Privacy Concerns Amidst OBA and the Need for Alternative Models

Farah Chanchary, Yomna Abdelaziz, Sonia Chiasson
Carleton University
Ottawa, Canada

April 22, 2017

Abstract

What convinces users to share information about themselves in an ad-based online world? We explore users' willingness to share data in the context of online behavioural advertising (OBA) and tracking prevention tools. We find positive responses for OBA and clear preferences for which types of information users would like to disclose. Factors including privacy attitudes and control mechanisms impact people's decisions to share their information with websites. These factors yield a discussion of the ad-based business model of the Internet and the balance with user preferences.

Keywords: online privacy, user study, human-computer interaction, OBA

Advertising networks use online tracking to create user profiles based on online activities and preferences without consent from users. The main mechanism for online tracking is third-party HTTP cookies by advertising domains [1]. Since users directly interact with first-party websites and may be unaware of hidden third parties, the data collection process may violate their online privacy. Thus, users' understanding of online behavioural advertising is critical in helping them protect their privacy. Research shows that users' data sharing willingness is influenced by familiarity with the advertising companies [2], the type of first-party website [3], and the third parties collecting the data [4]. We re-investigate users' perceptions of OBA, contributing to the discussion by examining different kinds of first-party websites and incorporating the impact of privacy attitudes, privacy practices, and technical background on data sharing willingness. Furthermore, we explore users' understanding of privacy tools and ad-blocking tools. Using an online survey, we collected responses from a geographically diverse sample of 368 participants on their willingness to share 24 types of information.

Our participants showed a relatively consistent willingness to share personal information with websites. Privacy fundamentalists were least willing to share any type of information. Participants with technical background showed increased willingness to share information for OBA. Tracking-prevention tool (TPT) features also made participants more inclined to share data. Overall, participants preferred seeing relevant ads and would share their personal information with online advertisers to receive these ads if they could control what information to share and with whom.

1 Online Behavioural Advertising in Context

Research by Blase Ur et al. [2] finds that participants deem OBA to be simultaneously useful and invasive to their privacy. Participants also had strong concerns about data collection, depending on the context. Elisa Costante et al. [5] show that users' perception of trust varied with application domains and users' IT related knowledge. In 2014, Emilee Rader [6] finds that despite having profound knowledge of first-party data tracking, participants were much less aware of automatic collection, collaboration and data aggregation across various websites.

Lalit Agarwal et. al [7] report that online tracking made users concerned about the content of online ads and how their browsing behaviour could lead to embarrassment. Moreover, the authors mention third-party-indifference as a major finding since they did not observe a significant difference in participants' trust levels across first and third-party websites. Yang Wang et al. [3] find that the type of first-party website significantly affects people's willingness to share data.

Users' lack of knowledge of tracking prevention tools, as well as usability issues with such tools affect their intentions to adopt suitable privacy practices. A survey by Aleecia McDonald and Jon Peha [8] in 2011 suggests a large gap between the actual implementation of "Do Not Track" in web browsers and what users expect from it, e.g., stopping complete data collection across websites. Pedro Leon et al. [1] also demonstrate that users misunderstand how tracking prevention tools work and mistakenly believe they are protected against tracking. To improve privacy tools for users, William Melicher et al. [9] propose design guidelines and explore the use of a classifier to automate privacy settings for users. Jagdish Achara et al. [10] argue that the economic ad-based model of the Internet is in danger because of ad-blocking tools. As a solution, they provide a browser extension that gives users fine-grained control based on website categories as opposed to the complete blocking of all ads.

In investigating data sharing willingness, Leon et al. [4] explore two versions of a first-party health website. In contrast to Agarwal et al.'s results, participants' willingness to share information is influenced by the privacy practices of the third party that collected the data, rather than the first-party site. Because health information is particularly sensitive and unique, we investigate whether users truly had no concerns regarding first-party tracking by exploring non-health related websites of varying sensitivity.

2 The Study

We conducted a between-subjects online study, partially following the research methodology adopted by Leon et al. [4]. We recruited participants from around the world using an online crowdsourcing service, CrowdFlower. Participants were between the ages 18 and 73 (mean=31.7 and σ=9.5), 27% were female and 72% male. Their occupations varied, but about half (49%) had a computer or IT-related background through education or work experience. To avoid biased answers, our recruitment material gave no indication that privacy would be a component of this study. This study was approved by our institutional Research Ethics Board. We note that the data we collected are self-reported values based on participants' views towards OBA and perceived willingness to share personal information in hypothetical scenarios. We are unable to confirm how well this maps to users' actual behaviour. These limitations are common with several other related studies available in the literature.

