SCIS 2021 2021 Symposium on
Cryptography and Information Security
On-line, Jan. 19 – 22, 2021
The Institute of Electronics,
Information and Communication Engineers

Investigating User Intention to Use a Privacy Sensitive Information Detection Tool

Vanessa Bracamonte* Sebastian Pape† Shinsaku Kiyomoto*

Abstract: Privacy sensitive information (PSI) detection tools have the potential to help users protect their privacy when posting information online, i. e. they can identify when a social media post contains information that users could later regret sharing. However, although users consider this type of tools useful, previous research indicates that the intention of using them is not very high. In this paper, we conduct a user survey (n=147) to investigate the factors that influence the intention to use a PSI detection tool. The results of a logistic regression analysis indicate a positive association of intention to use a PSI detection tool with performance expectation, social influence, and perception of accuracy of the tool. In addition, intention is negatively associated with privacy concerns related to the tool itself and with the participants' self-perceived ability to protect their own privacy. On the other hand, we did not find significant association with the participants' demographic characteristics or social media posting experience. We discuss these findings in the context of the design and development of PSI detection tools.

Keywords: Privacy sensitive information detection, privacy tools, user intention, user survey.

1 Introduction

With the expansion of online services such as social media, the amount of data that users share with others has increased. The consequence of this massive sharing of information is that in some cases users post privacy sensitive information, accidentally or without understanding possible consequences. This privacy sensitive information (PSI) can then be misused by others [1] and users have indicated that they regretted posting such information [2, 3].

There have been recent efforts to implement privacy enhancing technologies related to this problem: Privacy Detective [4] and PrivacyBot [5] are two examples of privacy enhancing tools that can detect privacy sensitive information in unstructured texts. Although these tools have potential to help users to make informed decisions and protect their privacy, little research has been conducted on how users perceive them. In an initial research in this area, it was found that participants considered PSI detection tools useful and interesting, but their level of intention of using the tool was not high [6].

For these type of tools to succeed in their objective, there is a need to understand which factors help or hinder their use. Therefore, in this paper, we conduct a user survey to quantitatively investigate the factors that positively or negatively influence user intention to use PSI detection tools.

* KDDI Research, Inc., 2-1-15 Ohara, Fujimino-shi, Saitama, Japan. va-bracamonte@kddi-research.jp
† Goethe University Frankfurt, Theodor-W.-Adorno Platz 4, Frankfurt, Germany.

2 Related Work

We first report about studies on the adoption of general or other privacy enhancing technologies. In the second subsection, we specifically highlight related work on PSI detection tools.

2.1 Adoption of Privacy Enhancing Technologies

User studies on adoption of privacy enhancing technologies (PETs) [7] and transparency enhancing technologies [8] that support user decisions indicate that these type of tools have promise in guiding users to make informed decisions. Although there is a broad technical discussion how to implement and build PETs, the investigation of acceptance factors and the users' intention to adopt PETs is still scant [9]. Recent studies mostly utilised the Technology Acceptance Model [10] such as work on Tor [11, 12, 13, 14], attribute-based credentials [15, 16, 17] or virtual private networks [11, 12, 18]. However, all of the investigated PETs have in common that they have their main popularity among privacy experts – maybe with the exception of virtual private networks which are also used to circumvent geoblocking for streaming services [19].

2.2 PSI detection tools

PSI detection is a growing research area [20]. Personally identifiable information prediction models [3, 21] and automatic recognition processes [21] have been developed for email data; mechanisms to detect privacy sensitive health information have also been proposed [22]. In the case of social media, approaches for PSI detection in tweets have been proposed [23, 24, 5],

to identify details such as vacation plans [23], for example, or whether the tweet contains information classified based on the EU GDPR [5].

Research on how to improve PSI detection is expanding, but on the other hand there is limited research on how users perceive this privacy technology. A user survey was previously conducted to initially explore how users perceived a PSI detection tool, what are their opinions about such a tool, and whether features such as explanation of the results of the tool influenced their opinion [6]. The study found that the PSI detection tool was considered useful and interesting, but that participants had a neutral intention to use the tool. In addition, one of the most frequent type of qualitative response was the opinion that the tool was useful, but for others rather than themselves.

The limitation of that study is that there was no quantitative validation of the impact of participants' negative and positive opinions on their intention to use the tool. In this paper, we address this limitation by conducting a logistic regression analysis with the factors identified in [6].

