Users’ Information Disclosure Determinants in Social Networking Sites: A Systematic Literature Review

Wajdan Al Malwi, Karen Renaud, Lewis Mackenzie

Abstract—The privacy paradox describes a phenomenon whereby there is no connection between stated privacy concerns and privacy behaviours. We need to understand the underlying reasons for this paradox if we are to help users to preserve their privacy more effectively. In particular, the Social Networking System (SNS) domain offers a rich area of investigation due to the risks of unwise information disclosure decisions. Our study thus aims to untangle the complicated nature and underlying mechanisms of online privacy-related decisions in SNSs. In this paper, we report on the findings of a Systematic Literature Review (SLR) that revealed a number of factors that are likely to influence online privacy decisions. Our deductive analysis approach was informed by Communicative Privacy Management (CPM) theory. We uncovered a lack of clarity around privacy attitudes and their link to behaviours, which makes it challenging to design privacy-protecting SNS platforms and to craft legislation to ensure that users’ privacy is preserved.

Keywords—Privacy paradox, self-disclosure, privacy attitude, privacy behaviour, social networking sites.

I. INTRODUCTION

SNSs encourage their users to share quantities of personal data. Moreover, the diffusion of the smartphones throughout society, as well as the ubiquity of a variety of Internet-enabled services, has exacerbated this sharing. Besides the undoubted benefits provided by these platforms, major privacy and safety concerns have emerged worldwide, especially in the Social Networking domain.

The PEW Research Center reports on a growing number of SNS users (from 5% in 2005 to 72% in 2019). Millions of people around the world use SNSs as part of their daily routine. The reason behind this international popularity is the incredible convenience offered by these networks in allowing users to stay in touch, establish relationships and create social capital [1]. The explosive growth of SNSs has driven many scholars to pose questions regarding disclosure of private personal information as a kind of quid-pro-quo [2]-[7].

We carried out a SLR to synthesize research in this area. Building on the approach proposed by [8], our study delivers a set of key factors that influence users’ information disclosure decisions when using SNSs. Our findings contribute to understanding the complicated nature of online privacy-related decisions.

II. ONLINE PRIVACY IN SNSs

SNSs are a distinctive online social environment, where personal information provided voluntarily by users and then collected, processed and analysed [9]. Such information can be easily imported, copied, saved and distributed with others. A range of privacy violation practices have emerged as online information collection has ramped up over recent years.

The disclosure of personal information on SNS platforms has received considerable critical attention from researchers. While SNS platforms deliver undeniable benefits, the aligned privacy concerns have been keenly debated. A phenomenon identified as the privacy paradox has emerged: labelling the seemingly paradoxical difference between users’ privacy-related behaviours and their self-reported high levels of privacy concerns.

The term privacy paradox was initially coined by Barnes [3]. It is called paradoxical because, in this case, those who claim that they value their privacy do not deploy the expected privacy protection strategies that such concerns would seem to indicate [10], [11]. The research suggests that SNS users do indeed engage in information privacy management and are not “passive users” as formerly referenced [12]. This means they are making deliberate trade-offs, and also that their stated privacy concerns are not as influential as anticipated in this context.

In prior studies addressing the privacy paradox phenomenon, sharing personal information on SNSs is seen as information self-disclosure occurring during social platform enabled connections between individuals [70], [13]. Personal information, feelings and thoughts are exchanged in most communication, reinforcing social ties [70]. This also occurs in the context of SNSs when users share private personal information with can be consumed by multiple users [14].

Individuals share their personal information for many reasons. They might want to construct social capital, enhance self-presentation, engage in impression management, and for entertainment purposes [15]. Basically, the information disclosure practices are aligned to the specific benefits offered by SNSs [16]. Because benefits will differ from person to person, such self-disclosure practices differ in scope, degree of intimacy, precision, purpose and awareness [70], [71]. CPM theory suggests that self-disclosure is typically a product of the cost/benefit evaluation process [13].

