Towards New Privacy Regulations in Europe: Users' Privacy Perception in Recommender Systems

Itishree Mohallick[1], Katrien De Moor[2], Özlem Özgöbek[1](✉), and Jon Atle Gulla[1]

[1] Department of Computer Science, NTNU, Trondheim, Norway
itishrem@stud.ntnu.no, {ozlem.ozgobek,jon.atle.gulla}@ntnu.no
[2] Department of Information Security and Communication Technology, NTNU, Trondheim, Norway
katrien.demoor@ntnu.no

Abstract. Despite the fact that recommender systems are becoming increasingly popular in every aspect of the web, users might hesitate to use these personalization-based services in return of their personal information if they believe their privacy is compromised in any possible way. While new privacy regulations in Europe bring more transparency and control over data collection to users, this study aims to provide a better understanding of the users' perception over privacy in recommender systems domain over several aspects such as behavioral preferences, privacy preferences, trust, data ownership and control over own data through an on-line survey. The results indicate that the majority of the respondents consider that recommender systems violate user privacy in different ways. Further, the results indicate that increased control and perceived sense of ownership over one's own data may help to decrease the negative attitudes towards recommender systems and providers and to re-instate and increase users' trust. However, the findings also indicate that users' trust may be hard to re-establish in cases where the thought of "apparently"/in theory go hand in hand with more transparency and user control will in reality/in practice not lead to drastic changes.

Keywords: Privacy · Recommender systems · Privacy perception EU GDPR · User study

1 Introduction

The presence of personalization-based systems in every corner of the World Wide Web has marked itself as a powerful tool. However, this optimistic outlook of such technologies possess a severe threat to user privacy due to their need of collection, processing and transfer of personal data [14]. The growing concern about the violation of privacy with the rise of ubiquitous recommendation technologies, is a potential research area for many researchers.

G. Wang et al. (Eds.): SpaCCS 2018, LNCS 11342, pp. 1–12, 2018.
https://doi.org/10.1007/978-3-030-05345-1_27

Recommender systems need to collect, store and process user's personal information in order to provide tailored services for individual users. This indeed is the primary source of user's privacy invasion in recommendation domain. Users are concerned about their online privacy which is found from various earlier researches [1–3,11,14]. Most of these research indicate that the privacy breach of user data is a result of directly accessing user data and indirectly, inferring from user's preference data or unauthorized usage by the external entities. A detailed research consisting of numerous survey results performed in the past decades concerning user's privacy perception supports the aforesaid user concern about privacy [10,19,20].

On the other hand, the protection of users' private data is a similarly important research topic in the modern information society. However the users' privacy behavior varies in practice as compared to the theoretical preferences. This is a privacy paradox [4,5,21] in the personalization-based systems and it is well studied through user-centric research. Generally, surveys on user privacy focus on information disclosure to on-line service providers. However, the sharing of user profiles across services, trust to the service provider, and control and ownership over personal data are less studied areas of user-centric research related to privacy and/in the recommender system domain.

The primary purpose of this study is not only to gain better insights into the privacy preferences, perception, and behavior of users in the context of recommender systems but also to explore the ways of increasing the trust on the service provider and user control over personal data. With the current enforcement of the EU GDPR (European Union General Data Protection Regulation), this research can shed light on the promises of the GDPR and remaining challenges in this respect. The new set of GDPR rules aim to bring two primary changes over the existing regulation. This sets a higher bar for the collection of personal data by the various service providers. In addition, the informed and explicit consent of the user is made mandatory while obtaining the user data. Secondly, the penalties for the non-followers (service providers who will not adhere to the new GDPR) are made severe enough. This study is based on an on-line survey on user's privacy perception in a season of rising privacy concerns in personalized systems, in which 200 participants of 28 different nationalities participated. The research was conducted at the Norwegian University of Science and Technology in Trondheim, Norway.

This paper is structured as follows. A brief background study related to user perception on privacy in recommender systems is given in Sect. 2. Section 3 explains the methodological approach and study set-up. Section 4 reports on the main findings from the study. Finally, Sect. 5 further discusses and summarizes the results and suggests a number of potential directions for future research on user perception in recommender systems.

2 Related Work

The concept of privacy is an explicit human perception which is inherently associated with data collection, data distribution (sharing) and re-use of the disclosed

data. In general, users prefer to share a fair amount of personal data for personalized recommendation purposes [4,7,13,23]. But at the same time, users express their concern for invasion of their information privacy through excessive data collection, incorrect inference and inappropriate after usage by these recommender systems.