3 Method

In this section, we describe in detail which factors were considered in the analysis.

3.1 Model Design

A previous study classified both positive and negative opinions about a PSI detection tool [6]. The most frequent positive opinions were on the tool's usefulness, the perception of the tool as a "good idea" or interesting. On the other hand, the most frequent negative opinions were on the tool's privacy risk, the perception of the tool as not useful for the participants themselves, too high sensitivity in information detection, concerns about performance, and inconvenience. In this study we wanted to quantitatively investigate the influence of these variables on user intention.

In addition, we considered social media experience and demographic characteristics (age and gender) as independent variables. Finally, we considered *Social influence*, a factor of the Unified theory of acceptance and use of technology (UTAUT) [25] that was not identified in the participants' comments, but that is relevant for this type of privacy tool [9].

3.2 Questionnaire

We presented the participants a questionnaire to answer about a privacy alert tool for social media. We first asked participants to imagine that there is a tool that detects privacy sensitive information in online social media posts, that works by automatically analyzing text and showing an alert message if the text contains privacy sensitive information, so that the user can decide whether to continue posting, rewrite or not post.

We described the tool as free and provided by a reputable organization not associated with the social media site. We also indicated that the use of the tool is voluntary, not obligatory, and that the provider indicated that they would not sell the users' data or use it for purposes other than detecting privacy sensitive information disclosure. We included these descriptions as a way of establishing a level of trustworthiness of the tool based on a realistic scenario, since these concerns have been raised by participants [6] and trust in a PET has been shown as one of the main drivers in previous studies [8, 26, 27].

We then showed examples of alerts of the tool for different types of PSI (Figure 1) and asked the participants to answer the questionnaire based on their perception.

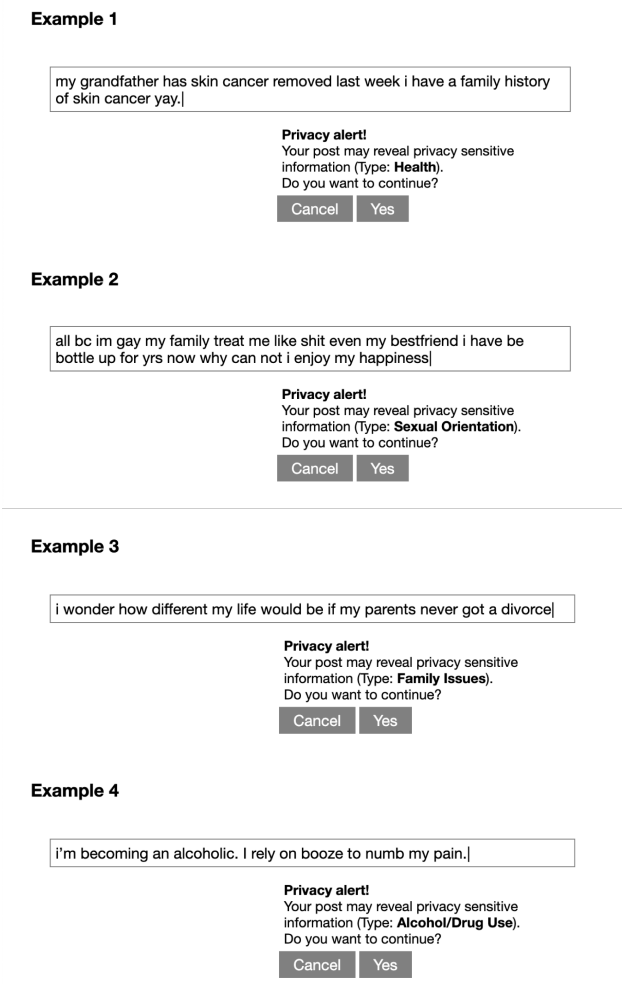

Figure 1: Alert examples of the PSI detection tool shown to the survey participants. Types: Health, Sexual Orientation, Family Issues, Alcohol/Drug Use.

Table 1 shows the details of the questionnaire's main items.

Besides the questionnaire items based on the qualitative findings of previous research [6], we also included questions about the participants' experience of regret of posting privacy sensitive information, and their past social media posting frequency, and frequency of intentionally posting personal information (Table 2).

In addition to the main items, we also asked par-

Table 1: Questionnaire items related to perception, on a 7-point scale from Strongly Disagree to Strongly Agree.