Our respondents come from 56 countries. The top geographical location is India (10% of respondents), followed by Canada (8%), Venezuela (6%), Serbia (6%), United Kingdom (5%), Spain (5%), Portugal (5%), Vietnam (4%), Turkey (4%), and USA (4%). The remaining 46 countries are each represented by 4% or less of participants.

2.2 Structure of the Questionnaire

Our survey questionnaire covered six parts:

1. Demographic information
2. Basic understanding of online advertising
3. Informational video on OBA [11], along with two test questions to screen out participants who weren't paying attention
4. Willingness to share 24 types of information. For this part only, participants were evenly distributed into four groups and assigned to one type of website: online banking (OB), online shopping (OS), search engine (SE) or social network (SN) site. Participants disclosed their willingness to share information with their assigned first-party site
5. Understanding of tracking prevention tools; and
6. Privacy attitudes and practices

In Part 3, we found 32 participants with incorrect answers. Our analysis used responses from 386 participants who passed both test questions.

2.3 Analysis

We performed statistical tests to identify significant patterns among several data elements we collected. All statistical tests assumed a significance level of $p<0.05$.

Factor Analysis:

To investigate how the categories of websites influenced participants' willingness to share 24 types of information, we performed factor analyses to reduce these 24 types to a smaller number of output variables. Factor analysis is a process that evaluates underlying associations of closely related variables and combines them into a single latent factor. We conducted further analysis based on these latent factors instead of the individual variables.

Our exploratory factor analyses found that 17 data types could be grouped into 4 factors and the remaining 7 data types did not conform to any particular factor. As in Leon et al. [4], we considered a variable part of a factor if it had a factor loading of at least 0.6 for the particular group, as well as factor loadings under 0.4 for all other groups. We named the resultant factors: *(1) Demographic Information*, *(2) Personal Identification Information (PII) & Financial Information*, *(3) Location Information*, and *(4) Computer Information*. We used Cronbach's alpha (α) value for each factor to estimate the internal reliability of the factor analysis test. All four resultant factors had alpha values higher than 0.8, which is the standard to support high correlations between group members. We conducted all further analyses using the four resultant factors. We created an index variable for each factor by averaging participants' responses to all the questions included in the factor.

Westin Index Analysis:

The Westin Index [12] is a set of three questions designed to quickly segment users into three groups: (1) Privacy Fundamentalists, who view privacy as having a high value which they feel very strongly about; (2) Privacy Pragmatists, who have strong feelings about privacy but also see the benefits of surrendering some privacy in situations where they believe care is taken to prevent the misuse of this information; and (3) Privacy Unconcerned, who have no real concerns about privacy or about how other people and organizations use information about them. We found that 30.4% of our participants were Privacy Fundamentalists, 45.9% were Privacy Pragmatists and 23.6% were Privacy Unconcerned. This conforms to typically observed demographics [12]. We explored how participants' privacy attitudes influenced their data sharing willingness. Where appropriate, we further analysed these correlations according to categories of websites.

3 Results

3.1 Practices, Understanding and Perception

Half of participants were aware of OBA, but most were oblivious to the functionalities of tracking prevention tools. In general, participants were dissatisfied with receiving targeted ads based on their online activities. While half of participants appreciated the idea of user-customized targeted ads, half (not mutually exclusive) reported generally ignoring current targeted ads.

Practices:

Participants' privacy attitudes significantly affected their data sharing willingness for two out of four overall factors: personal identification, financial, and demographic data. Privacy fundamentalists were most protective of their information.

Participants were given a list of privacy-preserving practices and could select all that applied. Most users (>80%) took specific steps to preserve their online privacy. The top practice was deleting cookies, followed by refusing to give out unnecessary personal information. The third-ranking practice was terminating online transactions when they were uncertain about the data retention, followed by reading a website's privacy policies, and finally, activating the *Do Not Track* option in web browsers or installing tracking prevention tools on their computers (58%). We found that while users are taking steps to prevent online data leakage, they use only a subset of available safeguards.