Periodically, multiple surveys are performed in the past to understand user's online attitude with ubiquitous computing [1,15,19]. The largest survey conducted till date by the European Commission to find out European citizens attitude towards data protection, user privacy and identity management reveals the awareness, views and wish regarding the European user's data protection [8]. Given the three primary privacy threats in the recommender systems, namely: the recommender systems itself, other users of the systems and external entities, several privacy solutions have been adopted [13,16] over the last years. These measures and solutions range from technical solutions, algorithmic solutions, and more recently also legislative approaches such as EU GDPR. The recent commencement of EU GDPR is an approach to address the concerns expressed in the aforesaid survey and to protect the European user's personal data. The various key concerns to protect the fundamental rights and freedom of individual users added in GDPR [9] includes the following set of rules: (a) The right to be forgotten, (b) better control over who holds ones private data, (c) the right to switch ones personal data to another service provider, (d) the right to be informed in clear and plain language, (e) the right to know if your data has been hacked, (f) clear limits on the use of profiling, (g) special protection for children. This legislature also works towards the free movement of European data across all the member states.

As user's attitude towards privacy differs in different situations, it is important to study user perception on privacy within diverse cultural backgrounds and with different demographic aspects to explore the changes on user's privacy concern and behavior. In one such prior research [17], we have studied the privacy perspective of Norwegian users as compared to the other non-Norwegian users in a recommendation environment. As a continuation of the aforesaid approach, we have continued the data collection for the similar survey on diverse set of people. With an objective of understanding user's privacy attitude and information sharing behavior against different demographic setting, the user-centric survey is conducted after the GDPR is implemented.

3 User Study: Methodology

The primary objective of a survey strategy is to gather similar data from a group of people in an organized manner. Then the found statistical patterns are utilized to establish a general trend for a larger population [18]. In this study, data has been collected by using the on-line survey method in order to further investigate the users' broader perceptions of (violation of) user privacy, and attitudes towards recommender systems and service providers. In the following section we briefly present the included topics, before describing a number of key characteristics of the respondents.

3.1 Survey Design and Distribution

The survey was designed to include the privacy perception in recommender systems in general and in relation to specific recommender systems domains. Moreover, we also aimed to explore potential preferences in cross-domain recommendations and willingness of users to let their profile be shared across domains.

The survey consisted of 25 (both closed and open-ended) questions in English. From previous experience we know that users are not always aware of the fact that they are using recommender systems, as they often look like a natural part of the web page. This means that a regular Internet user may be exposed to several recommendations in a day, yet she or he may not be aware of it. Therefore, a brief description of recommender systems with a number of screen-shots of frequently used services was provided at the beginning of the survey, before any questions related to the use of and attitudes towards recommender systems were asked.

This short description was followed by a number of basic socio-demographical and general use-related questions. Next, the survey contained a number of questions related to perceived privacy issues and violation of privacy by and related to the use of recommender systems. Further, preferences in terms of recommender systems and cross-domain recommendations, as well as willingness to share one's user profile in a cross-domain environment were queried. Finally, questions concerned with trust and trustworthiness of recommender systems/providers and with data control and ownership were included.

The data have been collected in three phases in a total time span of 6 weeks. The first two phases were held in 2017 (May and July 2017) and the last phase was held in February 2018. The total number of participants to the user survey was 200 (52 participants in the first phase, 48 participants in the second and 100 participants in the third phase). A convenience sampling approach was used: the link to the survey was distributed via the involved researchers' own networks and existing channels. This approach has the advantage that it is more affordable as it draws on the recruitment of easily accessible subjects to the study. It has however also disadvantages as it may introduce a certain bias and does not allow for generalizations towards the entire population [12]. Yet, as the main goal with the study was not to generalize, but to further explore and gain better insights into users' perceptions and attitudes towards privacy issues with recommender systems, we consider the use of this sampling approach is justified here.

3.2 Sample Description

As mentioned above, in total 200 respondents filled out the entire survey. While they stem from 28 different countries, the majority are Norwegian (which is due to the fact that the survey was distributed through channels at the Norwegian University of Science and Technology). The top 4 nationalities participated to the user survey is as follows: Norwegian (39%), Indian (22.5%), Turkish (8.5 %) and Chinese (7%). In terms of gender, the sample consists of 68.5% male participants, whereas 31.5% are female. In terms of age, it should be noted that most of the participants (52%) are between 25 and 34 years old. This is followed

by the age group 35–44 which includes 22.5% of all the participants. Even though no formal question on the educational level of the participants was included, we assume that the sample primarily consists of participants with a relatively high education level, as the survey invitation was primarily distributed via networks and channels linked to the university.