Construct	Question
Intention	I would use this tool to receive this type of privacy alert.
Performance expectancy	This tool would be helpful to decide whether or not to post certain information on social media.
Effort expectancy	Using this tool would be inconvenient.
Privacy concern	I am concerned about the privacy risks of using this tool.
Social influence	I would use this tool only if people I know used it.
Perceived accuracy	Based on the examples, I think the tool correctly detects privacy sensitive information.
Privacy self-efficacy	I can avoid accidentally posting privacy sensitive information on social media.

Table 2: Questionnaire items related to past social media experience.

Construct	Question
Posting regret experience	Have you ever had the experience of posting personal information on social media and then regretting it? (Yes/No)
Social media posting freq.	I post on social media. (Never - Very frequently)
Personal information posting freq.	I post personal information on social media intentionally (not by mistake) (Never - Very frequently)

ticipants about the social media sites they used most frequently and the type of provider they thought would be appropriate for this type of privacy tool.

3.3 Data Collection

The survey was conducted between November 27-28, 2020, using the Amazon Mechanical Turk platform (AMT). We recruited AMT workers with the following characteristics: USA, Canada, UK or Australia workers, a 99% task approval rate, and a minimum of 1000 worked tasks.

We initially collected 150 participant responses. We reviewed the answers to the attention check question and identified 3 invalid responses, which were eliminated from the data. This resulted in 147 valid responses.

4 Results

This section discusses the samples' characteristics, describes the results and presents a logistic regression.

4.1 Sample Characteristics

The participants' gender distribution was 26% females (38) and 54% male (79), with 20% blank responses (30). The participants' age distribution is shown in Figure 2.

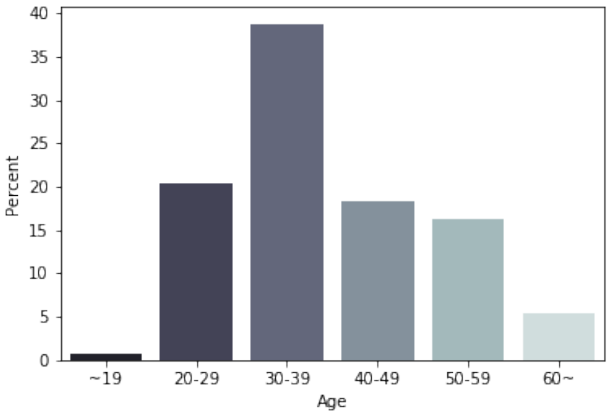

Figure 2: Percentage distribution of participants' age.

Our sample distribution contains fewer female users, but matches the age structure and social media site preference compared to recent reports [28].

The results also show that Facebook was the social media site most frequently used by participants, followed by Twitter (Figure 3). Other social media sites mentioned included YouTube and Reddit.

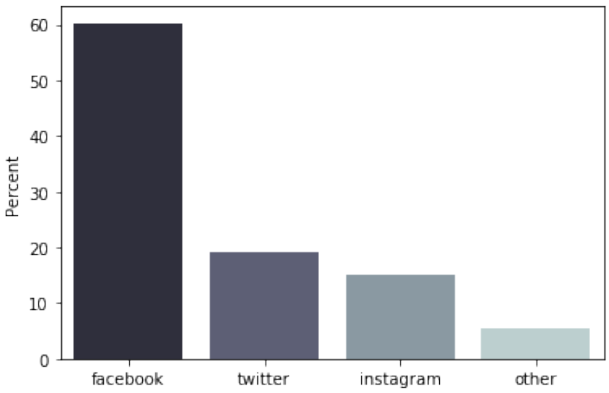

Figure 3: Answers to the questionnaire item "Please indicate the social media site you post to most frequently"

Finally, as Figure 3 shows, participants chose the social media site itself as an appropriate provider for this type of privacy tool more frequently than other types of providers. Other responses included non-profits.

