GDPR Case Study: 2020-12-15, Twitter International Company, Ireland

Barry Zhang (mengze_zhang@brown.edu)

Brown University

Abstract

On 8 January 2019, Twitter International Company (TIC) notified a data breach to Ireland Data Protection Commission [1]. The breach was due to a flaw in Twitter's Design, and it has caused a leakage of over 88,000 European Union and European Economic Area Twitter users' private information [2], namely, their protected tweets. After a nearly two-year inspection, the final decision of the Data Protection Commission was made on 9 December 2020, and Twitter was charged a fine of €450,000 [2] for violation of General Data Protection Regulation (GDPR).

In this paper, I examine the case and propose two improvements that could be made by Twitter and other similar entities, one is focused on redesigning the underlying infrastructure, the other is focused on optimizing bug tracking and reporting system. Apart from that, I also evaluate the GDPR enforcement process and discuss the effectiveness and fairness it has shown in this case.

1. Background

Twitter is a microblogging company that allows its users to share their opinions and moments of their life through posting tweets [3]. It offers a service known as 'Protect your Tweets', if a user chose to activate this service, then all the tweets of this user are considered as protected tweets and only accessible to this user's followers.

However, due to a bug in the code introduced in 2014 [1], the protected tweets of users who use Twitter on Android devices may be exposed to the public, if the users made certain changes to their accounts like changing their emails [4]. It is clear to see that the data objects here in the case are the users who use Twitter on Android and the personal data exposed is their protected tweets.

After receiving a data breach notification from Twitter, the Ireland's Data Protection Commission started to inspect into this situation. In their final decision, they decided that TIC violated GDPR Article 33(1) and 33(5) for not fulfilling their duties as the Controller. Although, from what have been confirmed by TIC [1], it is not hard to find out that the Processor in this case, Twitter Inc., should also be blamed for this violation.

2. GDPR Violation

Even though the root cause of Twitter's violation is the leakage of tens of thousands of users' private information, the direct reason why Twitter was punished is that they failed to comply with Article 33(1) of GDPR, which requires the Controller to report any data breach to the supervising authorities on time [5], and Article 33(5), which states that the Controller needs to properly document any data breach they find [5]. The difference between the root cause and the direct cause suggests two different approaches to avoid violations like these in the future.

2.1 Details of the violation

I consider there to be three violations committed by Twitter in this case.

2.1.1 Violation of Article 33(1) and 33(5)

According to what TIC has presented to the Data Protection Commission, TIC reported the data breach on 8 January 2019 because its Processor Twitter Inc. first assessed this breach on 3 January 2019 but did not notify it until 7 January 2019 [1]. The Commission decided that TIC should have reported the breach 'at the latest by 3 January 2019' [1], thus TIC's delayed notification constituted a violation against Article 33(1).

Also, during the investigation, the Commission found out that the documentation provided by TIC was insufficient since '… the report does not contain any reference to, or explanation of, the issues that led to the delay in TIC being notified of the Breach. In addition, the Incident Report does not address how TIC assessed the risk, arising from the Breach, to affected users.' [1]. This lack of detailed record was considered a direct violation of Article 33(5).

It's not hard to find out that both these two violations are the result of human errors: engineering team at Twitter Inc. failed to follow privacy breach protocol and delayed report of a potential leakage; legal team at TIC failed to thoroughly examine GDPR requirements and provided insufficient documenting records to the regulating authorities.

2.1.2 Leakage of protected tweets

In the final decision, the data breach itself was not deemed as a direct violation to any GDPR article. However, since it has caused the leakage of many users' protected tweets and thus directly violated users' privacy, I would still like to consider it as a violation by Twitter.

The data breach was caused due to a bug introduced in 2014 [1], this bug made that 'if a user operating an Android device changed the email address associated with that Twitter account, their tweets became unprotected and consequently were accessible to the wider public without the user's knowledge' [1]. Since Twitter did not share the details of this bug or the structure of their backend infrastructure, it is hard to tell if this bug is the result of a flawed design like using changeable attributes such as email address to identify users or the result of a human factor like a sloppy implementation by one or multiple engineers.