4 Results

Before addressing attitudes related to (violation of) user privacy, trust, and data ownership and control, we briefly present a number of general usage-related aspects.

The majority of the participants (78.7%) can be considered as frequent users of recommender systems. Around 8 out of 10 frequent users report that they use recommender systems and personalized services on a daily basis (up to several times a day). We found no significant differences between male and female participants and the different age groups in this respect. Even though we did not collect fine-grained data on which specific services participants use (and how), data were collected on the importance of recommendations/personalized services in different domains (on a scale from 1 to 10). Overall, recommendations are considered more important in the domains of books, movies, and music compared to the domains of news and tourism. We found no gender differences in this respect, but for frequent users, analyses using the Mann-Whitney U test indicated that getting recommendations for books ($U = 2855, p < .05$, Median = 6), movies ($U = 2691, p < .01$, Median = 7) and music ($U = 2153, p = .000$, Median = 7) is significantly more important than it is for sporadic users.

As part of the further analysis, we systematically checked for potential differences in attitudes linked to usage frequency, basic demographical variables such as gender, age group, and where relevant, attitude-related variables. Finally, one more behavior-related variable was systematically included in this respect, and where relevant, we compare participants that already have requested to see their user profile or other information the provider has about them (32%) with those that have never done so (68%).

4.1 Violation of User Privacy

First we present the results related with the violation of user privacy. As it is shown in Fig. 1, less than 1 out of 10 respondents think that recommender systems that they use/have used respect laws and regulations on privacy and security. While around 1 out of 3 respondents clearly disagree here, it is interesting to note that more than half is in a grey zone: they do not know or are not sure.

When asked to which extent they think that existing recommender systems violate user privacy (on a scale from 1 to 10, where the lowest rating is 1 and the highest is 10), this uncertainty is also reflected. The average rating is 5.96, but the variance and spread of the values is rather large, indicating that the opinions are

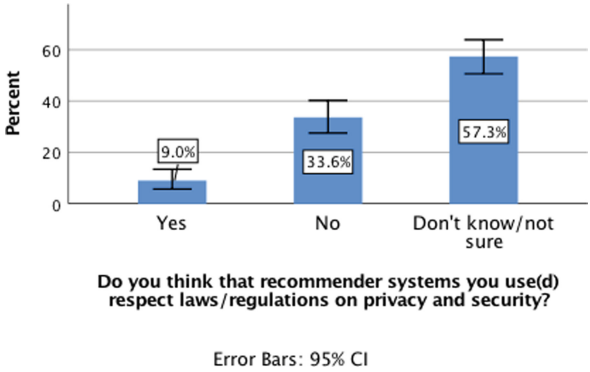

Fig. 1. Users' perception towards respect/violation of privacy laws and regulations

rather mixed. Even though no differences could be observed between the different age groups and frequent versus sporadic users, the data show a tendency that men are more skeptical in this respect and think to a significantly higher extent that recommender systems do violate user privacy ($U = 3781, p < .01$). This can be observed in Fig. 2, which also indicates that the opinions overall are rather mixed, also among men and women.

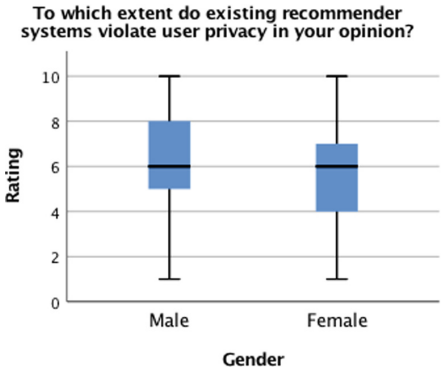

Fig. 2. Extent to which recommender systems are perceived to violate user privacy, by gender

As can be observed in Fig. 3, main concerns of users are the sharing of data with third parties (66%), collection of more data than what has been approved (49%) and the use of data for other purposes than what was approved (44%). Respondents also had the possibility to further elaborate on this issue in an open question. Here several participants noted that privacy terms and conditions represent a challenge in itself, as consent is often implicitly hidden in terms and conditions. As one respondent puts it: *"Often it can be difficult to know what is*

approved. You may not read all the terms before accepting because of long text and it being hard to understand".

