



## The Rise of Crowdsourcing Using Social Media Platforms: Security and Privacy Issues

Noor Hafizah Hassan<sup>1</sup> and Fiza Abdul Rahim<sup>2\*</sup>

<sup>1</sup>Advanced Informatics School, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia

<sup>2</sup>College of Computer Science and Information Technology, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia

### ABSTRACT

The increasing adoption of social media is a viable means in crowdsourcing. It can facilitate the connectivity of collaboration between different organisations, people and society to produce innovative and cost-effective solutions to many problems. Social media have opened up unprecedented new possibilities of engaging the public in meaningful ways through crowdsourcing. However, the growing number of security and privacy issues in social media may weaken the efficacy of crowdsourcing. This study aims to provide a basic understanding of security and privacy issues in line with the growth of crowdsourcing using social media platforms. This study also illustrates how crowdsourcing and social media data can lead to security and privacy issues in different environments. Lastly, this study proposes future works that may serve as direction for scholars to explore security and privacy in crowdsourcing through social media platforms. Secondary sources obtained from journals, conference papers, industry reports and books were reviewed to gather information.

*Keywords:* Crowdsourcing, privacy, security, social media

### INTRODUCTION

The term ‘crowdsourcing’ is becoming well known among organisations in various fields

as well as the interdisciplinary research community. Howe (2006) coined the term by combining the words ‘crowd’ and ‘outsourcing’. This study uses the definition of crowdsourcing by Halder (2014a, pp377-393) as “the process of finding needed information and service for a common goal from a large number of people.” Over the last few years, types of crowdsourcing have expanded rapidly, allowing people to connect with one another through crowd voting, crowd creation, crowd wisdom and crowd funding.

#### ARTICLE INFO

##### *Article history:*

Received: 09 March 2017

Accepted: 18 September 2017

##### *E-mail addresses:*

noorhafizah.kl@utm.my (Noor Hafizah Hassan),

fiza@uniten.edu.my (Fiza Abdul Rahim)

\*Corresponding Author

In recent years, social media have become ubiquitous and important in social networking as well as content sharing. Social media, such as social networking, media sharing, blogging and tweeting, have been used worldwide as an important tool for communication (Rahim, Ismail, & Samy, 2014). Social media also act as an intermediate platform for raising support. By establishing relationships, social media manage to create value with the participation of its users. People connect with one another based on the values that they share. Both crowdsourcing and social media owe their presence to the influence of networks. When different individuals perform a specific common task, the output may vary from one to another. However, when people come together to work towards a common goal, great outcomes can be achieved.

Crowdsourcing rides on this wave of creativity, empathy and support that combines the contributions from a crowd of people to build completely new ventures and sharply radical ideas. For example, crowdsourcing can take the smallest ripple in marketing and form great waves of ideas that result in new ways of thinking by democratising the process of starting and running a business. The power of crowds makes any ambitious concept become possible. Social media help to spread the word about crowdsourcing projects and reach out to intended users who might be able to contribute meaningfully towards a new project.

The use of crowdsourcing in different domains not only makes it possible to mine, aggregate and categorise data but also helps in the readiness to face a certain situation, reaction during the situation and recovery after the situation (Halder, 2014a). However, the increasing usage of crowdsourcing in social media has also given rise to certain issues and challenges. This study investigates security and privacy issues related to crowdsourcing activities that use social media. The first section of this paper introduces crowdsourcing usage in social media platforms. The second section reviews the various usage of crowdsourcing in social media platforms. The third section describes the methodology used to achieve the aim of this study. The fourth section discusses the security and privacy issues related to crowdsourcing in social media. Finally, the conclusion summarises the output of this study.

## **Crowdsourcing Usage in Social Media Platforms**

Social media platforms are believed to be one of the most influential sources in capturing the latest information involving many participants in crowdsourcing. ‘YouPartTime’ and ‘MyTeksi’ are examples of crowdsourcing platforms that are linked to social media in Malaysia for searching for part-time jobs and taxi booking (Habbal et al., 2015). In this section, before identifying the security and privacy issues, some aspects of crowdsourcing in social media will be briefly described. Some crowdsourcing activities that utilise social media platforms are disaster relief, facial attributes and depression prediction.