Understanding:

Results of open-ended questions showed that 55% of participants could define website advertising. Only 6% of participants mentioned that website advertising was beneficial, while others thought it was spam, annoying, and false information. Even though almost half of participants had degrees or work experience in computer related fields, we found that overall awareness of how targeted ads and privacy protection tools work was relatively low. We asked them to explain how targeted ads worked and 46% had partially correct answers. Only 38% of participants could correctly explain how tracking prevention tools worked.

Website Ads and Online Tracking:

Half of participants agreed that website advertising is necessary to enjoy free services on the Internet, 42% found website advertising useful, and 42% believed that website advertising relevant to their interests can save time. However, half also said that they did not normally notice the ads that appeared on the websites they visited.

Approximately half of participants were aware of the various tracking capabilities. Nearly one-fifth believed it was impossible for online tracking systems to track all websites visited, and some wrongly believed that companies did not track individuals' online activities without users' permission (27%). More than half of participants correctly assumed that advertisers collect their personal information and track their location, visited websites, and online behaviour.

Targeted Ads:

Only 23% of participants liked receiving targeted ads reflecting their online activities, while 37% expressed clear dislike, and the remainder were neutral. In response to our open-ended question, *"Explain what, if anything, would make you feel more comfortable with receiving targeted ads?"*, participants displayed a variety of reactions, including criticisms for currently generated targeted ads. They saw much more value in seeing ads based on their actual expressed interests. This was clearly articulated by participants in our study: *"Most of the time I get ads that have nothing to do with me, being a girl doesn't mean I'm looking for makeup or trying to get skinny or whatever other stereotyped information that make ads show up"* or *"I'm tired of keep getting ads that I searched over 1 month ago."*

3.2 Impact of First and Third Parties

There was no significant difference regarding first parties on participants' willingness to share data, contrasting Wang et al.'s results [3]. Except for online banking (OB), participants were equally concerned between first and third-party tracking.

Although the type of first-party site did not impact participants' data sharing willingness, participants do distinguish between different types of information. Figure 1 shows participants' responses based on a 5-point Likert scale across all websites. Participants expressed relatively consistent preferences across the four website categories. Approximately half were willing to share items at the bottom of the figure and few wanted to disclose items near the top, although it is unclear why anyone would want to share details such as their credit card number. More participants were willing to share their location information (46%), demographic and computer information (39%) than their personal identification (PII) and financial information (13%). These results are consistent with that of Leon et al. [4] with health websites. However, our results confirmed that preferences also hold for other types of first-party websites.

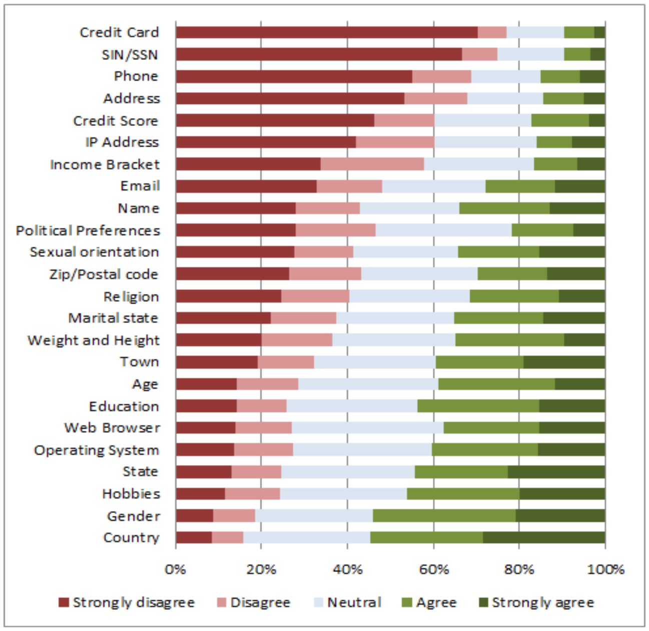

Figure 1: Willingness to disclose to a first-party website.

Concern for First- and Third-Party Tracking:

In general, participants from the OS, SE, and SN groups expressed approximately equal levels of concern for both first and third parties. Only participants of the online banking (OB) group expressed increased concern (55%) for third-party tracking than for first-party tracking of the OB site (37%). We suggest two possible reasons for this result. First, online banking sites generally do not show many online ads compared to other sites so any ads may be viewed suspiciously. Secondly, users manage highly sensitive financial data through OB sites, and reasonably wish to avoid third-party tracking of such data.