4.2 Trust and How to Increase It

The findings presented above where only a minority of respondents think that privacy-related laws and regulations are respected by service providers and a majority perceives that privacy is violated in several ways, imply that trust in recommender systems/providers and their practices in terms of e.g., data sharing and usage, is rather low. This raises the important question of how trust to service providers could be improved and what recommender system users consider as essential elements that could contribute to the trustworthiness of recommender system providers. As shown in Fig. 4, the top 3 elements in this respect are: asking permission before using/sharing the users' data (77%), respecting the applicable laws and regulations (70%) and having the option to change and/or delete one's user profile (66%). Entries to the follow-up open question however also indicate that there is a user segment who will never really trust providers, as *"they are dependent on collecting and selling users' data"*. In addition, as some respondents point out, even if providers respect the laws and regulations, there is a risk of *surveillance* by governmental bodies. Finally, some respondents fear that even with some of the above options (e.g., being able to change and/or delete one's user profile), there will still not be real transparency and control over own data. As one respondent puts it: *"I don't trust 'delete' to be a true delete. Data is still kept"*. Another respondent even coins that some well-intended measures may *"give the appearance of transparency and control"*, while they in practice essentially may not change much.

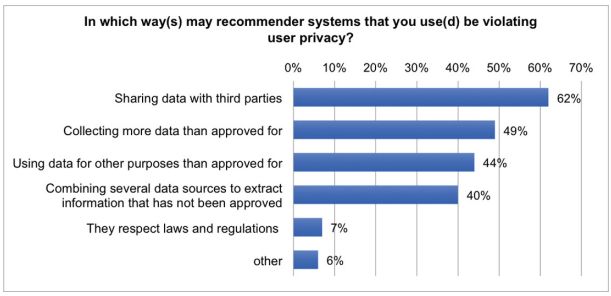

Fig. 3. Violation of user privacy by recommender systems, as perceived by the survey participants

While follow-up analyses did not yield any statistical differences between men and women or depending on the usage frequency or users' age in terms of the potentially trust-enhancing characteristics, we can observe a significant difference between respondents that already have at some point asked to see

their user profile and information that the provider has about them and those who never did so: the former consider the option to modify and delete one's user profile to a larger extent as a characteristic associated with a trusted provider ($\chi^2(1) = 4.602$, $p < .05$).

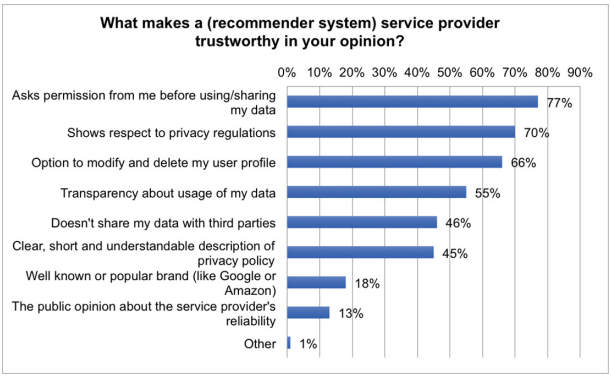

Fig. 4. Characteristics of trusted and trustworthy providers

The extent to which a service provider is trusted also seems to be relevant in the context of cross-domain recommendations. When asked whether they would be willing to let their user profile be shared across applications, 61% of the respondents reply negatively and they indicate that even if they would be able to choose the applications themselves, they do not want their profile to be shared. However, when asked to which extent this would be the case if the service provider is one that they trust, the number of respondents that says no drops to 41% and the willingness to let the user profile be shared under certain conditions increases considerably.

4.3 Control over Personal Data

Ownership and control over user's own personal data plays a crucial role in empowering users in the recommendation domain. As has been shown in previous studies, a minimal level of control and ownership increases the trust on the service provider and decreases the privacy concerns of users [22].

In the study presented here, it is also clear that a lack of trust is implicitly linked to lack of control over own data (sharing data with other parties, using data for other purposes than approved etc.). Providing users with real possibilities to control their own data (e.g., possibility to inspect, modify and delete their user profile) could increase both users' agency and empowerment, but could also result in more positive attitudes. As shown in Fig. 5, this possibility would make most respondents less worried about privacy risks when using recommender systems (63%) and would also increase their trust to service providers (52%). No

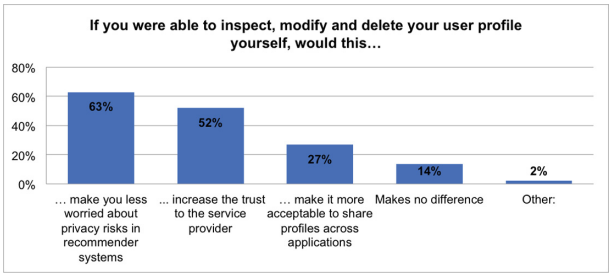

Fig. 5. Anticipated implications of being able to inspect, modify, delete own data and user profile

significant differences between groups (in terms of gender, usage frequency, age) could be observed here.

