Who Shares What with Whom?
Information Sharing Preferences
in the Online and Offline Worlds

Chang Liu[1(✉)], Hsiao-Ying Huang[1], Dolores Albarracin[2],
and Masooda Bashir[3]

[1] Illinois Informatics Institute, University of Illinois at Urbana-Champaign,
Urbana, USA
`{cliul00,hhuang65}@illinois.edu`
[2] Department of Psychology, University of Illinois at Urbana-Champaign,
Urbana, USA
`dalbarra@illinois.edu`
[3] School of Information Sciences, University of Illinois at Urbana-Champaign,
Urbana, USA
`mnb@illinois.edu`

Abstract. Today people reveal a substantial amount of personal information both online and offline. Although beneficial in many aspects, this exchange of personal information may pose privacy challenges if the information is disseminated outside the originally intended contexts. Through an online survey, this study investigates people's online and offline information sharing preferences in a comparative fashion. Our analysis reveals that people generally have similar sharing preferences in online and offline contexts, except that they have different preferences for sharing information with their friends and family offline than they do for sharing with personal networks online. We also found that people share their gender and ethnicity less online than offline. Moreover, sharing religious affiliation was similar to sharing daily activities offline, whereas it was similar to sharing political beliefs online. Our findings corroborate Nissenbaum's (2011) theory of contextual integrity and shed light on preferences for sharing certain information with certain recipients.

Keywords: Information sharing behavior · Information sharing preferences
Information privacy · Privacy on cyberspace

1 Introduction

People reveal a substantial amount of personal information as they engage in their daily activities online and offline, including sending and receiving packages or emails, applying for jobs or schools, shopping on a website or at a nearby grocery store, and subscribing to a digital or paper-based magazine. Information scholars have yet to understand information sharing in the digital era in a comprehensive way. Considerable benefits stem from data-driven technologies and services that exploit personal information, but these same platforms can pose serious privacy challenges A recent example

is the 2017 Equifax data breach, in which sensitive information, including social security and driver's license numbers, were stolen from 143 million U.S. customers [1]. Unsurprisingly, people report increasing concerns about the privacy of their information [2]. Due to extensive social impacts of privacy breaches in the digital era, privacy has become an important subject of research in various areas, including legal, social, technical, and psychological aspects of cybersecurity.

Privacy should be a consideration when information is taken out of its original intended context and is shared and used in other contexts [3, 4]. Particularly online, individuals often lose control over their information once they have handed it to a recipient, whether an individual or entity [5]. Further, people are often unclear as to how their data will be used and for what purposes [5]. Nissenbaum [4] argues for the importance of contextual considerations for people's information sharing, using the term "contextual integrity" in her seminal work about information privacy. She maintains that the informational norms and general moral principles that have been established in various physical-world contexts should not be discarded as our communications and interactions step into the cyberspace. For instance, when we give out personal health information in online health-related platforms, users should be allowed to expect their information to be treated in the same way as it would in the traditional healthcare context. Following the idea of contextual integrity, we propose that understanding people's online and offline preferences for information sharing can provide insights into today's privacy challenges as well as suggest useful implications for public policy and the design of technologies for information privacy and security.

This study investigated people's online and offline information sharing preferences in a comparative fashion. To best of our knowledge, this study is the first to test the idea of contextual integrity with empirical data and is consistent with a long tradition of specifying the target, action, context, and time (TACT) of a behavior or goal [6]. Our study is also unique in comparing people's online information-sharing behavior with offline behavior as a benchmark. Although prior studies [7–12] have suggested that information sharing preferences depend on the recipients and the context, very few studies, if any, have investigated this subject in a systematic way. We have conducted an online survey and present the results and implications in the following sections.

