example, they wanted to have control over turning notifications on or off in specific environments. P273 wanted the tool to have “*the ability to turn it on or off only for certain subreddits would be nice,*” rather than in all Reddit communities they engage in. Similarly, P207 thought the tool could be less intrusive “*if it can be altered with filters.*”

User interface design. Twenty-four participants suggested changes and improvements to the tool’s user interface (UI) design, including modifying font colors, using bold text, changing to clickable buttons, and rearranging the location of notifications. For example, P74 noted,

It is an interesting tool, maybe they could improve the layout because this text box (pop-up) in the middle of the screen is a bit scary for many users.

⁸As our participants may have mentioned more than one topic in their response, total number of utterances can be inconsistent with our total number of participants.

Expand the tool's features. Six users expressed a desire for additional features, including details about how many users viewed their posts, and the total number of users visiting a subreddit each day. One participant expressed interest in utilizing the tool across other social media platforms.

Discussion

Our mixed-methods approach sheds light on how contextual information can be integrated into the development of PETs. Following, we discuss the outcomes of our studies and how they inform the design of PETs to support improved SD decision-making. We note the limitations of this work and propose future directions for research.

Design of PETs

The majority (78.6% and 85.4%, respectively) of our participants stated that our prototype tool could help them manage potential privacy risks and enhance their privacy risk awareness. Our work highlights key considerations for the design of PETs to support users' SD decision making.

Contextual factors and self-disclosure (SD). We observed that SD varies across topics of discussion and social circles. Prior literature also supports this observation (Wang, Burke, and Kraut 2016; Umar, Squicciarini, and Rajtmajer 2019). We also observe the central role of SD in personal and relationship conversations, highlighting heightened privacy risks in these communities. This suggests implementing privacy controls tailored to specific communities, with differing sensitivity thresholds.

Our work also suggests the importance of audience size for SD decisions (Wang, Burke, and Kraut 2016; Vitak and Kim 2014). Providing users with insights into their audience's characteristics would allow individuals to better make informed self-disclosure decisions based on their social connections and desired privacy levels.

Furthermore, our work underlines the influence of group norms on SD, noting that users seek to align with the actions of their peers in a community. Exposing and making explicit community norms of SD through PETs can help users to understand processes that may be implicit but play a critical role in user privacy, i.e., via the framework of contextual integrity (Nissenbaum 2004).

Design specifications. Survey responses shed light on design implications. Designers of PETs should provide SD notifications that highlight specific parts of the text containing personal or sensitive information. Information with different levels of sensitivity could be highlighted with, e.g., a temperature scale (Tahaei, Frik, and Vaniea 2021). Many participants sought suggestions for alternate text which can make posts less sensitive while preserving their meaning.

Designers should consider the *explainability* of PETs for improved user trust (Roundtree 2023; Liao and Sundar 2022). Our findings indicate that participants desire greater transparency regarding the algorithms driving our insights, e.g., how they categorizes information and why certain information is considered personal. While explainability of outputs from machine learning algorithms is a persistent

challenge, designers should keep in mind that integrating available solutions where possible is desirable.

Our findings indicate that PETs for SD should clarify the consequences of posting sensitive information and provide guidance for users on how to avoid privacy risks. This could be achieved, for example, by implementing *hovering* features in the tool. Consistent with prior work (Limerick, Coyle, and Moore 2014; Schaub et al. 2015) our participants also sought greater control over our prototype tool, e.g., allowing them to turn notifications on and off for specific subreddits rather than across all Reddit communities in which they participate. Providing these options not only empowers users but also enables them to make meaningful decisions (Schaub et al. 2015).

Lastly, designers can enhance our UI design by using visual cues to highlight more important information with different text colors (e.g., red), or by adjusting the color of statistical information based on whether the numbers are high or low to capture users' attention. Additionally, our tool can foster a better user experience by changing our interface from typing to clickable buttons.

Limitations and Future Work

To offer participants a tangible context when asking the scenario questions, we asked all participants to take on the given persona. However, the provided persona is not gender inclusive. This lack of inclusivity may introduce bias into our results. Future work could involve a more immersive approach to study user self-disclosure behaviors through our developed context-aware browser extension.