The privacy paradox phenomenon has enjoyed a great deal of attention in the research literature, with conflicting findings. A number of studies confirm the existence of the paradox e.g., [2], [3], [17]. Other studies challenge its existence e.g., [6], [7], [12] Such conflicting findings suggest that the paradox,
how it manifests and the factors that trigger it, are not yet well understood.

III. METHODOLOGY

A SLR is a high-level survey of a range of published studies on a specific topic. The SLR systematically identifies, assesses and investigates all related research evidence in order to answer a pre-defined research question or questions [8]. The main characteristic of a SLR is that it starts by developing a review protocol that defines the research grounds to systematically perform the review. There should be a clear set of objectives with pre-defined inclusion criteria for the studies. These characteristics make the SLR different than the regular literature review. SLRs aim to summarise and synthesise the research and investigations of objectives with pre-defined inclusion criteria for the studies. There should be a clear set of objectives with pre-defined inclusion criteria for the studies. In our SLR, we utilised the approach proposed by [8] to reach the intended goal. Kitchenham’s methodology [8] involves seven main stages. Fig. 1 depicts the review process.

A. Research Questions Development

A list of research questions was developed based on our original survey of SNSs privacy paradox literature, which are:

- RQ1: What key factors have been reported in privacy studies to determine users’ information disclosure decisions on SNSs?
- RQ2: What platforms have been selected for investigations?
- RQ3: What are the participants’ geographic distribution in the relevant studies?
- RQ4: What are the research methods adopted by researchers in the relevant studies?

B. Keyword Identification

Keywords used for the current SLR were specifically selected based on their frequency in the related literature. The selected keywords are a combination of terms including: "self-disclosure", "privacy concerns", "privacy paradox", "privacy behaviour", "privacy attitude" and "social network sites". Table I shows the search queries details.

C. Databases Selection

The search process was carried out on electronic databases, including Web of Science, SCOPUS, ACM, IEEE, Science Direct and Google Scholar.

D. Inclusion and Exclusion Criteria

To be included, the selected studies must focus primarily on information disclosure behaviour in SNSs. Moreover, our search was restricted to papers published in English between 2003 and 2020. The reason we limit our search with this time range, is that SNSs started gaining popularity worldwide in 2003 [1].

E. Search Method

A saturation sampling approach has been implemented to ensure a comprehensive exploration. As for every search query (SQ), we extracted and recorded related papers manually. Then, when new factors were detected, they were recorded, and combined with the growing list of factors. For each SQ, we had begun with a preliminary sample including 50 search results (SR), since this sample size was a controllable starting point. Next, all the retrieved papers were reviewed, evaluated, explored, and categorised based on the inclusion criteria, yielding a set of related papers (RP). Then, the other studies were examined in rounds, where each round contains five studies (with a 10% increase of the initial round). The exploration process terminated once we have two rounds of non-relevant papers. Table II demonstrates the applied saturation sampling approach.

F. Data Extraction and Recording

For each relevant paper, the following information was documented including author(s), year of publication, journal and theoretical background. Each study was classified based on the used methods, population of the research participants and the selected social networking platforms. In addition, for
them into seven key factors, as outlined in Fig. 2 and Table II. Having identified the reported factors, we further categorized impact users’ information disclosure directly or indirectly. These factors may in users’ privacy decision making on SNSs. These factors may be classified as an essential privacy decision criterion defined by CPM.

IV. FINDINGS

This section outlines the results of the relevant papers' analysis and presents answers to the pre-defined research questions.

RQ1) What Key Factors Have Been Reported in Prior Privacy Studies to Determine Users’ Information Disclosure Decisions on SNSs?

Our review identified 17 factors that play an essential role in users’ privacy decision making on SNSs. These factors may impact users' information disclosure directly or indirectly. Having identified the reported factors, we further categorized them into seven key factors, as outlined in Fig. 2 and Table III in the appendix section.