3.3 Impact of Tracking Prevention Tool Features

Participants clearly distinguished between tracking prevention tools and ad-blocking tools, and the majority (55%) considered tracking prevention tools more useful than ad-blocking tools (37%). Overall, 72% of participants chose tracking prevention tools as their preferred tool over ad-blocking tools. Nearly half of participants, across all websites and irrespective of their privacy attitudes, were more willing to share data if they could restrict both first and third parties from collecting data, select types of information to share, and customize topics of targeted ads.

3.4 Other Factors Affecting Willingness to Share

We conducted post-hoc exploration of other potential factors. Although the statistical analysis is left out of this short article, results showed statistically significant differences for the following factors.

Frequency of Website Visit:

Frequency of visiting a website significantly influenced overall willingness to share for location data and PII and financial data. Daily visitors were more willing (34%) to share location information than infrequent visitors (12%). We also found that frequent visitors of search engines (SE) were less likely to share PII and financial data than participants who visited SE websites only a few times in the last year.

Many websites provide location-based selection or search facilities for their client services. Frequent Internet users might perceive this as a useful feature and hence be more willing to share these data. However, financial or PII data are too sensitive and frequent users appear aware of the risk of online exposure.

Computer Related Background:

IT participants were significantly more willing to share their PII data and financial information and computer related information than the non-IT users (42% compared to 27%). Similarly, 13% of non-IT participants refused to share computer related data compared to 5% of IT participants. It appears that those in computer related fields are more confident in their abilities to handle the risk of information leaking and are more willing to share these data.

Intentions to Explore Online Ads:

More users who clicked links to explore online ads (25%) would like to receive targeted ads based on their online activities than users who did not explore ads (17%).

We further found significant impact of this intention on participants' concern for third-party tracking. Users who clicked links to explore online ads (52%) knew that they might be at risk and showed increased concern for third-party tracking compared to users who did not explore ads (38%). However, participants' intentions to explore ads did not influence their concern for first-party tracking.

Access to Collected Data:

We found that 25% of participants were more willing to share information if they were given access to collected data for reviewing, editing, or permanently deleting. Interestingly, most participants did not change their data sharing willingness. Many of them expressed distrust of such a mechanism. Some did not care for the correctness of their online profiles, and others simply did not want their information collected.

Furthermore, most our participants were unwilling to pay to stop targeted ads (61%) or online tracking (51%). This result supports our findings on the impact of tracking prevention tool features and also matches results published in Leon et al. [4].

3.5 Summary of Results

To summarize, we identified four factors that greatly influenced participants' willingness to share various types of PII and non-PII data (see Table 1): (1) participants' privacy attitudes, (2) frequency of visiting a specific type of website, (3) technical background, and (4) intention to explore online ads. Our participants also showed preferences for the types of data they were willing to share online. The choice of first-party websites had no impact on participants' data sharing willingness, confirming the findings by Leon et al. [4], but contrasting those of Wang et al. [3].

Some factors influenced a subset of our participants, such as options that allowed access to modify user profiles, and tracking prevention tool features to restrict data collection or to select topics for targeted ads.

Table 1: Factors affecting participants' sharing willingness.

Factors	Impact Level
Type of first-party website	None
Control features of tracking prevention tools	Moderate
Access to collected data	Moderate
Privacy attitude	High
Frequency of website visit	High
Computer/IT background	High
Exploring Online ads	High

4 Privacy Concerns: Solutions and Alternatives

Confirming and extending prior studies [2], [4], [7] we found that the type of first-party website had no major impact on participants' willingness to share data. A recent study contrasts with this finding [3], suggesting the need for future study on people's preferences in relation to website types. Furthermore, participants expressed equal concern for both first and third-party tracking. Privacy fundamentalists were unwilling to disclose personal, financial, and demographic data for any type of website. Other types of data were also of concern to smaller segments of the population; providing opportunity to voice a preference would be beneficial in these cases. In response, consent mechanisms should offer some assurance that opt out preferences are being observed.

Other recent work confirms our findings; a 2015 Pew Research survey [13] finds that many Americans are willing to disclose information depending on the value being offered and the risk of doing so. Participants displayed a wide range of sensitivity to disclosing information, but some information was inherently more private than others.