2   Related Works

People may consider certain types of information to be more sensitive and private than others. Khalil and Connelly [7] explored people's sharing patterns in a telephony situation and found that they tended to share more of certain types of information. People also exhibit multiple different sharing behaviors for certain types of information. Benisch et al. [8] found that people's willingness to share their locations depended on the time of the day, the day of the week, and the location. Olson et al. [9] asked 30 participants to indicate their willingness to share 40 different types of personal information with 19 different entities. They found that people's information sharing preference varied as a function of the entities that would receive the information. Patil and Lai [10] studied MySpace users' information sharing preferences and reported that the users tended to assign privacy permissions by groups such as "family," "friends," and

"project members." They also found that, although the users considered location to be the most sensitive information, they were comfortable sharing it with their coworkers during business hours. Lederer, Mankoff, and Dey [11] reported that people's information sharing preferences tended to be consistent for identical inquirers in different situations, and for different inquirers in identical situations. An international study by Huang and Bashir [12] also revealed cultural differences in online users' information sharing preferences. These studies all suggest that information sharing preferences depend not merely on the type of information but also on the entities and contexts in which the information is being shared. Nevertheless, to the best of our knowledge, no prior studies have sufficiently investigated how people's information sharing preferences may be similar or different across online and offline contexts. We believe that understanding information sharing preferences can provide useful insights into the information sharing behavior.

Using the term "contextual integrity," Nissenbaum [4] indicates that that the informational norms and general moral/principles that have been established in various physical-world social contexts should not be discarded as our communications and interactions step into the digital space. For instance, when we give out personal health in-formation to online health-related platforms, we should be able to expect our information to be treated in the same way as it would in the traditional healthcare context. Nissenbaum [4] suggests that, to achieve better relationships with individuals, it is necessary for data collectors to establish a framework of supporting assurances that enable users to have faith in collectors' goals and practices. People who provide their personal information for healthcare or human subject research, for example, are equally unlikely both to fully understand the details of how the data will be used and to have faith in the collecting entities. Solove [3] suggests requiring data collectors to engage in fiduciary relationships – the relationships observed between doctors and patients, lawyers and clients, or corporate officials and shareholders – with individuals. Maguire et al. [13] propose a context-aware metadata-based architecture that can help users to express information about and thus guide sharing preferences and usages for different contexts through metadata information about users' preferences. They point out that understanding the nuances and contexts remains a challenge to address in the future.

Based on existing works, our study examined people's expectations about what they should share, and with whom, and explored how these expectations compare across the traditional, offline, physical-world interactions and the digital, online interactions. Adapting the famous question posed by the Yale School of social psychologists analyzing communication processes – "Who says what to whom?" – our research question was: What is shared with whom in which world?

3   Methodology

To address the research question, we conducted an online survey between August 2016 and September 2017. The survey link was distributed through Amazon Mechanical Turk [14, 15]. Only participants within the United States were recruited. In the survey, participants were asked to answer questions about their information sharing preferences

in online and offline contexts and their demographics. Each participant received $1 US dollar for completing the survey.

3.1   Survey Design and Measurement

We developed a questionnaire to study participants' information sharing preferences. We generated 28 types of information that are commonly shared by individuals during daily activities. For each type of information, we provided examples to help participants understand more specifically what was being discussed. For instance, we included "home address, phone number, email address, etc." as the examples of "contact information" and "psychological and emotional well-being, depression, anxiety etc." as the examples for "mental health information." Then participants were asked to indicate which entities they would like to share with these 28 types of information in online and offline contexts respectively. Considering the different nature of the online and offline world, we generated 16 offline entities and 20 online entities. Similarly, we provided concrete examples for these entities so that participants could have a better sense about which entities they shared. For instance, we presented "retail workers, transportation workers, food service workers, etc." as the examples of "service industry workers" in offline context and "online banking services, Chase Online, Bank of America Online Banking, etc." as the examples of "bank and financial institutions" in online contexts. For this paper, we focused on comparing people's sharing preferences for 7 entities in both online and offline contexts, including informal social group, employers, government, healthcare providers, educational institutions, financial institutions, and business services.

3.2   Survey Design and Measurement

A total of 201 participants completed the survey. 112 participants (55.7%) were female and 89 participants (44.3%) were male. Participants were from different age groups: 14.4% were between 18 to 24 years old; 45.8% were between 25 to 34 years old; 21.4% were between 35 to 44 years old; 12.4% were between 45 to 54 years old; 6.0% were 55 years old or older. For education, 9.0% of the participants were high school graduates; 80.6% had college degrees; 10.0% had advanced degrees. We recruited participants from the U.S. based on U.S. Census Regional Division [16], including New England (5.5%), Mid Atlantic (16.4%), EN Central (14.4%), WN Central (3.5%), South Atlantic (17.9%), ES Central (5.0%), WS Central (10.0%), Mountain (7.5%), and Pacific (19.9%).