Fig. 2 Classification of the Retrieved Factors

A. Information Disclosure Costs

Information disclosure costs are inhibitors. This includes perceived privacy-related issues which negatively impact individual’s decision to disclose personal information on SNSs. Since SNSs are designed as public platforms, sharing personal information on such platforms should have remarkable privacy costs. Our analysis reveals that there are three main disclosure costs: (1) privacy concerns e.g., [18]-[20], (2) perceived privacy risks e.g., [21]-[23] and (3) perceived damage [11], [24], [25].

A.1 Privacy Concerns: this indicates the extent to which SNS users value their privacy. This factor refers to an individual’s concerns about the privacy issues associated with information disclosure in most studies, this construct represents the cost element in the privacy calculus process. It has been frequently utilised as a predictor of self-disclosure. Specifically, once users decide to disclose private information, they are concerned about information misuse by a) SNS providers and their partners (institutional privacy concerns), and b) other platform users (social privacy concerns). Therefore, privacy can be categorized into two dimensions: (1) social and (2) institutional privacy concerns.

A.2 Perceived Privacy Risks: this relates to users’ perceptions of the privacy threats related to information disclosure. It is another factor that may discourage users from sharing information on SNSs. Several studies found an evidence that users with high privacy risk perceptions reduce their information disclosure [5], [26], [27].

A.3 Perceived Damage: this factor describes users’ assessment of the level of damage that may occur in the event of privacy violations. Very few studies found evidence that perceived damage could lead to a user refraining from sharing their information [11], [24], [25].

B. Privacy Cost Mitigators

Our analysis uncovered several factors that might play a key role in the online privacy decision-making process, in terms of privacy cost mitigation. Three main factors emerge as privacy cost mitigators, which are: (1) privacy management e.g., [10], [28], [29], (2) trust e.g., [26], [30], [31] and (3) perceived control e.g., [11], [25].

B.1 Privacy Management: this term refers to the actions and strategies applied by SNS users to regulate and manage their online privacy. It has been empirically proven that SNS users employ different tactics to protect their privacy and mitigate their privacy concern, while keeping disclosing information to balance between the associated costs and benefits of SNSs. There are several privacy protection strategies reported in the related studies, such as: (1) updating the default privacy settings, (2) audience segmentation, (3) posts deletion, (4) unwanted user blocking, (5) messages encoding, (6) information falsification and (7) using multiple profiles with nicknames.

B.2 Trust: the role of trust has become one of the key instruments in the related privacy studies. Trust acts as an essential determinant of privacy-related decisions. Since information disclosure behavior embraces a certain degree of risk, trust acts as a mitigator and reduces risk perceptions. Two dimensions of trust have been identified, which are: (1) trust in institutions (institutional trust), and (2) trust in individuals (social trust). Accordingly, users’ privacy risk perceptions will be mitigated by trust in the SNS providers.
(institutional trust) and trust in other platform members (social trust).

B.3 Perceived Control: this factor indicates the extent to which SNS users feel that their released information is under their control. Perception of control over information is considered as privacy cost mitigator, as the high control perception will lead to more relaxed privacy concerns. On most of SNSs, certain measure of information control is offered to the public. However, users might still have concerns about the final control.

C. Online Privacy Literacy

Online privacy literacy is among the most important factors that impact SNS users’ privacy behaviours. Several studies claim that paradoxical behaviour may be justified by users’ lack of privacy literacy. Our analysis classifies online privacy literacy into two main constructs: (1) privacy awareness and the impact of privacy policy, e.g., [10], [32] and (2) privacy invasion experience e.g., [29], [33].

C.1 Privacy Awareness and Privacy Policy: privacy awareness is a key determinant that reflects the level of users’ understanding and knowledge about privacy practices and violations in SNSs. It is a core factor that is tightly linked to information disclosure and privacy related decisions in SNSs, as it defines the grounds of users’ perception of privacy regulation in SNSs. In addition, the presence of the platform privacy policy gives the users the impression that their information is confidential. However, [10] measures users’ consumption of privacy policy and have suggested that thorough reading of the privacy policy provided by Facebook negatively impacted users’ disclosure.