### **Disaster Relief**

One of the crowdsourcing activities that use social media is disaster relief. People are able to report and update latest information on disasters that may happen or have already happened for immediate response. Data analysis and collection relating to the disaster can also be conducted. From a few cases involving catastrophic earthquakes in Haiti, Afghanistan, Kenya and Mexico, a crowdsourcing tool known as ‘Ushahidi’ ([www.usahidi.com](http://www.usahidi.com)), for instance, was

able to gather latest information from multiple sources including social media (Gao, Barbier, & Goolsby, 2011). Previously, Zook, Graham, Shelton and Gorman (2010) demonstrated that different types of crowdsourcing information and mapping services can greatly enhance the logistical systems upon which relief efforts are ultimately grounded.

Another crowdsourcing tool involving social media is called ‘Tweak the Tweet’, which uses the Twitter platform for updating crowdsourced information during mass emergencies (Starbird & Palen, 2011). This tool even provides rewards for the participants involved in reporting any related incidents. An open-source web platform developed by Rogstadius et al. (2013) known as ‘CrisisTracker’ is able to track keywords on Twitter and construct stories by clustering related tweets from lexical similarity to promote disaster awareness. The Crowdsourcing Disaster Support platform (CDSF) developed by Yang et al. (2014) is able to provide real-time assistance that extracts input from social networks, learning from the past and historical data for instant communication services.

**Facial Attributes.** Crowdsourcing in social media also contributes facial attributes that provide information about a person based on the communities’ uploaded images and video collections. Facial attributes can be detected from leveraging photos in social media and removing ‘noisy photos’ by mid-level feature voting from crowdsourcing methods (Chen, Hsu, & Liao, 2011). Not only that, facial expression and the affect recognition system that utilises crowdsourcing methods gather information from videos uploaded online and are able to view natural and spontaneous responses over the web (McDuff, El Kaliouby, & Picard, 2011).

**Predicting Depression.** Social media platforms are also crowdsourcing tools for identifying depression among users based on the postings. De Choudhury, Gamon, Counts and Horvitz (2013) examined the ability of image postings on Twitter that can be used to measure and predict depression among individuals with 70% accuracy. In another study by Nadeem, Horn and Coppersmith (2016), crowdsourcing platforms were employed to compile a list of Twitter users who had professed to being diagnosed with depression. They successfully demonstrated the potential of using Twitter as a tool for measuring and predicting major depressive disorders in individuals. The ability to estimate, generalise and interpret daily variations in depression may be used in a medical context to identify clinical depression from the behaviour of social media users. The potential in identifying depression prior to the onset of mild depression can also help save lives.

## METHOD

The review of security and privacy issues of crowdsourcing in social media platforms was based on secondary sources. The information obtained from journals, conference papers, industry reports and books was summarised to develop understanding on security and privacy issues. The online databases that were given particular attention include Springer Link, Science Direct, Emerald Library, IEEE Explorer Digital Library, Taylor & Francis Online and EBSCO host. Search terms were used to find relevant articles included crowdsourcing, social media, security and privacy. The inclusion criteria for the articles were predetermined to be: 1)

Articles published in English, 2) Full-text articles, and 3) Security or privacy issues related to crowdsourcing in social media platforms. Because of the limited number of literature found, this study looked at the basic understanding of security and privacy that has been previously discussed and the issues that society may face as crowdsourcing in social media continues to grow.

## **Security and Privacy Issues**

Crowdsourcing can bring additional risks since these services are often outsourced or indirectly managed. With the increasing number of crowdsourcing activities available through social media, participants are exposed to security and privacy issues. In the next section, the security issues discussed pertain to data security, confidentiality and integrity, while privacy issues focus on privacy violations, identifiable information and information sharing.

### **Security Issues**

The use of crowdsourcing platforms through social media has become one of the main choices by organisations for providing services. However, using social media as a platform may expose users to a few security issues. Information security makes use of some measurement in order to protect information supported to crowdsourcing platforms (Cilliers & Flowerday, 2015).

**Data security.** With the increasing usage of crowdsourcing in many different forms, many organisations may be able to virtually spy on political persons (Halder, 2014b). Rapid use of mobile applications for crowdsourcing activities may expose users to security issues. This includes revealing location information involving mobile movement and also personal data stored in mobile devices. The availability of such data through crowdsourcing platforms that use social media can lead to a rise in illegal activities engineered by cybercriminals (Shiffman & Gupta, 2013). For example, if a cybercriminal is able to spoof one identity, he or she would be able to gain access to many accounts such as personal email, bank or health records.