2.2 Measures to avoid similar violations

In this section, I consider the previous violations and present two possible technical solutions that Twitter and other companies could take to avoid similar problems.

2.2.1 Separation of tweets settings and account settings

One of the most fundamental way of preventing violations mentioned in 2.1.2 is to separate the privacy settings of a tweet from the account settings of its owner. Even though the data storing structure of Twitter remains opaque to me, I believe it is fair to make the inference that there was a strong relation between user account settings (like registered email addresses) and privacy settings of tweets (like access permission), for if such a relation did not exist, then a scenario where changes made to user accounts would cascade to access of tweets should be impossible. So, if a separation of tweets access and account settings can be created, then no matter what kind of changes made to the user accounts, intentionally or inadvertently, the accessibility of tweets will remain unchanged.

Nevertheless, it could turn out to be extremely expensive to achieve this since if the accessibility of a tweet is strictly kept to the tweet itself, then whenever a user decides to change the accessibility of all the tweets, thousands or even tens of thousands of updates to the database are required.

2.2.2 GDPR aware bug reporting system

TIC has acknowledged that the delay of their report was due to a failure by a developing team, to follow a specific protocol [1]. In order to prevent such a violation from happening again, there are many approaches Twitter can take, like providing better GDPR training for their developers so that they may recognize potential data breaches or enlisting the help of GDPR professionals to examine every reported bug. However, my insight is that solutions like these are either ineffective or too expensive, a better solution is to introduce a GDPR aware bug reporting system to assist developers in finding data breaches.

It is common practice in the industry to use systems like JIRA [6] to track bugs and their repair status. If these systems are modified to be GDPR aware and able to first guide developers to recognize potential data breaches when they submit tickets for bugs, then automatically check if any privacy data may have been compromised based on the information included in the tickets, finally decide whether to notify the Controller and supervising authorities depending on the results of automatic check, then companies could sharply reduce the occurrence of violations to Article 33(1) with minimum costs. Even more, a data breach documentation feature could be integrated into such systems to help produce GDPR compliant documentations, by guiding legal teams to fill in the necessary information or even auto filling the information using what is included in the tickets.

3. Discussion

I believe this case has shown the power of GDPR in regard of protecting users' information, Twitter was charged a fine of 450,000 euros [2], I believe this amount is significant enough to warn all similar companies to change their practices. However, it is my opinion that this amount is still not enough when the damage caused by Twitter is considered. The data breach involved in this case was introduced as early as 2014 and more than 880,000 users live in EU alone were affected [1], but as I mentioned in this paper, Twitter was not charged for such a catastrophic leakage, it was only charged for not reporting on time and failure to present a detailed documentation. I do believe that even though Twitter has taken several measures like recovering privacy settings for users who were affected and putting up a notice in their help center to raise awareness of the problem [4], it should have been charged for insufficient technical or organizational support to protect users' privacy.

Also, the efficiency of the GDPR enforcement process is questionable to me. Although the Commission commenced their inspection in less than 15 days after they received the report from Twitter in January 2019, the final decision was passed on December 2020 [1], the whole process took nearly two years to finish. As much as I appreciate the thoroughness shown by the regulators in this case, I believe the procedure could be optimized and made more efficient.

References

[1] DPC case IN-19-1-1 *Twitter International Company v. Ireland Data Protection Commission.* [2020] (https://edpb.europa.eu/sites/edpb/files/decisions/final_decision_-_in-19-1-1_9.12.2020.pdf)

[2] https://www.siliconrepublic.com/enterprise/data-protection-commissioner-twitter-fine

[3] https://en.wikipedia.org/wiki/Twitter

[4] https://help.twitter.com/en/protected-tweets-android

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679#d1e6226-1-1

[6] https://en.wikipedia.org/wiki/Jira_(software)