C.2 Privacy Invasion Experience: this factor highlights that users who have experienced privacy violation incidents, have an advanced level of online privacy literacy. Specifically, they will be more knowledgeable on how to assess privacy risks and adopt more privacy regulation strategies.

D. Information Self-Disclosure Motivators

Among others, the perceived benefits of using SNSs appear to be a significant facilitating factor. Building on the CPM theory, SNS users assess both benefits and costs when disclosing personal information. It has been found that information disclosure occurs when perceived benefits outweigh possible costs. Interestingly, SNS users’ information disclosure depends on each individual’s specific goals. In fact, findings indicate that social motivations increase the use of SNSs and correspondingly increase sharing more sensitive data. Likewise, self-presentation plays a central role in SNSs participation as users can post photos, achievements, and show their list of friends. Overall, there are some perceived benefits that motivate information disclosure in SNSs, such as:

- Relationship initiating;
- Relation maintenance;
- Social rewards: social capital, social validation, social control;
- Self-presentation;
- Enjoyment and entertainment;
- Convivence.

E. Social Impact Factors

This term has been used to describe the socially related factors, which refer to how other individuals (whether community or close connections) might influence users’ privacy behaviours in SNSs. We have identified (1) privacy social norms and (2) close connection influence (peers, family) as two dimensions.

D.1 Privacy Social Norms: this factor incorporates how community members might influence user’s perceptions about information privacy value. Users vary in the level of privacy valuation and their privacy social norms play a part in that aspect. For instance, users draw privacy lines depending on their perception of what others are expecting.

D.2 Close Connections Influence: this construct refers to the impact of close peers or family members privacy behaviour on the privacy decision making process. In other words, individuals tend to coordinate their privacy behaviours to be compatible with their close connections (peers, family members). Interestingly, it has been empirically proven that SNS users are more likely to adjust their profiles status into private, if their peers have already done so.

F. User’s Network Characteristics

A few studies have attempted to integrate the concept of privacy decision-making in SNSs with individuals’ network characteristics [31], [33], [43], [53]. That is, network characteristics have found to be impacting the level of information disclosure in SNSs. We identified network size and diversity and the perceived network’s responsiveness as sub factors.

F.1 Network Size and Diversity: the increase of the network size is correlated with a higher amount of information disclosure. Likewise, there is a greater chance that individuals reveal personal information on Facebook when they have a larger list of followers in their networks [4]. However, there is an interaction between users’ network diversity and information disclosure intimacy. Thus, users adjust their privacy regulation depending on their network categories (friends, family members, work colleagues).

F.2 Perceived Network’s Responsiveness: a link has been found between online communication frequency and information disclosure in the context of SNSs. These platforms provide individuals with an interactive venue for receiving peer comments and feedback. Specifically, users tend to engage in an intensive information disclosure in order to receive more social rewards. Therefore, the process of information disclosure is strengthened by the number of responses and impressions received from other network members.

G. Personal Characteristics

A few personal psychological characteristics have been detected in the related studies as antecedents of information disclosure. Narcissism, extraversion, risk taking, and self-esteem were linked positively to information disclosure and shown to boost information disclosure intentions in SNSs.
RQ2: What Platforms Have Been Selected for Investigations?

The findings obtained from our SLR showed that Facebook was mainly the most investigated platform in the relevant studies with (43 RPs) followed by SNSs in general (14 papers) as demonstrated in Fig. 3, which also shows other platforms examined by the RPs. Remarkably, some of the relevant studies considered two platforms in their exploration, for instance [34].

![Fig. 3 Platforms Used by Relevant Papers](image)

RQ3: What Are the Participants’ Geographic Distribution in the Relevant Studies?

Our analysis indicates that the majority of the relevant papers have studied USA users’ privacy behaviour intensively. Privacy paradox research in our SLR was carried out by researchers in 15 countries with USA dominating (28 RPs) followed by Europe, most were from Germany (11 RPs), then China (6 RPs). We also observed that some of the examined studies take the perspective of cross-country, such as [35] and [36]. These findings suggest that a deeper understanding of different international samples is needed in privacy paradox research, as well as a richer insight into privacy concerns from other populations other than the United States. Fig. 4 depicts the participants’ geographic distribution in the included papers.