4   Results

Our analysis explored people's information sharing preferences with different types of entities in the online and offline worlds. We categorized the offline and online entities from the questionnaire into seven corresponding groups: (1) informal social groups (friends and colleagues), (2) employers, (3) government, (4) healthcare providers, (5) educational institutions, (6) financial institutions, and (7) business services. We used

a heat map (Fig. 1) to visualize the participants' sharing preferences for the 28 types of information with the aforementioned recipient groups. Figure 1 visualizes the number of respondents who expressed willingness to share each type of information with each category of entity.
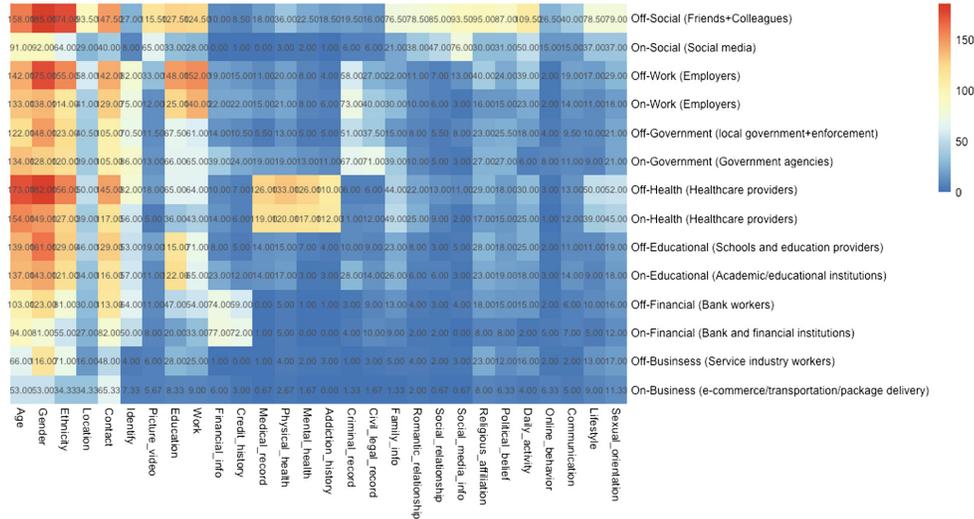

Fig. 1. Visualization of information sharing preferences in the online and offline contexts.

4.1   Overview of Sharing Preferences in Online and Offline Context

As exhibited in Fig. 1, we found that people tend to share their age, gender, and ethnicity more than other types of information. We also found that people are inclined to share certain kinds of information with certain entities. For instance, more people share their health-related information, including medical record, physical health, mental health, and addiction history with their healthcare providers in both offline and online contexts. Also, more people share their educational and working information with their employers. This suggests that people may have specific preferences for sharing certain information. In other words, when sharing information, people may tend to distinguish whether it makes sense for certain entities to have the information. We further compared whether certain types of information are shared more in offline or online contexts using a paired t-test. The result shows that people are more willing to share their gender (t = 4.38, p = .005) and ethnicity (t = 2.71, p = .035) in the offline world than in the online world.

