What WhatsApp knows about your Contacts: A GDPR Case Study

Benjamin Kilimnik
Brown University

Abstract

On August 20th, 2021 the Data Protection Commission of Ireland concluded that WhatsApp unlawfully stored phone numbers of non-users via its "Contact Feature", thereby violating articles 5, 12, 13, and 14 of the GDPR. The commission levied a fine of 225,000,000 Euros on the company and imposed 8 corrective actions which WhatsApp is required complete within 3 months of the decision.

1 Background

The Data Protection Commission of Ireland investigated WhatsApp in response to 88 complaints submitted by the privacy authorities of Germany, the Netherlands, Austria, Spain, the UK, France, Finland, and Poland [1]. The complaints reported transparency issues regarding the data processing activities of WhatsApp and raised concerns over the possible sharing of personally identifiable information between WhatsApp and other Facebook-owned companies [2].

During the investigation, WhatsApp Ireland Limited - the company's headquarters in Europe - was deemed the **data controller** for EU users by the commission and was therefore responsible for determining the means and purposes of processing of user data. Users of WhatsApp located in the EU were deemed **data subjects** while unnamed "sub-processors" of WhatsApp were considered **data processor(s)** [1].

1.1 Technical Details of the Violation

One core finding of the investigation was that WhatsApp has a "Contact Feature" that allows the Service to access phone numbers stored on a user's device in order to identify which ones already use WhatsApp and which do not. The feature works as follows:

(i) When a user activates the "Contact Feature", WhatsApp accesses all mobile phone numbers on a user's device,

(ii) transfers those numbers to WhatsApp's servers,

(iii) generates cryptographic lossy hashes of the non-user phone numbers,

(iv) stores the hashes in a list ("the Non-User List"), linking each user who uploaded the non-user's number to the generated lossy hash of the non-user's number, and finally,

(v) deletes the underlying phone numbers of non-users [1].

By comparing the hashes in the Non-User List with the hashed numbers of newly joined users, WhatsApp can update the address book of existing users with the details of the new WhatsApp user. On this basis, the Data Protection Commission determined that the purpose of processing non-user contacts implied the identification of individuals. An important point here is the commission's ruling that the phone number of a non-user remains personal data even after hashing and hence falls under the jurisdiction and protection of the GDPR. In particular, the commission determined that the "table of lossy hashes together with the associated users' phone numbers as Non-User List constitutes personal data" [1].

2 GDPR Violations

2.1 Infringements

1. *Depriving data subjects of their right to information (GDPR §12-14):* During the investigation, it was determined that WhatsApp did not provide non-users with information that would have enabled them to make a fully informed decision about whether or not to become a user of the service [1]. Moreover, WhatsApp did not provide non-users with a method of giving consent to the collection of their phone numbers through the "Contact Feature." In particular, the investigators identified violations in the following articles of the GDPR:

   (a) §12(1), which requires that information be conveyed transparently, intelligibly and easily accessible (particularly for children),

(b) §13(1)(c) which requires that the data controller provide information about the "legal basis for the processing,"

(c) §13(1)(d), which requires the provision of information about the "legitimate interests being pursued," and

(d) §13(1)(e), which requires that specific information be provided about the categories of recipients of personal data.

The Data Protection Commission found WhatsApp's privacy policy to be lacking specifics about the purposes for which the company shares personal data with third-parties. WhatsApp also failed to elaborate on the specific "circumstances in which their [data subjects'] personal data [such as phone numbers] would be transferred and to whom such transfers are made" [1]. Finally, the investigation determined that the information WhatsApp provided to data subjects does not enable an understanding of what data is sent to which category of recipient, nor why such data transfers are done, resulting in a violation of §12(1).

2. *Violation of the Transparency Principle §5(1)(a):* The commission found that WhatsApp's privacy policy was lacking in detail and did not provide sufficient information to establish the "legal basis for the processing" of user data. In fact, it was discovered that WhatsApp provided only 41% of the prescribed information under GDPR to users and none at all to non-users of the service whose phone numbers were being stored on WhatsApp's servers in a hashed format [1]. This is particularly problematic given the sheer number of data subjects affected (estimated in the millions) and the length of time the infraction has been occurring for (at least since April 2018 and perhaps longer) [1].

3 Ruling

WhatsApp was found guilty of GDPR violations and fined 225,000,000 Euros [2]. In addition, the company was given 8 "Terms of Order" to bring the service into compliance with GDPR within a three month period. These actions broadly state that WhatsApp must increase the amount and quality of information available to users and non-users regarding the purposes of data processing, the data gathering process, as well as the recipients of personal data.

4 WhatsApp's Response

WhatsApp's objections to the ruling may be categorized into two general types:

Different interpretation of the GDPR: WhatsApp fundamentally disagrees with the Data Protection Commission's statement that information about the categories of personal data to be received by identified categories of recipients should be made available to users, claiming that such a requirement falls beyond the scope of the Transparency Guidelines set out in §13(1)(e). WhatsApp denies that any real damage has been done, claiming, for instance, that the "Contact Feature" was designed for the convenience and benefit of users and potential users.

Technical limitations: WhatsApp states that it lacks the ability to access the actual mobile phone number of non-users during processing and does not have the technical infrastructure nor desire to link a non-user's phone number to a natural person. So even if phone numbers are stored on WhatsApp's servers in some form, there is no practical way of linking them to the people they belong to.

No official press release appears to have been made by WhatsApp or Facebook discussing the ruling and its implications for their business. Perhaps that is still forthcoming. Regardless, there is certainly no shortage of media coverage on the subject, particularly because this case involves one of the largest GDPR fines to date.

5 Prevention

WhatsApp's greatest issue was a general lack of transparency in its privacy policy regarding the use and recipients of personal information, leading to a series of complaints by individuals and data protection authorities about the possible storing and sharing of personal data with other Facebook owned companies. To prevent complaints, the service could have provided more in-depth information about the company's data gathering practices to allow users to make an informed choice of whether to use the service. For example, WhatsApp could have stated in their privacy policy that phone numbers would be stored on their servers through use of the "Contact Feature" and could have developed a way for non-users to opt out.

On a technical level, WhatsApp could have designed their "Contact Feature" in a way that does not store non-user phone numbers (even in a hashed form) on WhatsApp's servers without their consent.

6 Discussion

Considering the sheer number of data subjects affected and the potential consequences of leaking vast stores of non-user phone numbers, the fine appears appropriate. This is compounded by vague descriptions and omissions in WhatsApp's privacy policy which might have persuaded a privacy conscious user not to enroll in the service.

WhatsApp states that based on the way their systems currently operate, it is "not technically feasible" to extract non-user phone numbers from the hashing process - that would require design and code changes in order to "log additional information" [1]. I have little doubt that this is true - nonetheless, it does not entirely remove the possibility of such identification taking place, which would be a large-scale privacy debacle.

References

[1] Helen Dixon. *Decision of the Data Protection Commission made pursuant to Section 111 of the Data Protection Act, 2018 and Articles 60 and 65 of the General Data Protection Regulation*. Data Protection Commission, 2021. https://dataprotection.ie/en/news-media/press-releases/data-protection-commission-announces-decision-whatsapp-inquiry.

[2] Natasha Lomas. Whatsapp faces $267m fine for breaching europe's gdpr, 2021. https://techcrunch.com/2021/09/02/whatsapp-faces-267m-fine-for-breaching-europes-gdpr/.