![Fig. 4 Participants’ Geographic Distribution](image)

RQ4: What Are the Research Methods Adopted by Researchers in the Relevant Studies?

A wide range of methods were used to investigate the privacy paradox phenomenon. Most of the relevant studies adopted questionnaires as a primary method. Additionally, mixed methods were adopted to collect data, basically questionnaires along with another method. However, a few studies used interviews to understand users’ privacy perceptions. In summary, the majority of the studies (83.87%) adopted quantitative research approaches, with a few using a qualitative research approach (5%), and mixed methods (9.6%). Fig. 5 illustrates the used research methods in the RPs.

![Fig. 5 The Used Research Methods in the Relevant Papers](image)

V. DISCUSSION

The present SLR aimed to summarise the online privacy decision-making process adopted by SNSs users through a wide set of factors. Our analysis identifies seven main determinants of information disclosure behaviour in SNSs, which are: (1) information disclosure costs, (2) motivators, (3) privacy costs mitigators, (4) online privacy literacy, (5) user’s network characteristics, (6) social impact factors and (7) personal characteristics. It is important to highlight that prior studies might have overlooked particular privacy costs and discarded other important determinants in the context of SNSs. More specifically, the relationship between privacy concerns and information disclosure in SNSs has been thoroughly investigated. However, majority of studies have mainly emphasised privacy concerns stemming from mistrust in institutions and neglected the privacy concerns triggered by concerns about other users (social privacy concerns). Social privacy concerns have not received as much attention, with only a few studies attempting to take the first step [6], [7], [37]. Krasnova et al. [37] found that SNSs users were more concerned about social threats and their information disclosure was accordingly impacted. However, institutional privacy concerns did not affect their disclosure practices. This might help to explain the observed behavioural discrepancies discovered on some studies, which has been attributed to the privacy paradox.

The results suggest that SNSs users employ different strategies to protect their privacy and mitigate their privacy concerns.
concern, while keep disclosing enough information to balance between the associated costs and benefits of SNSs [4]. There is evidence that the available privacy settings did not address SNS privacy concerns. Accordingly, users have established further strategies to regulate their privacy such as: falsification of information, messages encoding, using multiple profiles. This has important implications for SNS providers and policy makers. We argue that understanding how privacy management impacts both privacy concerns and self-disclosure practices can provide new insights to privacy makers and SNS providers to help them to design systems that effectively reflect the privacy needs of SNS users.

Our analysis goes beyond previous reports, showing that most research on the privacy paradox phenomenon was conducted in USA specifically. There have been a few empirical investigations that conducted a cross-country perspective, and these often examine the variances between the US and another country. Since the original SNSs rise was in the United States, perhaps it was not unexpected that most prior research have investigated users’ behaviours in America followed by Europe. However, given that SNSs have become popular worldwide, there is a crucial need for more information privacy studies to shed the light on multiple countries while non-student samples are preferable. Thus, a deeper understanding of different international samples is needed as well as a richer insight into privacy concerns from other populations other than the United States.

The findings demonstrate that Facebook has been identified as a dominant platform, and this was confirmed. There might be a need to examine users’ online privacy practices on a broader set of platforms to understand whether users’ behaviour might change in response to the platform’s affordances and standards.

VI. CONCLUSION AND FUTURE WORK

The SLR provides a comprehensive overview of research into the privacy paradox phenomenon in SNSs. We cast a light on information disclosure determinants that have been observed in the online privacy behaviour literature. Altogether, we were able to summarise the key factors that predict information disclosure in SNSs in order to highlight research gaps. Such findings contribute to understanding the complicated nature of online privacy-related decisions in SNSs. It supports the design of privacy-protecting systems and tools, as well as formulation of privacy policies that accommodate users’ privacy concerns. Furthermore, such a study will deliver important insights to SNSs providers by identifying key determinants of users’ information disclosure behaviour. Such providers could then create designated instruments to resolve the defined privacy issues and thus ensure platform sustainability. The current SLR could be further employed as a framework for new studies. In our future work, based on the frequency of the retrieved factors, a set of core determinants has been detected. A comprehensive research model has been created incorporating all the identified factors. The model will be validated via a triangulation study, implementing an explanatory sequential design procedure where each phase builds on the results of the previous phase.