4.2 Descriptive statistics

Figure 5 shows the percentage distribution of responses to main items. One can see that for the construct *Intention* the majority answered with *Yes*. For the perception constructs, the majority of the participants disagreed with the statements on effort expectancy,

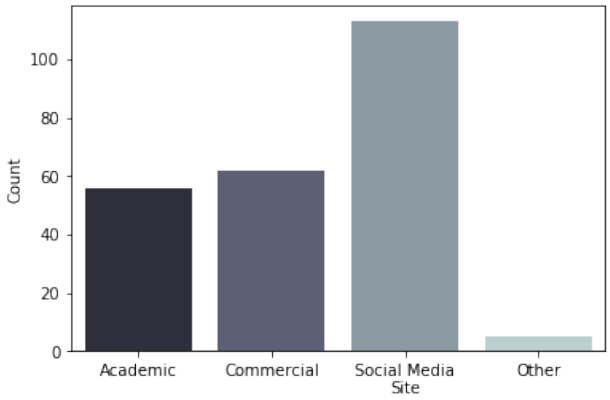

Figure 4: Answers to the questionnaire item "Which type of provider would be appropriate for this type of privacy tool?"

privacy concern and social influence, but agreed to performance expectancy, perceived accuracy and privacy self-efficacy. For the experience variables, the majority of participants answered *No* regarding their *Posting regret experience* and also responded that they did not frequently shared personal information on social media, although most participants indicated that they posted frequently in general.

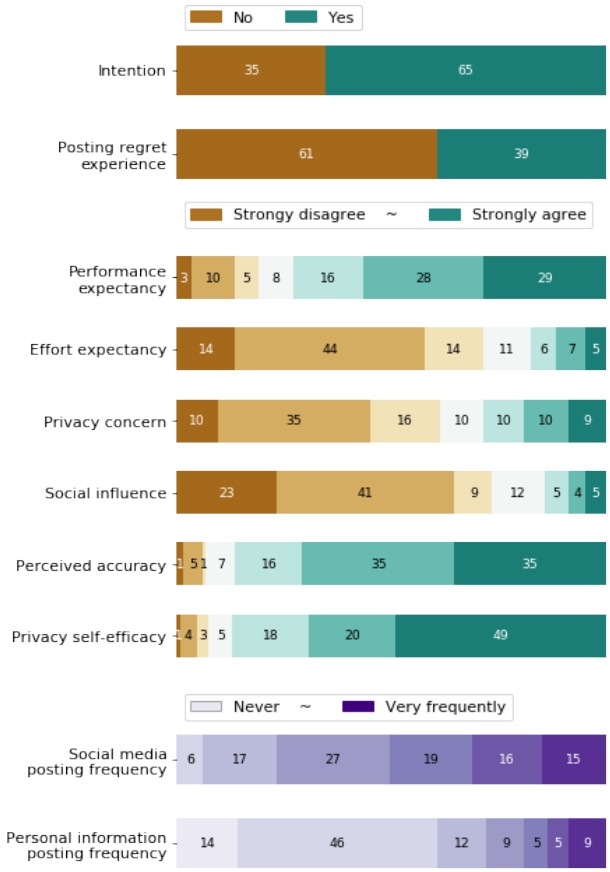

Figure 5: Percentage distribution of responses to main items.

4.3 Logistic Regression Analysis

We first converted the *Intention* variable to a Yes/No variable, removing 9 cases with a neutral response. In addition, since there were blank responses for age, we used the mode (male) to impute missing data for the purposes of this analysis. The independent variables were treated as numeric with the exception of *Posting regret experience (Yes/No)*, gender and age.

The result of the logistic regression analysis shows that five variables were significant: performance expectancy, perceived accuracy, social influence, privacy concern and privacy self-efficacy. Table 3 shows the detail of the results.

Table 3: Logistic regression results (significant coefficients).

Variable	Estimate (Std. Err)	p-value	Odds ratio (95% CI)
Performance expectancy	1.28 (0.26)	<0.001	3.60 (2.27-6.45)
Perceived accuracy	0.55 (0.27)	0.042	1.73 (1.03-3.00)
Social influence	0.55 (0.26)	0.033	1.73 (1.09-3.02)
Privacy concern	-0.44 (0.18)	0.015	0.65 (0.44-0.90)
Privacy self-efficacy	-0.54 (0.25)	0.034	0.58 (0.34-0.94)

Of these, performance expectancy, that is, the perceived usefulness of the tool, had the strongest positive association. Previous findings shows that most positive opinions about the PSI detection tool were related to its potential usefulness [6]. This result suggests that participants can clearly imagine that the tool would be helpful to make privacy decisions.

Social influence, whether the participant would use the tool if someone else they knew used it, and perceived accuracy, whether the participant thought the tool correctly detected privacy sensitive information, also had a positive association of a similar strength. Perceived accuracy, in particular, is a variable over which developer of the tool has influence. The results suggest that improving the reliability of these tools, and validating that users consider them accurate is important to promote their use.