When it comes to data ownership and the use of recommender systems, the findings indicate that overall, respondents consider it as rather important to own their own data (average: 7.23, standard deviation: 2.31). For respondents who think that the applicable laws and regulations are not respected, data ownership is significantly more important than for those who do not know or are not sure ($H = 7.14(2), p < .05$).

However, what does "owning your own data" in the recommender systems domain actually mean for users? The results indicate that only a limited number of respondents (14%) associate this with storing data on one's own device. For the majority, ownership in this respect means being able to decide how one's data is shared (72%) and being able to modify and delete one's data (77%). Still, as mentioned in the beginning of this section, only slightly under 1 out 3 has already requested to see their online user profile or other information that the provider has about them, which indicates a potential discrepancy between the attitudes towards control and ownership over own data and respondents' actual behavior. In this respect, one respondent also underlined the importance of *"educating users"* by making them more aware of *"the value and amount of data that is collected about them on a daily basis"* and by making them more aware of their rights and possibilities when it comes to data control and ownership.

5 Discussion and Conclusions

In this paper, we have presented the results from an online survey on privacy perception of users in the recommender systems domain within three main categories: violation of user privacy, trust and data ownership. The main findings of the study can be summarized as it is shown in Table 1.

According to the results, there is no significant age group of usage frequency difference that affects the findings. However, gender seems to have an effect on the perception of privacy violation of recommender systems, where men are found to be more skeptical about their personal data than women.

Table 1. Summary of main findings.

Violation of user privacy	Only 10% of the users think that the recommender systems respect existing laws and regulations
	There is no significant perception difference between age groups and frequent versus sporadic users, however men are more skeptical than women about recommender systems violating user privacy
	Main concerns of the users are that the recommender systems share personal data with third parties, collect more data and use the personal data for other purposes than what has been approved
Trust	Trust to the recommender system service providers is found to be low
	Incorporating the users more by asking permission for sharing their data and giving the option for users to modify/delete their user profiles can increase trust
	Trust affects the common data usage for cross domain recommendations
Data ownership	For most of the users, ownership of personal data is important
	According to the users, ownership means deciding how the personal data is used/shared and being able to modify/delete it
	Ownership can increase the level of trust to the service provider by reducing user's privacy concern

With the new privacy regulations in Europe, EU GDPR seems to be the answer for some of these privacy concerns. As mentioned in Sect. 2, GDPR gives a better control to users over their own data which can decrease the privacy concerns of users in recommender systems domain. However, the concern about the recommender systems (or service providers) not respecting the laws and regulations may not be affected by the change of regulations. As some of the survey participants noted that a delete of a user profile by its owner may not be a true delete and it can still be available somewhere in the service provider's database implies that the service provider may not follow the regulations. Similarly, as noted by some participants, giving the appearance of transparency and control may not be a real solution to the privacy violations. Even though the GDPR is a big step towards the protection of privacy and fundamental rights of the users in Europe, the control of the service providers' respect to the regulations is important. EU GDPR strengthens the conditions of consent where the service providers have to provide an intelligent and easily accessible form for the users which must be written in simple language. The users no longer have to go through the hassle of long illegible terms and conditions full of legalese anymore

while giving their consent. Similarly, withdrawal of the user consent is made easier with the new regulation.

Recommender systems should primarily focus on privacy-driven user-centric approaches to accomplish the milestones of EU GDPR. In addition to this, recommender systems might have to change or adjust many of the currently used algorithms for individual decision making including profiling in order to comply with "the right to be forgotten". The practice of "privacy by design" [6] concept has the potential to attain the privacy regulation (GDPR) in recommender systems and the various service providers as well. However, the limitations of the aforesaid concept might restrain to gain complete privacy protection for the recommender systems. Sensible user engagement is thus required by these service providers, depending on the context. Also the privacy compatibility of worldwide online services remains several questions in the minds of users. The multicultural background of participants in this study shows that the users from all over the world share similar privacy concerns and requirements.

Acknowledgments. This work is a part of the master thesis which is supported by the NTNU SmartMedia program on news recommendation.