We found that a remarkable number of participants were open to targeted ads, if they had some control over what information is being collected. Several researchers have recently proposed alternatives to address these issues raised in our study. For example, Achara et al. [10] propose a nuanced solution to give users that control without eliminating online ads. Their solution allows users to block ads and tracking based on the category of content on a webpage-level because many websites combine content of varying sensitivity. This solution is put forth as a compromise between user privacy and the ad-based economy of the Internet.

In preventing invasion to privacy while allowing desired tracking, Melicher et al. [9] explore the use of a classifier to automatically detect users' privacy preferences based on the type of website being visited. The researchers propose design guidelines to improve existing tools: automate common preferences, give users control in specific situations, and inform users about how online activities impact the information that might be inferred about them. Melicher et al. trained a classifier to automatically identify the correct privacy preferences in 50% of cases.

A 2016 literature survey [14] finds that in addition to ad-blocking, other privacy tools also employ other protection strategies including obfuscating or sandboxing user data. For example, Adnostic [15] combines obfuscation and sandboxing of user data, allowing users to receive targeted ads without revealing their information to third-party trackers. Adnostic forms online profiles locally in the users' browser, based their search and browsing history. Furthermore, it enables ad networks to charge the advertisers without knowing which ads were displayed to users. Although this architecture mitigates some privacy risks, Estrada-Jiménez et al. [14] argue that by limiting its data sources, Adnostic produces content that is less relevant to users, thereby decreasing its attractiveness to advertisers.

For users who do not want ads altogether, alternative economic models are needed. Tension between the privacy concerns and preferences of users and the economic realities of the online world means that this issue is unlikely to be resolved soon. When users block ads or behavioural tracking, the revenue potential of visited websites is diminished. Many sites are fighting the technological advances that give users more control with counter-measures. For example, Facebook has recently decided to prevent third-party ad-blocking tools (ABT) [16]. Open-source solutions to counter Facebook's ABT-blocking feature are being continuously developed, creating a familiar ongoing struggle between both sides.

Alternative sources of revenue are possible, including through paid content. However, many users are currently unwilling to pay a fee for online content or services. Approximately only half of our survey participants report willingness to pay a fee, and other recent polls show similar or lower numbers of users currently accessing paid content [17]. However, some early signs suggest a shift in attitudes in the newer generation; 78% of millennials have paid for at least one type of entertainment content [18]. However, even in these cases, paid content is typically for "special requests," viewed as an infrequently paid premium or for special types of content, rather than something done for access to all online content. Moving to a pay-per-use or subscription model for most online content would lead to a fundamental shift in the nature of the web.

Another alternative economic model is counter-cyclical offering [19] in which content providers offer more free content at times of high demand, thus attracting more users and reaping the benefits of advertising revenue. In times of low demand, content providers can offer paid options which will cater to the segment of users who are generally willing to pay for content regardless of the demand wave. Such an economic model assumes that consumers are diverse in how they value content and in their willingness to pay for it.

5 Conclusions

The vast majority of users indicate an active intention to preserve their online privacy and our findings reveal that they are more concerned about tracking than online ads. While a significant portion of users do not oppose targeted ads, they have differing privacy attitudes with complex privacy needs that necessitate usable control mechanisms to meet them. Research on privacy attitudes and preferences informs the design of such tools. The complexity of the online economic model combined with varied user preferences and tolerance for data sharing suggest that a uniform approach is unlikely to gain traction. Ultimately, a hybrid model consisting of several of the above approaches will likely continue for the foreseeable future.

6 Acknowledgements

F. Chanchary is a recipient of an Ontario Graduate Scholarship. S. Chiasson holds a Canada Research Chair in Human Oriented Computer Security. She acknowledges the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding the Chair and her Discovery Grant.

References

[1] P. G. Leon, B. Ur, R. Balebako, L. F. Cranor, R. Shay, and Y. Wang, "Why Johnny Can't Opt Out : A Usability Evaluation of Tools to Limit Online Behavioral Advertising," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, pp. 589–598, 2012.

[2] B. Ur, P. G. Leon, L. F. Cranor, R. Shay, and Y. Wang, "Smart, Useful, Scary, Creepy: Perceptions of Online Behavioral Advertising," *Symp. Usable Priv. Secur.*, vol. 2012, p. 1, 2012.

[3] Y. Wang, H. Xia, and Y. Huang, "Examining American and Chinese Internet Users- Contextual Privacy Preferences of Behavioral Advertising," *Proc. 19th ACM Conf. Comput. Coop. Work Soc. Comput. - CSCW '16*, pp. 538–551, 2016.