Individuals may have malicious intent towards the goal, sponsoring organisation or individuals, individual participants or underlying infrastructure in a crowdsourcing solution. The anonymous data stored from different sources in social media may endanger consumer data, thus violating data security regulations (Wolfson & Lease, 2011). Hence, legal issues related to data security must be addressed in order to avoid misuse of personal information that will further lead to other issues. With the open-call format concept of crowdsourcing, conflict of data between participants in crowdsourcing platforms and crowd developers may also lead to intellectual property issues (Simula, Töllmen, & Karjaluo, 2015).

**Confidentiality.** Confidentiality is related to data security and is associated with ensuring information exposed in social media will not be misused in the crowdsourcing platform. An example of security risks reported on sensitive data exposure is the Netflix cases. Netflix is a type of social media that provides video content delivery services and has being sued by consumers for illegal use of consumer data (Lieberstein, Tucker, & Yankovsky, 2012).

However, Sarwar and Khan (2013) stated that once a citizen has posted information on a crowdsourcing platform, the ownership of the data is no longer that of the citizen. This shows that the confidentiality of data has not been secured appropriately in crowdsourcing platforms. This increases the chances of identity theft among consumers. Therefore, there is a need for a security policy that can protect information belonging to participants of crowdsourcing. Lease et al. (2013) identified Amazon's Mechanical Turk as one of the powerful crowdsourcing tools that have claimed that online work is anonymous but sensitive information regarding participants can be revealed.

**Integrity.** Users were also concerned about the integrity of the data gathered in crowdsourcing platforms. Since the data from social media can be in any form, it has been reported that users were very concerned about modification of stored information (Cilliers & Flowerday, 2015). Hence, there should be administrative control of the crowdsourcing platform that is able to ensure data integrity and accuracy posted through social media (Barbier, Zafarani, Gao, Fung, & Liu, 2012). For example, using crowdsourcing platforms in the medical environment requires accuracy for the data collected. Therefore, it is very important that the integrity of the data posted through the crowdsourcing platform be maintained to ensure that the crowd are receiving correct information.

**Privacy Issues.** Social media can be a powerful tool for influencing and educating (Jean Barry MSN, 2012). However, the emergence of various social media sites has changed the perception of privacy (Halder, 2014b). Furthermore, it may also potentially impact negatively on professionalism, ethics and privacy in some professions (Rahim, Ismail, & Samy, 2014). Using crowdsourcing in the social media platform may also potentially pose a threat to the privacy and protection of the users' personal and sensitive data.

**Privacy violations.** As the popularity of crowdsourcing escalates, the rate of 'crowd attacking' will also grow. Important information providers are being attacked both physically and virtually. Massive data including status updates, personal thoughts about certain issues or ideas, also known as 'crowdsourced information' or 'user generated content', are also unprotected. For example, governments and other parties are able to virtually spy on any person if they wish to.

Detailed information about individuals, mobile numbers, IP addresses and geographical locations can be violated as masses of data are easily collected through social media platforms (Halder, 2014a). Different types of information are being collected through crowdsourcing that may lead to identity profiling, where health status, age, gender, race, religion, political ideology, sexual orientation and other details of individuals can be guessed. Thus, the information providers of crowdsourcing can become potential victims of privacy violation.

**Identifiable Information.** Geography tagging has become a popular feature on several social media platforms, such as Facebook and Instagram. It is also known as geo-tagging, referring to a process of adding geographical identification on the online post. Facebook users can geo-tag photos that can be added to the page of the location where the picture is tagged. Users

may use this feature to find nearby Facebook friends by spawning a list of people based on the location tracker in their mobile devices. Instagram uses a map feature that allows users to geo-tag photos, thus allowing users to link specific photos on a world map.

This shared geo-tag information may also contribute to negative impact with regards to privacy (Huang & Gartner, 2016). For instance, there is a possibility of accidentally revealing a home address when a photo has been geo-tagged in front of a home. Any geo-tagged photo that reveals a school logo may possibly expose the location of the school to potential child-crime offenders. Such geo-tagged photos or information can reveal personal information, and this can cause harm if the information falls into the wrong hands.