4.2   Information Sharing Preferences for Online and Offline Entities

In order to identify people's preferences for sharing entities, we conducted a cluster analysis for all sharing entities in online and offline contexts. As displayed in Fig. 2, unsurprisingly, most online entity groups were grouped as clusters with their physical-world counterparts, except for the recipient of "informal social groups." This indicates
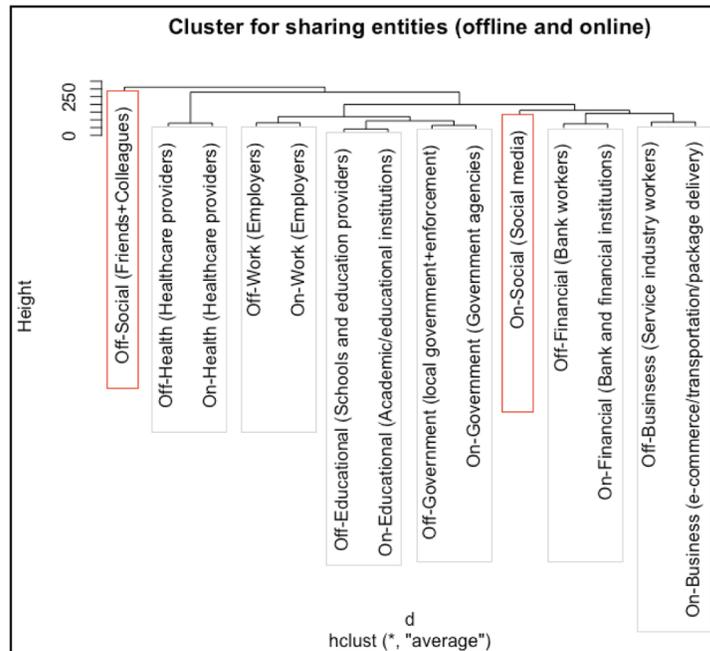
Fig. 2. Cluster analysis for offline and online sharing entities.

that people have similar sharing preferences for the same type of recipient in offline and online contexts. Interestingly, the online social group is close to the financial and business domain. This means that people tend to have similar sharing preferences for online social media and financial institutions as well as business services. On the other hand, people have rather different sharing preferences for the informal social groups in the offline context.

4.3   Sharing Preferences for Information Types

We further explored people's sharing preferences for 28 types of information online and offline. According to the cluster analysis (Fig. 3a and b), we identified 7 clusters that people have similar sharing preferences in both contexts including demographic information (age, gender, ethnicity, contact information), professional information (work and education), legal information (criminal records and civil legal records), financial information (finance and credit score), online behavior/communication, location, and health information (medical record, mental health, addiction history, and physical health). It is worth noting that for health-related in-formation, people tend to have similar sharing preferences for their medical records and mental health in the offline world. In the online world, people have similar sharing preferences for their medical records and physical health instead of mental health, which is grouped with addiction history in the offline context.

Our analysis also reveals that people have different offline and online sharing preferences for certain types of information. For instance, the results show that people
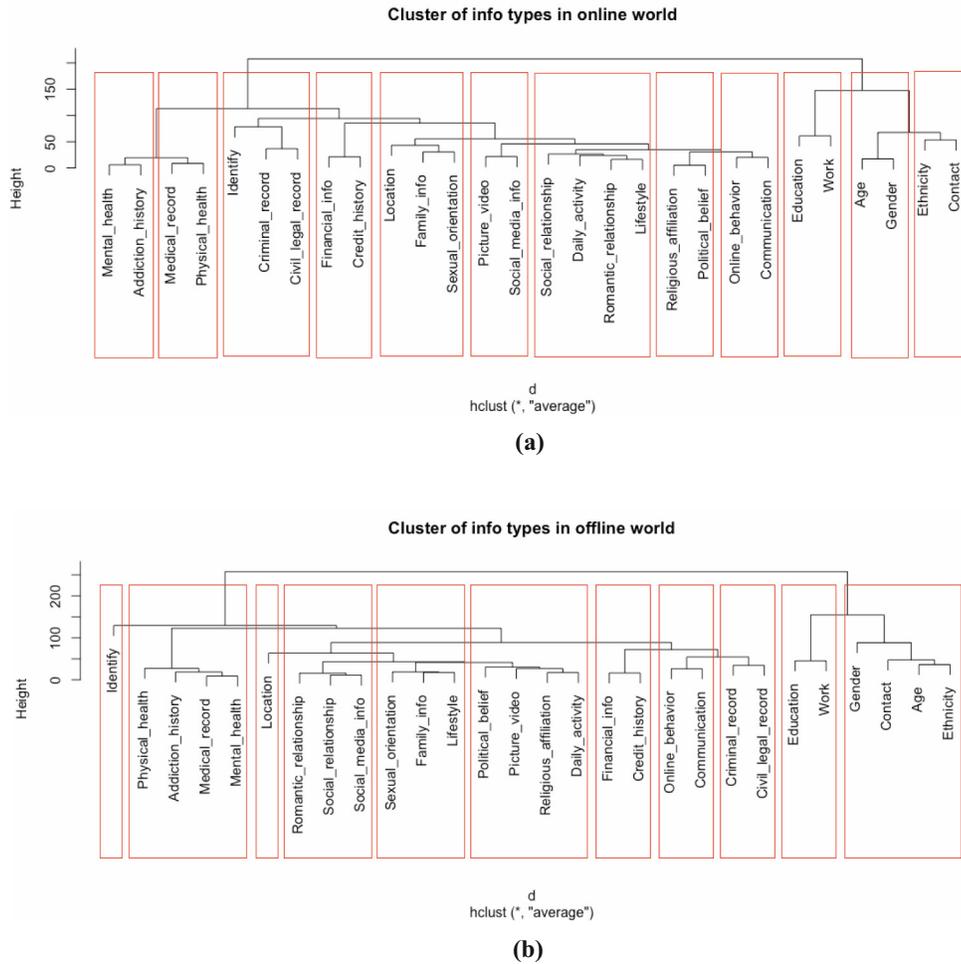
Fig. 3. (a) Hierarchical cluster analysis of information types in the online context. (b) Hierarchical cluster analysis of information types in the offline context.