In the case of the variables with negative associations, the result show that privacy concern and privacy self-efficacy were significant. These variables can also potentially be addressed directly: privacy concern, the worry that the use of tool itself involves a privacy risk, is related to the concern that the tool will store and misuse the users' text [6]. This concern may also be the reason why the social media site itself was chosen as the most appropriate provider for a PSI detection tool, since the the users are already posting their information on these platforms.

In this study, we did not measure whether indicating that the provider would not misuse the data reduced

the level of privacy concern, but one way of addressing this issue could be to provide technical and legal assurances that the users' privacy would be protected; future work could investigate this by manipulating the type of assurances given by the provider and evaluating its effect.

In the case of privacy self-efficacy, the tool could be designed to provide the user some evidence of their actual privacy efficacy. For example, the tool could measure the privacy risk score of the users' past social media posts, similar to the approach proposed by [29], or give the user examples of previous posts where they might have revealed privacy sensitive information.

Finally, from the non-significant results we see that previous social media experience and demographics were not significantly associated with intention. In addition, effort expectancy was also not significantly associated. In this study, the participants were only presented with mockups of the alert; it is possible that the effort or potential "annoyance" of actually receiving the alerts would have a stronger negative effect.

4.3.1 Limitations

Since the study was aiming to prepare the ground for conducting a larger scale study, we only used limited resources, i. e. we used single-items to measure the factors of interest. Additionally, the sample size was on the lower side for this type of analysis, where a rule-of-thumb is 10-20 participants per variable of interest. This means that only medium sized to large effects could be detected. Future work will address these limitations by conducting a larger scale study to validate the results of the current study.

5 Conclusions

In this paper, we investigated the factors associated with user intention to use a PSI detection tool. We conducted a user survey and obtained responses from 147 participants. The results of a logistic regression analysis indicate that intention to use the tool is positively associated with performance expectancy (usefulness), social influence and perception of accuracy. On the other hand, intention is negatively associated with privacy concerns about the tool, and with the perception that participants could avoid privacy risks by themselves (privacy self-efficacy). Future work is planned to evaluate the effect of these variables and investigate their strength on different conditions of the PSI detection tool.

References

[1] A. Acquisti and C. Fong, "An experiment in hiring discrimination via online social networks," *Management Science*, vol. 66, no. 3, pp. 1005–1024, 2020.

[2] M. Sleeper, J. Cranshaw, P. G. Kelley, B. Ur, A. Acquisti, L. F. Cranor, and N. Sadeh, ""i read my twitter the next morning and was astonished": A conversational perspective on twitter regrets," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, (New York, NY, USA), p. 3277–3286, Association for Computing Machinery, 2013.

[3] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, and L. F. Cranor, ""i regretted the minute i pressed share": A qualitative study of regrets on facebook," in *Proceedings of the Seventh Symposium on Usable Privacy and Security*, SOUPS '11, (New York, NY, USA), Association for Computing Machinery, 2011.

[4] A. Caliskan Islam, J. Walsh, and R. Greenstadt, "Privacy detective: Detecting private information and collective privacy behavior in a large social network," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, WPES '14, (New York, NY, USA), p. 35–46, Association for Computing Machinery, 2014.

[5] W. B. Tesfay, J. Serna, and K. Rannenberg, "Privacybot: Detecting privacy sensitive information in unstructured texts," in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 53–60, 2019.

[6] V. Bracamonte, W. B. Tesfay, and S. Kiyomoto, "Towards Exploring User Perception of a Privacy Sensitive Information Detection Tool," in *7th International Conference on Information Systems Security and Privacy*, 2021 (forthcoming).

[7] A. Acquisti, I. Adjerid, R. Balebako, L. Brandimarte, L. F. Cranor, S. Komanduri, P. G. Leon, N. Sadeh, F. Schaub, M. Sleeper, *et al.*, "Nudges for privacy and security: Understanding and assisting users' choices online," *ACM Computing Surveys (CSUR)*, vol. 50, no. 3, pp. 1–41, 2017.

[8] M. Janic, J. P. Wijbenga, and T. Veugen, "Transparency enhancing tools (tets): an overview," in *2013 Third Workshop on Socio-Technical Aspects in Security and Trust*, pp. 18–25, IEEE, 2013.