APPENDIX

| TABLE III |
| DETAILS OF THE RETRIEVED FACTORS |

<table>
<thead>
<tr>
<th>Factors</th>
<th>Citation</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Self-Disclosure Costs</td>
<td>Privacy Concerns: [4]-[7], [10], [11], [14]-[20], [24], [25], [29], [34], [35], [38], [53]</td>
<td>30</td>
</tr>
<tr>
<td>Perceived Privacy Risks</td>
<td>[5], [7], [15], [21], [23], [26], [27], [36], [41], [44], [50], [54], [55]</td>
<td>14</td>
</tr>
<tr>
<td>Privacy Costs Mitigators</td>
<td>Perceived Damage</td>
<td>[11], [24], [25]</td>
</tr>
<tr>
<td>Privacy Management</td>
<td>[4], [5], [10], [17], [22], [27], [30], [34], [35], [42], [45], [47], [48], [53], [56], [61], [64], [69]</td>
<td>22</td>
</tr>
<tr>
<td>Trust:</td>
<td>Trust in Providers</td>
<td>[11], [18], [21], [24], [26], [28], [30], [31], [39], [40], [43], [48], [49], [51], [54], [55], [58], [59], [62]</td>
</tr>
<tr>
<td>Trust in SNS Members</td>
<td>[49], [51], [54], [55], [58], [59], [62]</td>
<td>11</td>
</tr>
<tr>
<td>Perceived Control</td>
<td>[7], [11], [15], [24], [25], [39], [49], [50], [54], [55], [62]</td>
<td>11</td>
</tr>
<tr>
<td>Online Privacy Literacy</td>
<td>Privacy Awareness</td>
<td>[18], [25], [27], [41], [47], [54], [63]</td>
</tr>
<tr>
<td>Privacy Policy</td>
<td>[10], [18], [23], [50]</td>
<td>4</td>
</tr>
<tr>
<td>Privacy Invasion Experience</td>
<td>[23], [29], [33], [41], [65]</td>
<td>5</td>
</tr>
<tr>
<td>Privacy self-efficiency</td>
<td>Privacy Self-efficiency</td>
<td>[19], [35]</td>
</tr>
<tr>
<td>Information Self-Disclosure Motivators</td>
<td>Information Self-Disclosure Disclosure</td>
<td>[5], [7], [11], [19], [20], [22], [26], [30], [31], [35], [36], [46], [52], [57], [59], [60], [65], [68]</td>
</tr>
<tr>
<td>Perceived Benefits</td>
<td>[5], [7], [11], [19], [20], [22], [26], [30], [31], [35], [36], [46], [52], [57], [59], [60], [65], [68]</td>
<td>2</td>
</tr>
<tr>
<td>Social Impact Factors</td>
<td>Privacy Social Norms</td>
<td>[18], [53]</td>
</tr>
<tr>
<td>Social Influence</td>
<td>[29], [59], [62]</td>
<td>3</td>
</tr>
<tr>
<td>User’s Network Characteristics</td>
<td>Network Size and Diversity</td>
<td>[4], [23], [34], [58]</td>
</tr>
<tr>
<td>Perceived Network’s Responsiveness</td>
<td>[47]</td>
<td>1</td>
</tr>
<tr>
<td>Personal Characteristics</td>
<td>Self-Esteem</td>
<td>[31]</td>
</tr>
<tr>
<td>Risk Taking</td>
<td>[43]</td>
<td>1</td>
</tr>
<tr>
<td>Extraversion</td>
<td>[33]</td>
<td>1</td>
</tr>
<tr>
<td>Narcissism</td>
<td>[53]</td>
<td>1</td>
</tr>
</tbody>
</table>

REFERENCES