are inclined to have similar sharing patterns for daily activities and religious affiliation in the offline world. However, the religious affiliation is clustered with political belief in the online world, meaning that people tend to share their religious affiliation and political belief with similar entities online. In addition, the result shows that social media information is clustered with pictures and videos online. That is, people have similar sharing preferences for their social media information and pictures as well as videos online, which corresponds to the phenomenon in which people share abundance of pictures and videos on social media.

Another interesting finding is that family history information is clustered with lifestyle offline, but with sexual orientation online. This may suggest that people have a distinctive mindset when it comes to offline and online sharing of family history. Furthermore, we found that people have specific sharing preferences for location and

identity in the offline context; therefore, these two types of information have no particular clusters. In the online context, identity is clustered with criminal and civil legal records, and location is clustered with family history and sexual orientation.

5   Discussion

5.1   Sharing Preferences for Recipient Entities

We explored people's sharing preferences for 7 types of recipient entities, including informal social groups (e.g., friends and colleagues), employers, government, healthcare providers, educational institutions, financial institutions, and business services. We found that people have similar sharing preferences for the same type of recipients in both online and offline contexts, except for informal, communal social groups. This means that, no matter whether they are in online or offline contexts, people expect to have similar information sharing experiences with formal entities with which people have exchange instead of communal relationships, such as employer, healthcare providers, and financial institutions. This finding corroborates Nissenbaum's concept of contextual integrity [4]. People hold normative expectations for types of recipients. Based on those expectations, people decide whether to share their information.

Interestingly, unlike other recipient entities, people have distinct sharing preferences for online and offline social groups. This finding may further suggest that people view their informal social group in the online world as different from their informal social groups in the offline world. A possible explanation is that the informal social group in the online context (e.g., social media) is a new territory that does not exist in the offline world. Therefore, people may develop new sharing preferences for their online social groups instead of applying the existing sharing preferences based on their offline relationships.

5.2   Sharing Preferences for Types of Information

We further examined people's sharing preferences for 28 types of information in online and offline contexts. Our findings revealed that more people share their gender and ethnicity in offline than online interactions. A potential explanation for this high prevalence of communicating gender and ethnicity information in the offline context is because it is inevitable. It is generally difficult if not impossible for people not to share their gender and ethnicity offline. However, when people can control whether to share their gender and ethnicity online, the willingness to share this information de-creases greatly.

In addition, we found that people have different preferences for sharing their religious affiliation online and offline. Offline, people share their religious affiliation and daily activities with similar recipient entities. In the online context, the religious affiliation and political beliefs are shared with similar entities. These results may suggest that in the offline world, religious affiliation is tied to people's daily activities. However, in online interactions, religious affiliation becomes more like a marker of

ideology that behaves like political beliefs. Thus, people share religious affiliation in different ways online and offline, displaying distinctive mindsets about online and offline sharing for the same type of information.