References

1. Ackerman, M.S., Cranor, L.F., Reagle, J.: Privacy in e-commerce: examining user scenarios and privacy preferences. In: Proceedings of the 1st ACM Conference on Electronic Commerce, pp. 1–8. ACM (1999)
2. Adams, A.: Users' perception of privacy in multimedia communication. In: Extended Abstracts on Human Factors in Computing Systems, CHI 1999, pp. 53–54. ACM (1999)
3. Antón, A.I., Earp, J.B., Young, J.D.: How internet users' privacy concerns have evolved since 2002. IEEE Secur. Priv. **8**(1), 21–27 (2010). https://doi.org/10.1109/MSP.2010.38. ISSN: 1540-7993
4. Awad, N.F., Krishnan, M.S.: The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Q. **30**(1), 13–28 (2006). http://www.jstor.org/stable/25148715
5. Barnes, S.B.: A privacy paradox: social networking in the United States. First Monday **11**(9) (2006). http://journals.uic.edu/ojs/index.php/fm/article/view/1394
6. Cavoukian, A., Fisher, A., Killen, S., Hoffman, D.A.: Remote home health care technologies: how to ensure privacy? Build it in: privacy by design. Identity Inf. Soc. **3**(2), 363–378 (2010)
7. Chellappa, R.K., Sin, R.G.: Personalization versus privacy: an empirical examination of the online consumers dilemma. Inf. Tech. Manag. **6**(2–3), 181–202 (2005)
8. European Commission: Attitudes on data protection and electronic identity in the European Union, June 2011. http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_359_en.pdf
9. European Commission: 2018 reform EU data protection rules, May 2018. https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

10. Cranor, L.F., Reagle, J., Ackerman, M.S.: Beyond concern: understanding net users attitudes about online privacy. In: The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy, pp. 47–70 (2000)
11. Fox, S., Rainie, L., Horrigan, J., Lenhart, A., Spooner, T., Carter, C.: Trust and privacy online: why Americans want to rewrite the rules. The Pew Internet and American Life Project, pp. 1–29 (2000)
12. Henry, G.T.: Practical Sampling, vol. 21. Sage, Thousand Oaks (1990)
13. Knijnenburg, B.P., Berkovsky, S.: Privacy for recommender systems: tutorial abstract. In: Proceedings of the Eleventh ACM Conference on Recommender Systems, pp. 394–395. ACM (2017)
14. Kobsa, A.: Tailoring privacy to users' needs[1]. In: Bauer, M., Gmytrasiewicz, P.J., Vassileva, J. (eds.) UM 2001. LNCS (LNAI), vol. 2109, pp. 301–313. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44566-8_52
15. Lederer, S., Mankoff, J., Dey, A.K.: Who wants to know what when? Privacy preference determinants in ubiquitous computing. In: Extended Abstracts on Human Factors in Computing Systems, CHI 2003, pp. 724–725. ACM (2003)
16. Mohallick, I., Özgöbek, Ö.: Exploring privacy concerns in news recommender systems. In: Proceedings of the International Conference on Web Intelligence, pp. 1054–1061. ACM (2017)
17. Mohallick, I., Özgöbek, Ö.: A survey on Norwegian user perspective on privacy in recommender systems. In: Proceedings of the 3rd Norwegian Big Data Symposium (NOBIDS 2017) in Conjunction with NxtMedia Conference 2017. CEUR Workshop Proceedings (2017)
18. Oates, B.J.: Researching Information Systems and Computing. Sage Publications Ltd., Thousand Oaks (2006)
19. Olson, J., Grudin, J., Horvitz, E.: Toward understanding preferences for sharing and privacy. In: Proceedings of the CHI (2004)
20. Olson, J.S., Grudin, J., Horvitz, E.: A study of preferences for sharing and privacy. In: Extended Abstracts on Human Factors in Computing Systems, CHI 2005, pp. 1985–1988. ACM (2005)
21. Pötzsch, S.: Privacy awareness: a means to solve the privacy paradox? In: Matyáš, V., Fischer-Hübner, S., Cvrček, D., Švenda, P. (eds.) Privacy and Identity 2008. IAICT, vol. 298, pp. 226–236. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03315-5_17
22. Taylor, D.G., Davis, D.F., Jillapalli, R.: Privacy concern and online personalization: the moderating effects of information control and compensation. Electron. Commer. Res. **9**(3), 203–223 (2009)
23. Toch, E., Wang, Y., Cranor, L.F.: Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. User Model. User-Adap. Interact. **22**(1–2), 203–220 (2012)