In addition, we treat all Reddit accounts as having the same level of anonymity. Research has studied the phenomenon of "throwaway accounts" on Reddit and indicated that there are differences in posting behaviors between pseudonymous and throwaway accounts (Leavitt 2015; Ammari, Schoenebeck, and Romero 2019). A study considering varying levels of anonymity would address this limitation.

Conclusion

Our work adopted a mixed-methods approach to investigate the influence of contextual factors on users' online self-disclosure and explore how they might inform the design of PETs to support improved SD decision-making. Our results align with existing research, demonstrating the influence of topics, audience, and norms on self-disclosure. Additionally, our prototype context-aware browser extension showcases the technology's potential to enhance users' awareness of self-disclosure. While our study offers insights, we recognize its limitations and propose future directions. Longitudinal and interactive data collection and studies considering varying levels of anonymity could yield a more comprehensive understanding of self-disclosure behaviors across different contexts. Our work contributes to the broader understanding of how context shapes online self-disclosure behaviors. As technologies continue to evolve, our findings can inform the design of tools that empower users to make informed sharing decisions while safeguarding their privacy.

Acknowledgements

Work of TD, JK, and SR was partially funded by the National Science Foundation Award 2247723. Work of AS was supported by (while serving at) the National Science Foundation. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- Abramova, O.; Wagner, A.; Krasnova, H.; and Buxmann, P. 2017. Understanding self-disclosure on social networking sites-a literature review.
- Acquisti, A.; Adjerid, I.; Balebako, R.; Brandimarte, L.; Cranor, L. F.; Komanduri, S.; Leon, P. G.; Sadeh, N.; Schaub, F.; Sleeper, M.; et al. 2017. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)*, 50(3): 1–41.
- Acquisti, A.; and Gross, R. 2006. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *International workshop on privacy enhancing technologies*, 36–58. Springer.
- Ammari, T.; Schoenebeck, S.; and Romero, D. 2019. Self-declared throwaway accounts on Reddit: How platform affordances and shared norms enable parenting disclosure and support. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW): 1–30.
- Andalibi, N.; Ozturk, P.; and Forte, A. 2017. Sensitive self-disclosures, responses, and social support on Instagram: The case of #depression. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 1485–1500.
- Attrill, A.; and Jalil, R. 2011. Revealing only the superficial me: Exploring categorical self-disclosure online. *Computers in Human Behavior*, 27(5): 1634–1642.
- Backes, M.; Berrang, P.; Goga, O.; Gummadi, K. P.; and Manoharan, P. 2016. On profile linkability despite anonymity in social media systems. In *Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society*, 25–35.
- Bak, J.; Kim, S.; and Oh, A. 2012. Self-disclosure and relationship strength in twitter conversations. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 60–64.
- Bakshy, E.; Rosenn, I.; Marlow, C.; and Adamic, L. 2012. The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, 519–528.
- Bazarova, N. N.; and Choi, Y. H. 2014. Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of communication*, 64(4): 635–657.
- Chang, D.; Krupka, E. L.; Adar, E.; and Acquisti, A. 2016. Engineering information disclosure: Norm shaping designs. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 587–597.
- Chang, Y.; Wong, S. F.; Libaque-Saenz, C. F.; and Lee, H. 2018. The role of privacy policy on consumers' perceived privacy. *Government Information Quarterly*, 35(3): 445–459.
- Cheung, C.; Lee, Z. W.; and Chan, T. K. 2015. Self-disclosure in social networking sites: the role of perceived cost, perceived benefits and social influence. *Internet Research*, 25(2): 279–299.
- Cheung, C. M.; Chiu, P.-Y.; and Lee, M. K. 2011. Online social networks: Why do students use facebook? *Computers in human behavior*, 27(4): 1337–1343.
- Choi, Y. H.; and Bazarova, N. N. 2015. Self-disclosure characteristics and motivations in social media: Extending the functional model to multiple social network sites. *Human Communication Research*, 41(4): 480–500.
- Clark-Gordon, C. V.; Bowman, N. D.; Goodboy, A. K.; and Wright, A. 2019. Anonymity and online self-disclosure: A meta-analysis. *Communication Reports*, 32(2): 98–111.
- Collins, N. L.; and Miller, L. C. 1994. Self-disclosure and liking: a meta-analytic review. *Psychological bulletin*, 116(3): 457.
- Coopamootoo, K. P. 2020. Usage patterns of privacy-enhancing technologies. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 1371–1390.
- Cropanzano, R.; and Mitchell, M. S. 2005. Social exchange theory: An interdisciplinary review. *Journal of management*, 31(6): 874–900.
- Crowley, J. 2019. Self-Disclosure. In *Oxford Research Encyclopedia of Communication*.
- De Choudhury, M.; and De, S. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Proceedings of the international AAAI conference on web and social media*, volume 8, 71–80.
- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Dwyer, C.; Hiltz, S.; and Passerini, K. 2007. Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. *AMCIS 2007 proceedings*, 339.
- Ferreyra, N. D.; Meis, R.; and Heisel, M. 2017. Online Self-disclosure: From Users' Regrets to Instructional Awareness. In *1st International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE)*, 83–102. Springer International Publishing.
- Garg, V.; Patil, S.; Kapadia, A.; and Camp, L. J. 2013. Peer-produced privacy protection. In *2013 IEEE International Symposium on Technology and Society (ISTAS): Social Implications of Wearable Computing and Augmented Reality in Everyday Life*, 147–154. IEEE.
- Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.
- Gürses, S.; and Berendt, B. 2010. PETs in the surveillance society: a critical review of the potentials and limitations of

