International Journal of Intelligent Information and Management Science
ISSN: 2307-0692, Volume 9, Issue 3, June, 2020

The Flow of Information Poses Challenges to Privacy

Yuedi Wu
College of Foreign Languages, Nankai University, Tianjin, 310000, China

Abstract: Big data era facilitates the flow of information, which could, accordingly, increase the risk of data breach. The essay aims at illustrating that the flow of information would pose challenges to privacy by discussing how the personal data that is directly collected and the data generated from pre-existing datasets can result in data breach. The first part analyzes how personal information trades and retreating phenomenon on social media cause the privacy leakage. The second part points out the threat that is presented by the inferred and derived data, which deserves more attention, yet few studies have covered that field.

Keywords: The flow of information; Data breach; Privacy

1. Introduction

Personal information that has been collected by a variety of methods can be applied and utilized in many fields; however, the fact is that even though the process of collecting personal information can be transparent, due to little knowledge of the mining process, the data subjects do not realize that their personal data could be disseminated and reused for many times, which might cause data breach and thus violate privacy (Manovich 6; Richterich 35; Cohen 1916). Many critics approach privacy challenges by analyzing how narrowing the limitation of access to the established datasets could help protect privacy (Manovich 5; Richterich 38). However, according to Boyd and Marwick, privacy is not simply binary – accessible or not (4). Rather, big data era facilitates the flow of information, which could increase the risk of data breach into a wider range. This essay argues how the flow of information, which bases on either the information collected directly from individuals or the generated personal data on the premise of pre-existing datasets could pose rising challenges to privacy[1-2].

2. Original Problem

Admittedly, access to personal information could still pose great challenges to privacy. Some unproper ways of accessing to personal information could cause data breach, and access to datasets in the excuse of researching or surveillance might conflict with privacy protection sometimes (Cohen 1914; Manovich 1). To begin with, permission-based privacy cannot completely protect personal data recorded on devices, for the applications may be attacked, which would result in personal information leakage (Park et al. 29). Despite security improvement, there is still a high risk of data breach, since by decoding the root privileges in one's phone, some malicious programs can run automatically and read personal data without getting the users' permission. Some data could be copied and even used in illegal ways [3-4]. Besides the unproper use, government surveillance upon electronic communication on the phones and conducting humanity research based on personal information could also spark off an intensive debate (Manovich 6). Specially, big data era brings a massive database to scholars who can work with "born-digital user-generated" content from a wide range of sources, like locations, photos uploaded by users and online communication, yet it might violate individuals' privacy (Manovich 2; Burgess and Bruns 3). After examining the controversial ways aforementioned, apparently, personal information is mined without the direct permission from the data subjects in these cases, though the collection process might gain their consent [5-6]. However, people seldom realize there would be multiple and unforeseen use after signing Terms of Services or other agreements, which would allow for data collection and possible analysis (Boyd and Crawford 673; Mittelstadt and Floridi 12; Latour 2). Therefore, even though there is a strictly regulated limitation of the access to datasets, due to a "grey area" lying in the informed consent that little agreement is notified explicitly , websites and applications' users would be unlikely to keep aware of the possible use of their data in the future (Mittelstadt and Floridi 12; Richterich 44).

3. Threat from the Flow of Information

Even though the fact that the regulation of certain access to personal data is to be strengthened, more attention should be paid to the flow of information or data sharing phenomenon, for single-instant consent in one website or application cannot stop the collected data from being shared by the dataset establishers to more platforms (Mittelstadt and Floridi 12). However, few Terms of Services clearly regulate concerning issues and the flow of infor-

mation becomes more and more common on social media [7-8]. On the one hand, the dataset establishers may use personal data in a commercial way by selling it as a product to a wide range of customers such as researchers who would use it as the media for humanity study (Berry 4). According to Mattioli, some companies might treat the datasets as an intellectual property and enhance the value of data by narrowing the access to the datasets (537). Normally, datasets are inaccessible to "civic actors", but by paying for them, researchers and customers could get access to the data (Richterich 38). However, when data subjects sign up the single-instant consent given from such companies, they might be uninformed or unaware of the potential commercial profits and unforeseen sharing of their data. And due to the purpose of economic benefits, companies would be more likely to trade on personal information via internet, despite the fact that accessibility to the data stored in one company does not mean that such data can be consumed and reused by others (Boyd and Crawford 672). Some companies tend to anonymize the information before carrying out the trade to make such a circulation more convenient, but it remains uncertain that such anonymization would last for long and still, the trade violates data subjects' privacy (Richterich 38).