5.3   Limitations and Future Directions

We acknowledge several limitations in this study. First, our results can only be generalized to the population from which participants were sampled. Since we recruited our participants via Amazon Mechanical Turk, most were younger and more educated than the general population, and from the west Pacific area of the U.S. In this study, we were unable to conclude whether the older or less educated populations would also express the same sharing preferences, but future studies should replicate our findings with more diverse groups. In addition, our study only provides an understanding of people's information sharing preferences without probing into the processes underlying the sharing preferences. The privacy research agenda should include uncovering those processes in the years to come.

6   Conclusion

This empirical study investigated information sharing preferences in online and offline contexts. Our results revealed several intriguing findings, including:

- People, in general, have similar sharing preferences for the same type of recipient entities in both online and offline contexts, except for informal social groups such as friends and colleagues.
- People's sharing preferences for online informal social groups differ from offline informal social groups.
- Fewer people share their gender and ethnicity in the online context than in the offline context.
- Sharing religious affiliation offline resembles sharing about daily activities, and sharing religious affiliation online resembles sharing political beliefs.
- Sharing medical records resembles sharing mental health information in the online context, whereas sharing medical records resembles sharing physical health information in the offline context.

Our findings corroborate Nissenbaum's theory of contextual integrity in information privacy and shed light on people's preferences for sharing certain types of information with specific recipient entities. While there is still a long way to go to gain a comprehensive understanding of people's preferences for information sharing and privacy, we hope this work can bring interesting ideas to this community and generate more discussions about this topic.

References

1. Bernard, T.S., Hsu, T., Perlroth, N., Lieber, R.: Equifax Says Cyberattack May Have Affected 143 Million in the U.S. https://www.nytimes.com/2017/09/07/business/equifax-cyberattack.html?mcubz=3. Accessed 18 Sep 2017
2. Madden, M., Rainie, L., Zickuhr, K., Duggan, M., Smith, A.: Public perceptions of privacy and security in the post-snowden era. Pew Research Internet Project (2014)
3. Solove, D.J.: The Digital Person: Technology and Privacy in the Information Age. NYU Press, New York (2004)
4. Nissenbaum, H.F.: A contextual approach to privacy online. Daedalus **140**(4), 32–48 (2011)
5. Ramirez, E., Brill, J., Ohlhausen, M., Wright., McSweeny, T.: Data Brokers: A Call for Transparency and Accountability. Federal Trade Commission (2014)
6. Ajzen, I., Fishbein, M.: The influence of attitudes on behavior. In: Albarracin, D., Johnson, B.T. (eds.) The Handbook of Attitudes **173**(221), 31 (2005)
7. Khalil, A., Connelly, K.: Context-aware telephony: privacy preferences and sharing patterns. In: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, pp. 469–478. ACM (2006)
8. Benisch, M., Kelley, P.G., Sadeh, N., Cranor, L.F.: Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. Pers. Ubiquit. Comput. **15**(7), 679–694 (2011)
9. Olson, J.S., Grudin, J., Horvitz, E.: A study of preferences for sharing and privacy. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1985–1988. ACM (2005)
10. Patil, S., Lai, J.: Who gets to know what when: configuring privacy permissions in an awareness application. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 101–110. ACM (2005)
11. Lederer, S., Mankoff, J., Dey, A.K.: Who wants to know what when? privacy preference determinants in ubiquitous computing. In: CHI 2003 Extended Abstracts on Human Factors in Computing Systems, pp. 724–725. ACM (2003)
12. Huang, H.Y., Bashir, M.: Privacy by region: Evaluation online users' privacy perceptions by geographical region. In: Future Technologies Conference, pp. 968–977. IEEE (2016)
13. Maguire, S., Friedberg, J., Nguyen, M.H.C., Haynes, P.: A metadata-based architecture for user-centered data accountability. Electron. Markets **25**(2), 155–160 (2015)
14. Amazon Mechanical Turk. https://www.mturk.com. Accessed 18 Sep 2017
15. Paolacci, G., Chandler, J., Ipeirotis, P.G.: Running experiments on amazon mechanical turk. Judgm. Decis. Making **5**(5), 411–419 (2010)
16. Census Regions and Divisions of the United States. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf. Accessed 18 Sep 2017