[9] F. Mangiò, D. Andreini, and G. Pedeliento, "Hands off my data: users' security concerns and intention to adopt privacy enhancing technologies," *Italian Journal of Marketing*, pp. 1–34, 2020.

[10] F. D. Davis, *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology, 1985.

[11] F. Brecht, B. Fabian, S. Kunz, and S. Mueller, "Are you willing to wait longer for internet privacy?," 2011.

[12] F. Brecht, B. Fabian, S. Kunz, and S. Müller, "Communication anonymizers: personality, internet privacy literacy and their influence on technology acceptance," 2012.

[13] D. Harborth and S. Pape, "Examining technology use factors of privacy-enhancing technologies: The role of perceived anonymity and trust," in *24th Americas Conference on Information Systems, AMCIS 2018, New Orleans, LA, USA, August 16-18, 2018*, Association for Information Systems, 2018.

[14] D. Harborth, S. Pape, and K. Rannenberg, "Explaining the technology use behavior of privacy-enhancing technologies: The case of tor and jondonym," *Proceedings on Privacy Enhancing Technologies (PoPETs)*, vol. 2020, pp. 111–128, 05 2020.

[15] Z. Benenson, A. Girard, I. Krontiris, V. Liagkou, K. Rannenberg, and Y. Stamatiou, "User acceptance of privacy-abcs: an exploratory study," in *International Conference on Human Aspects of Information Security, Privacy, and Trust*, pp. 375–386, Springer, 2014.

[16] Z. Benenson, A. Girard, and I. Krontiris, "User acceptance factors for anonymous credentials: An empirical investigation.," in *WEIS*, 2015.

[17] I. Krontiris, Z. Benenson, A. Girard, A. Sabouri, K. Rannenberg, and P. Schoo, "Privacy-abcs as a case for studying the adoption of pets by users and service providers," in *Annual Privacy Forum*, pp. 104–123, Springer, 2015.

[18] M. Namara, D. Wilkinson, K. Caine, and B. P. Knijnenburg, "Emotional and practical considerations towards the adoption and abandonment of vpns as a privacy-enhancing technology," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 1, pp. 83–102, 2020.

[19] S. Earle, "The battle against geo-blocking: The consumer strikes back," *Rich. J. Global L. & Bus.*, vol. 15, p. 1, 2016.

[20] W. B. Tesfay, J. Serna, and S. Pape, "Challenges in detecting privacy revealing information in unstructured text," in *PrivOn@ ISWC*, 2016.

[21] C. Bier and J. Prior, "Detection and labeling of personal identifiable information in e-mails," in *IFIP International Information Security Conference*, pp. 351–358, Springer, 2014.

[22] M. Sokolova, K. El Emam, S. Rose, S. Chowdhury, E. Neri, E. Jonker, and L. Peyton, "Personal health information leak prevention in heterogeneous texts," AdaptLRTtoND '09, (USA), p. 58–69, Association for Computational Linguistics, 2009.

[23] H. Mao, X. Shuai, and A. Kapadia, "Loose tweets: an analysis of privacy leaks on twitter," in *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*, pp. 1–12, 2011.

[24] S. R. M. Castillo and Z. Chen, "Using Transfer Learning to Identify Privacy Leaks in Tweets," in *2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC)*, pp. 506–513, Nov. 2016.

[25] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS quarterly*, pp. 425–478, 2003.

[26] D. Harborth and S. Pape, "Jondonym users' information privacy concerns," in *ICT Systems Security and Privacy Protection - 33rd IFIP TC 11 International Conference, SEC 2018, Held at the 24th IFIP World Computer Congress, WCC 2018, Poznan, Poland, September 18-20, 2018, Proceedings*, pp. 170–184, 2018.

[27] D. Harborth and S. Pape, "How privacy concerns and trust and risk beliefs influence users' intentions to use privacy-enhancing technologies – the case of tor," in *52nd Hawaii International Conference on System Sciences (HICSS) 2019*, pp. 4851–4860, 01 2019.

[28] Socialbakers, "Social media trends report Q4 2019." `https://www.socialbakers.com/web-api/wp/study/social-media-trends-report-key-insights-you-need-to-know?studyId=24325`, 2019.

[29] E. Aghasian, S. Garg, L. Gao, S. Yu, and J. Montgomery, "Scoring users' privacy disclosure across multiple online social networks," *IEEE Access*, vol. 5, pp. 13118–13130, 2017.