- the privacy as confidentiality paradigm. *Data Protection in a Profiled World*, 301–321.
- Ioannou, A.; Tusssyadiah, I.; Miller, G.; Li, S.; and Weick, M. 2021. Privacy nudges for disclosure of personal information: A systematic literature review and meta-analysis. *PloS one*, 16(8): e0256822.
- Jaidka, K.; Singh, I.; Lu, J.; Chhaya, N.; and Ungar, L. 2020. A report of the CL-Aff OffMyChest Shared Task: Modeling Supportiveness and Disclosure. In *Proceedings of the AAAI-20 Workshop on Affective Content Analysis*. New York, USA: AAAI.
- Joinson, A. N. 2001. Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European journal of social psychology*, 31(2): 177–192.
- Jourard, S. M. 1971. Self-disclosure: An experimental analysis of the transparent self.
- Kaushal, R.; Chandok, S.; Jain, P.; Dewan, P.; Gupta, N.; and Kumaraguru, P. 2017. Nudging nemo: Helping users control linkability across social networks. In *Social Informatics: 9th International Conference, SocInfo 2017, Oxford, UK, September 13-15, 2017, Proceedings, Part II* 9, 477–490. Springer.
- Kim, J.; and Dindia, K. 2011. Online self-disclosure: A review of research. *Computer-mediated communication in personal relationships*, 156–180.
- Knijnenburg, B. P.; and Kobsa, A. 2013. Making decisions about privacy: information disclosure in context-aware recommender systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 3(3): 1–23.
- Krasnova, H.; Spiekermann, S.; Koroleva, K.; and Hildebrand, T. 2010. Online social networks: Why we disclose. *Journal of information technology*, 25(2): 109–125.
- Leavitt, A. 2015. "This is a Throwaway Account" Temporary Technical Identities and Perceptions of Anonymity in a Massive Online Community. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 317–327.
- Li, D. C. 2011. Online social network acceptance: a social perspective. *Internet research*, 21(5): 562–580.
- Liao, Q. V.; and Sundar, S. S. 2022. Designing for responsible trust in AI systems: A communication perspective. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1257–1268.
- Limerick, H.; Coyle, D.; and Moore, J. W. 2014. The experience of agency in human-computer interactions: a review. *Frontiers in human neuroscience*, 8: 643.
- Lucero, A. 2015. Using affinity diagrams to evaluate interactive prototypes. In *Human-Computer Interaction-INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part II* 15, 231–248. Springer.
- Ma, X.; Hancock, J.; and Naaman, M. 2016. Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 3857–3869.
- McKenna, K. Y.; and Bargh, J. A. 2014. Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. In *Personality and Social Psychology at the Interface*, 57–75. Psychology Press.
- Nissenbaum, H. 2004. Privacy as contextual integrity. *Wash. L. Rev.*, 79: 119.
- O'Hagan, J.; Saeghe, P.; Gugenheimer, J.; Medeiros, D.; Marky, K.; Khamis, M.; and McGill, M. 2023. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(4): 1–35.
- Petronio, S.; Child, J. T.; and Hall, R. D. 2021. Communication privacy management theory: Significance for interpersonal communication. In *Engaging theories in interpersonal communication*, 314–327. Routledge.
- Proferes, N.; Jones, N.; Gilbert, S.; Fiesler, C.; and Zimmer, M. 2021. Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society*, 7(2): 20563051211019004.
- Rashidi, Y.; Kapadia, A.; Nippert-Eng, C.; and Su, N. M. 2020. "It's easier than causing confrontation": Sanctioning Strategies to Maintain Social Norms and Privacy on Social Media. *Proceedings of the ACM on human-computer interaction*, 4(CSCW1): 1–25.
- Roundtree, A. K. 2023. AI Explainability, Interpretability, Fairness, and Privacy: An Integrative Review of Reviews. In *International Conference on Human-Computer Interaction*, 305–317. Springer.
- Schaub, F.; Balebako, R.; Durity, A. L.; and Cranor, L. F. 2015. A design space for effective privacy notices. In *Eleventh symposium on usable privacy and security (SOUPS 2015)*, 1–17.
- Shelton, M.; Lo, K.; and Nardi, B. 2015. Online media forums as separate social lives: A qualitative study of disclosure within and beyond Reddit. *ICConference 2015 Proceedings*.
- Spottswood, E. L.; and Hancock, J. T. 2017. Should I share that? Prompting social norms that influence privacy behaviors on a social networking site. *Journal of Computer-Mediated Communication*, 22(2): 55–70.
- Stutzman, F.; Capra, R.; and Thompson, J. 2011. Factors mediating disclosure in social network sites. *Computers in Human Behavior*, 27(1): 590–598.
- Tahaei, M.; Friks, A.; and Vaniea, K. 2021. Deciding on personalized ads: Nudging developers about user privacy. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, 573–596.
- Tajfel, H. 1982. Social psychology of intergroup relations. *Annual review of psychology*, 33(1): 1–39.
- Tang, J.-H.; and Wang, C.-C. 2012. Self-disclosure among bloggers: Re-examination of social penetration theory. *Cyberpsychology, Behavior, and Social Networking*, 15(5): 245–250.