**Information Sharing.** Today, people share sensitive information about themselves on social media to obtain feedback from virtual community members (Denecke et al., 2015). For instance, Facebook users openly sought and shared information relating to behavioural, mental and genetic health by publicly revealing their name, photo and location when seeking sensitive health information using social media (Househ, 2011). Another example is when parents share their child's health information along with pictures online, before a public audience, as happened on CrowdMed (Denecke et al., 2015).

The purpose of information sharing is possibly to exchange information and share experience. At the same time, users can seek emotional support and request guidance and advice from healthcare professionals and others via social media. Through crowdsourcing using social media platforms, patients can ask for a second or third opinion of a diagnosis or treatment (CrowdMed, 2010). However, online interaction with patients may pose risks because of the uncertainty related with the written language, which may not represent the actual context of body language or the lack of awareness of potential abuse with regards to social media data (Chretien & Kind, 2013). These risks might lead to a wrong diagnosis or treatment suggestion and even place the individual in danger.

Therefore, there is a need to look into matters concerning privacy with regards to using social media. Privacy breaches may cause great harm when they occur online. This is because social media can potentially be accessed even globally because of the nature of digital information. More research should be done to develop privacy-preserving techniques for crowdsourcing and social media data analysis. The ethical issues concerning methods of sharing information, obtaining consent and using information for research or commercial purposes should be the primary concern in protecting the privacy of individuals who use social media platforms.

Table 1 summarises security and privacy issues of crowdsourcing using social media platforms that have been discussed in previous sections.

Table 1  
*Summary of security and privacy issues of crowdsourcing using social media platforms*

Issues	Details	References
<b>Security</b>		
• Data Security	Anonymous data stored from different sources in social media may endanger consumer, thus violating data security regulations.	Shiffman and Gupta, 2013; Simula, Töllmen, & Karjaluoto, 2015
• Confidentiality	Sensitive information related to individuals or organisations might be exposed in social media.	Lieberstein, Tucker, & Yankovsky, 2012; Sarwar and Khan, 2013
• Integrity	Maintaining the integrity of the data posted ensures that the crowd receive correct information.	Cilliers and Flowerday, 2015
<b>Privacy</b>		
• Privacy Violations	Collected information may violate privacy of individuals or organisations.	Halder, 2014a.
• Identifiable Information	Personal information, which can lead to harm if it falls into the wrong hands, can be revealed.	Huang and Gartner, 2016
• Information Sharing	Collected information may not represent the actual situation and the possibility of abuse of social media data might lead to wrong perceptions.	Househ, 2011; Denecke et al., 2015; CrowdMed, 2010

## CONCLUSION

Security and privacy are important matters to consider in crowdsourcing. As the use of crowdsourcing in many platforms increases, initiative should be taken to develop the guarantee of security and privacy in crowdsourcing from an ethical, legal and technological context. Three security issues were discussed in this study: data security, confidentiality and integrity. In addition, three privacy issues were reviewed: privacy violations, identifiable information and information sharing. The review highlighted key issues pertaining to crowdsourcing using social media platforms. This review provides useful information and knowledge, highlighting both security and privacy as crucial issues in crowdsourcing using social media platforms. Due to the limited number of security and privacy studies on crowdsourcing, this study was only able to refer to few studies available in the literature on social media platforms. Extensive literature is needed to accurately assess existing crowdsourcing tools to gauge the level of current security and privacy in this area. Therefore, further research in various fields is needed to investigate appropriate mechanisms to ensure the security and privacy of users of crowdsourcing using social media platforms.