[4] P. G. Leon *et al.*, "What Matters to Users? Factors That Affect Users' Willingness to Share Information With Online Advertisers," *Proc. Ninth Symp. Usable Priv. Secur.*, p. 1, 2013.

[5] E. Costante, J. Den Hartog, and M. Petkovic, "On-line Trust Perception: What Really Matters," *Proc. 1st Work. Socio-Technical Asp. Secur. Trust*, pp. 52–59, 2011.

[6] E. J. Rader, "'Awareness of Behavioral Tracking and Information Privacy Concern in Facebook and Google" *Symposium on Usable Privacy and Security (SOUPS)*, 2014, pp. 51–67.

[7] L. Agarwal, N. Shrivastava, S. Jaiswal, and S. Panjwani, "Do Not Embarrass: Re-examining User Concerns for Online Tracking and Advertising," *Proc. Ninth Symp. Usable Priv. Secur.*, p. 8:1--8:13, 2013.

[8] A. M. Mcdonald and J. M. Peha, "Track Gap : Policy Implications of User Expectations for the "Do Not Track"

Internet Privacy Feature," *Communications, Information and Internet Policy (TPRC)* pp. 1–36, 2011.

[9] W. Melicher, M. Sharif, J. Tan, L. Bauer, M. Christodorescu, and P. G. Leon, "(Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking," *Proc. Priv. Enhancing Technol.*, vol. 2016, no. 2, pp. 135–154, 2016.

[10] J. P. Achara, J. Parra-Arnau, and C. Castelluccia, "MyTrackingChoices: Pacifying the Ad-Block War by Enforcing User Privacy Preferences," *Proc. Work. Econ. Inf. Secur.*, 2016.

[11] Wall Street Journal, "How Advertisers Use Internet Cookies to Track You." [Online]. Available: http://www.tamingdata.com/2010/10/18/how-advertisers-use-internet-cookies-to-track-your-online-habits.

[12] H. Taylor, "Most People are 'Privacy Pragmatists' Who, While Concerned About Privacy, Will Sometimes Trade it off for Other Benefits," *Harris Poll*, vol. 17, no. 19, p. 44, 2003.

[13] M. Rainie, L and Duggan, "Privacy and Information Sharing," *Pew Research Center*, 2015. [Online]. Available: http://www.pewinternet.org/2016/01/14/2016/Privacy-and-Information-Sharing/.

[14] J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, and J. Forné, "Online Advertising: Analysis of Privacy Threats and Protection Approaches," *Comput. Commun.*, vol. 100, pp. 32–51, 2016.

[15] V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, and S. Barocas, "Adnostic : Privacy Preserving Targeted Advertising," *Proc. Netw. Distrib. Syst. Secur. Symp.*, pp. 1–21, 2010.

[16] E. Wire, "Facebook Blocking the Ad Blockers: The Industry Reacts," *Exchange Wire*. [Online]. Available: https://www.exchangewire.com/blog/2016/08/17/facebook-blocking-ad-blockers-industry-reacts/. [Accessed: 20-Jun-2008].

[17] TNS Political and Social, *Cross-Border Access to Online Content (Report)*. European Commission, 2015.

[18] American Press Institute, "How Millennials Get News: Paying for Content," 2015. [Online]. Available: https://www.americanpressinstitute.org/publications/reports/survey-research/millennials-who-pays/. [Accessed: 20-Jun-2008].

[19] A. Lambrecht and K. Misra, "Fee or Free: When Should Firms Charge for Online Content?," *Manage. Sci.*, no. May, pp. 1–36, 2015.

Farah Chanchary is a doctoral student in the School of Computer Science at Carleton University in Ottawa, Canada. Her main research interests include algorithms and data structures, computational geometry and usable computer security. Contact her at farah.chanchary@carleton.ca

Yomna Abdelaziz is pursuing a MASc in Human-Computer Interaction in the School of Computer Science at Carleton University in Ottawa, Canada. She is interested in user-centred design as it applies to computer security and user privacy. Contact her at yomna.abdelaziz@carleton.ca

Dr. Sonia Chiasson is the Canada Research Chair in Human Oriented Computer Security and an Associate Professor in the School of Computer Science at Carleton University in Ottawa, Canada. Her main research interests relate to the human aspects of computer security and privacy with the goal of making security mechanisms easier and safer for people to use. Contact her at chiasson@scs.carleton.ca