In addition to personal data trades, the flow of information is also caused by retweeting others' online information published on social media. Centralized social media such as Twitter or Facebook create a "Social Space" for users, where they can share daily life to the "intended recipients" (Kapanipathi 2). However, online messages that have been published can be retreated by their followers; even though users can manage the viewers, the phenomenon of retreating would make the protected tweets copied and spread to others without the consent given by the original author (Kapanipathi 2; Meeder 1). Users might learn about the great openness of social media, yet they remain unaware of the extent to which their information can be disseminated by retreating [9]. For instance, according to Meeder, who studied the privacy policy of Twitter in the registration process, detailed private settings often appear at the branching steps, where few people would pay attention to (Meeder 2; Boyd and Marwick 6). Users, therefore, have to thoroughly learn the long policy and specifically seek out the privacy terms if they want to set tweets protected. Moreover, even a protected tweet can still be retreated due to the fact that only a warning message would appear when others try to operate retreating, yet Twitter still enables the retreating (Meeder 3). Thus, even when access to users' online information can be strictly set by the data subjects, they cannot control data sharing and the flow of information caused by followers' retreating, and some of the retreating could leak private information and be deliberately used in other ways [10].

The flow of information that has been discussed above, however, only concerns with the information that has once been provided by individuals initially; no matter what ways of collecting personal data, individuals engage into the process of collection publicly. There is another type of personal data called inferred and derived data that is increasingly prevalent in the era of big data, which is obtained without individuals' involvement, yet it still brings to an increasing circulation of personal information via internet (Abrams 5). The data can either be valuable for researchers to further humanity study or increase the risk of privacy leakage. Specifically, the data is inferred and computerized from the pre-existing datasets (Abrams 5). This paragraph will not focus on the introduction of such data but briefly analyze how it can threaten privacy. Since such data is generated based on pre-existing information and often serves the purpose of commercial use, with a limited awareness of its generation, individuals seldom find their personal data processed and mined to gain profits by some companies, for example, in medial field, future health outcomes can be calculated based on the existing datasets and used to find potential medicine customers (Abrams 8). The process of data deduction and computerization from other datasets actually enables the flow of information. According to Eibl and Engel, load data can help conduct the information on appliances (1). Huafeng also claims that such way of generating data could make some personal information that has not been exposed appear (501). However, this field is still to be studied further, for few materials and essays can be found; on top of that, the challenges posed by the derived and inferred data need more experiment to tackle.

4. Conclusion

To conclude, on the basis of access to personal information discussed in the literature review, more attention and a deeper insight is given to the flow of information, for it poses new challenges to privacy but receives less concerns. Specifically, not only the trades on personal data carried out by some companies but also the retreating phenomenon on social media could cause data breach. Also, the data that is deduced and generated from other datasets could expose much private information to more viewers. Therefore, the awareness of controlling the flow of information should be raised. However, the study of the challenges posed by the derived and inferred data to privacy still remains in its infancy; thus, more research and experiment need to be implemented.

References

[1] Smaradottir B., Fensli R. A case study of the technology use and information flow at a hospital-driven telemedicine service. Studies in health technology and informatics. 2017, 244, 58-62.

[2] Harush U., Barzel B. Dynamic patterns of information flow in complex networks. Nature Communications. 2017, 8(1), 2181-2121.

[3] Carr Adrian. The challenge of critical theory for those in organization theory and behaviour: an overview. Neuroendocrinology. 2018, 1(1), 4-14.

[4] Imanimehr F., Fallah M.S. On transparent value-sensitive run-time monitoring for information flow policies. Computer Languages, Systems & Structures. 2018, 54(DEC.), 273-296.

[5] Do Q.H., Bubel R., Hahnle R. Automatic detection and demonstrator generation for information flow leaks in object-oriented programs. Computers & Security. 2017, 67(JUN.), 335-349.

[6] Wu L. A transformation-based multi-area dynamic economic dispatch approach for preserving information privacy of individual areas. IEEE Transactions on Smart Grid. 2019, 10(1), 722-731.

[7] Yong-Woon K., Namin C., Hye-Jung J. Trends in research on the security of medical information in korea: focused on information privacy security in hospitals. Healthcare Informatics Research. 2018, 24(1), 61-68.

[8] Dogruel L. Privacy nudges as policy interventions: comparing US and German media users' evaluation of information privacy nudges. Information Communication & Society. 2019, 22(7-8), 1080-1095.

[9] Presthus W., Vatne D.M. A survey on facebook users and information privacy. Procedia Computer Science. 2019, 164, 39-47.

[10] Ometov A., Bezzateev S., Davydov V., et al. Positioning information privacy in intelligent transportation systems: an overview and future perspective. Sensors. 2019, 19(7), 1-23.