## REFERENCES

- Barbier, G., Zafarani, R., Gao, H., Fung, G., & Liu, H. (2012). Maximizing benefits from crowdsourced data. *Computational and Mathematical Organization Theory*, 18(3), 257-279.
- Chen, Y. Y., Hsu, W. H., & Liao, H. Y. M. (2011, March). Learning facial attributes by crowdsourcing in social media. In *Proceedings of the 20<sup>th</sup> international conference companion on World wide web* (pp. 25-26). ACM.
- Chretien, K. C., & Kind, T. (2013). Social media and clinical care: Ethical, professional, and social implications. *Circulation*, 127(13), 1413–1421.
- Cilliers, L., & Flowerday, S. (2015). *The Relationship between Privacy, Information Security and the Trustworthiness of a Crowdsourcing System in a Smart City*. In HAISA (pp. 243-255).
- CrowdMed. (2010). *CrowdMed. Frequently-asked questions*. Does CrowdMed need to be HIPAA compliant (in the US)?
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). *Predicting Depression via Social Media*. *ICWSM*, 13, 1-10.
- Denecke, K., Bamidis, P., Bond, C., Gabarron, E., Househ, M., Lau, A. Y. S., ... & Hansen, M. (2015). Ethical issues of social media usage in healthcare. *Yearbook of Medical Informatics*, 10(1), 137.
- Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems*, 26(3), 10-14.
- Halder, B. (2014a). Evolution of crowdsourcing: Potential data protection, privacy and security concerns under the new media age. *Democracia Digital E Governo Eletrônico*, 1(10), 377–393.
- Halder, B. (2014b). Measuring security, privacy and data protection in crowdsourcing. *Privacy and Data Protection in Crowdsourcing* (2014, July 20).
- Househ, M. (2011). Sharing sensitive personal health information through Facebook: the unintended consequences. *Studies in Health Technology and Informatics*, 169, 616-620.
- Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine*, 14(6), 1-4.
- Huang, H., & Gartner, G. Using mobile crowdsourcing and geotagged social media data to study people's affective responses to environments. *European Handbook of Crowdsourced Geographic Information*, 385.
- Jean Barry MSN, R. N. (2012). Advancing nursing practice through social media: a global perspective. *Online Journal of Issues in Nursing*, 17(3), 1\_9.
- Lease, M., Hullman, J., Bigham, J. P., Bernstein, M. S., Kim, J., Lasecki, W. ... & Miller, R. C. (2013). Mechanical Turk is not anonymous. *Social Science Research Network*, 15.
- Lieberstein, B. M. A., Tucker, A., & Yankovsky, A. K. (2012). Crowdsourcing: Understanding the risks. *NYSBA Inside*, 30(2), 34–38.
- McDuff, D., El Kaliouby, R., & Picard, R. (2011). Crowdsourced data collection of facial responses. *The 13<sup>th</sup> International Conference*, 11–18.
- Nadeem, M., Horn, M., & Coppersmith, G. (2016). Identifying depression on Twitter. *CoRR*, 1–9.

- Rahim, F. A., Ismail, Z., & Samy, G. N. (2014). Information Privacy Concerns in the Use of Social Media Among Healthcare Practitioners: A Systematic Literature Review. *Advanced Science Letters*, 20(10-11), 2176-2179.
- Rogstadius, J., Vukovic, M., Teixeira, C. A., Kostakos, V., Karapanos, E., & Laredo, J. (2013). CrisisTracker: Crowdsourced Social Media Curation for Disaster Awareness. *IBM J. Res. Dev.*, 57(5).
- Sarwar, A., & Khan, M. N. (2013). A review of trust aspects in cloud computing security. *International Journal of Cloud Computing and Services Science*, 2(2), 116–122.
- Shiffman, G. M., & Gupta, R. (2013). Crowdsourcing cyber security: A property rights view of exclusion and theft on the information commons. *International Journal of the Commons*, 7(1), 92–112.
- Simula, H., Töllmen, A., & Karjaluo, H. (2015). Facilitating innovations and value co-creation in industrial B2B firms by combining digital marketing, social media and crowdsourcing. In L. Robinson Jr. (Ed.), *Marketing Dynamism & Sustainability: Things Change, things stay the same.... Developments in marketing science: Proceedings of the Academy of Marketing Science*. Springer, Cham.
- Starbird, K., & Palen, L. (2011). Self-organising by digital volunteers in times of crisis. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1071–1080).
- Wolfson, S. M., & Lease, M. (2011). Look before you leap: Legal pitfalls of crowdsourcing. *Proceedings of the ASIST Annual Meeting*, 48.
- Yang, D., Zhang, D., Frank, K., Robertson, P., Jennings, E., Roddy, M., & Lichtenstern, M. (2014). Providing real-time assistance in disaster relief by leveraging crowdsourcing power. *Personal and Ubiquitous Computing*, 18(8), 2025–2034.
- Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake. *World Medical and Health Policy*, 2, 7–33.