- 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): 141–161.
- Turner, J. C.; Hogg, M. A.; Oakes, P. J.; Reicher, S. D.; and Wetherell, M. S. 1987. *Rediscovering the social group: A self-categorization theory*. basil Blackwell.
- Umar, P.; Squicciarini, A.; and Rajtmajer, S. 2019. Detection and analysis of self-disclosure in online news commentaries. In *The World Wide Web Conference*, 3272–3278.
- Utz, S.; and Krämer, N. C. 2009. The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 3(2).
- Vitak, J.; and Kim, J. 2014. ” You can’t block people of fline” examining how facebook’s affordances shape the disclosure process. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, 461–474.
- Walther, J. B. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research*, 23(1): 3–43.
- Wang, Y.; Leon, P. G.; Acquisti, A.; Cranor, L. F.; Forget, A.; and Sadeh, N. 2014. A field trial of privacy nudges for facebook. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 2367–2376.
- Wang, Y.; Leon, P. G.; Scott, K.; Chen, X.; Acquisti, A.; and Cranor, L. F. 2013. Privacy nudges for social media: an exploratory Facebook study. In *Proceedings of the 22nd international conference on world wide web*, 763–770.
- Wang, Y.-C.; Burke, M.; and Kraut, R. 2016. Modeling self-disclosure in social networking sites. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, 74–85.
- Wisniewski, P. J.; Knijnenburg, B. P.; and Lipford, H. R. 2017. Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of human-computer studies*, 98: 95–108.
- Zhao, C.; Hinds, P.; and Gao, G. 2012. How and to whom people share: the role of culture in self-disclosure in online communities. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 67–76.
- Zhou, T. 2011. Understanding online community user participation: a social influence perspective. *Internet research*, 21(1): 67–81.