"Data Protection Can Sometimes Be a Nuisance" A Notification Study on Data Sharing Practices in City Apps

Jan Drescher
jan.drescher@tu-braunschweig.de
Technische Universität Braunschweig
Braunschweig, Germany

Jakob Moser
moser@cl.uni-heidelberg.de
Ruprecht-Karls-Universität
Heidelberg, Germany

Nicolas Strangmann
strangmann@stud.uni-frankfurt.de
Goethe-Universität
Frankfurt, Germany

Jonas Spinner
j.spinner@thphys.uni-heidelberg.de
Ruprecht-Karls-Universität
Heidelberg, Germany

Dominik Herrmann
dominik.herrmann@uni-bamberg.de
Otto-Friedrich-Universität
Bamberg, Germany

Melanie Volkamer
melanie.volkamer@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

Abstract

Despite the strict requirements regarding the justification of data sharing imposed by the General Data Protection Regulation (GDPR), many mobile apps, even those provided by European states, share user data with third parties without justification or consent. To assess data sharing of city apps, we analyzed 138 apps from German cities for non-compliance with the GDPR. We found that 70 of these apps contacted third-party services outside the European Union without user consent, making them potentially non-compliant with current European privacy regulations. To investigate what information helps app vendors to remediate the issue, we sent three types of notifications to potentially non-compliant vendors: A generic one, one with detailed technical guidance to achieve compliance, and one with a detailed legal explanation. We observed a response rate of 37% and fix rates of approximately 17% for the two groups that received detailed notifications. Thereby, we found that both technical guidance and legal explanations significantly increase the number of fixed apps, compared to just sending generic notifications. While the response rate was higher than during comparable studies, we observed high distrust in our messages, similar to related work. Surprisingly, we found that many of the app vendors who promised to remediate the issue, did not do so successfully, while others silently patched their app.

CCS Concepts

• **Social and professional topics** → **Privacy policies**; • **Security and privacy** → **Human and societal aspects of security and privacy**.

Keywords

data protection; compliance; notification; mobile apps

Mensch und Computer 2024 – Workshopband, Gesellschaft für Informatik e.V., 01.-04. September 2024, Karlsruhe, Germany

https://doi.org/10.18420/muc2024-mci-ws17-159

1 Introduction

Researchers have studied the compliance of apps and websites with data protection laws [9, 10] and best practices in security [3, 8, 17]. Among other things, it has been well documented that many apps track their users without explicit consent [5, 10, 11].

A recent study of governmental websites and mobile apps on the state-wide level revealed that up to 37% share data with third parties without consent [16]. The average citizen in their daily life, however, is much more likely to interact with the apps provided by their municipal government than apps provided by the state [15]. This is the reason why, we evaluate in our first research question: *Which percentage of city apps from Germany respects explicit consent before contacting third-party hosts?* To answer this question, we analyzed 138 apps from German cities and found that 70 of these apps contacted third-party services without consent. As public authorities cannot justify data processing as a legitimate interest per Article 6 (1) lit. f General Data Protection Regulation (GDPR) [14], these apps are likely non-compliant with current European privacy regulations.

When researchers contacted service providers to inform them about non-compliance with data protection regulations or security issues, they generally observed low response and fix rates [6, 10, 11]. A new field of research that aims to answer the question of how to effectively notify service providers and trigger improvements. For example, Maas et al. [9] examined the impact of the notification medium and sender on the remediation rate for non-compliant Google Analytics configurations. In line with related work, we ask which information helps or motivates the app providers to remediate the non-compliant data sharing. We hypothesize that information on the legal status of the data sharing practices motivates the receiver to fix the app. A specific technical guide that describes how to configure the app correctly might support the receiver in fixing the app.

This leads to our second research question: *What impact does a technically- or legally-focused mail have on the response?* Similar to related work [9, 10], we sent either a generic notification, technical guidance, or a legal explanation to the 70 affected app vendors and analyzed their responses. With the technical guidance we aim to support developers in improving the app. With the legal explanation, on the other hand, we aim to highlight the compliance

violation and create relevance. We report the quantitative effectiveness of the notification and reminder campaign, using the results of the notification campaign by Nguyen et al. [10] as a baseline.

Leveraging insights gained from the interaction with the app providers, we ultimately address the following question: *How can the responses in this study be used to optimize notifications about data protection violations in future studies?* We analyzed the responses using an open-coding-based qualitative approach to gain insight into the processes on the app vendors' side and tailor notifications better to the vendors' needs.

The paper is structured as follows: First, we review related work in Section 2. In Section 3, we describe the methods we used to collect apps, scan them for non-compliance, write the notifications and evaluate the responses. The results of our study are described in Section 4 and discussed in Section 5. We conclude the paper in Section 6.

2 Related Work

Nguyen et al. [10] conducted a comprehensive analysis of 86,163 Android apps, focusing on unauthorized data sharing with third-party providers. They employed a pipeline using mitmproxy to intercept network traffic, and search for personally identifiable information. No user interaction takes place as part of this pipeline, that we reuse. Thus, all the requests collected by the pipeline are sent without explicit user consent. Their data set consists of 16,163 "high-profile" apps, i.e., the top apps across different categories, and 57,299 "long-tail" apps. Up to 39.8% of apps sent personal data to third-party domains. In contrast, we analyze a smaller set of mobile applications. Because of their low monthly download numbers, our city apps would be considered as "long-tail" apps.

In a related vein, Samarasinghe et al. [16] investigated government websites and apps from 71 different countries, finding that 37% of the apps contained third-party trackers. In comparison, we analyze a smaller but more specific set of Android apps. We focus on apps from German cities, which have a lower profile than the top-level governmental apps that Samarasinghe et al. examined.

In a similar methodological approach to Nguyen et al., Kollnig et al. [6] manually examined 1,297 UK Play Store apps, monitoring network traffic to identify unauthorized contacts with known tracker hosts. They found that 71.3% of apps contacted known tracker hosts without consent. In addition, Kollnig et al. examined the APIs and guides that third parties provide to app developers to correctly obtain user consent. They found that very few providers implement consent in their SDKs. While compliance guides generally exist, they are hard to find, read, and implement. Similar to Kollnig et al., we also examine the documentation for various SDKs, to provide the technical guides for our notification study.

Building on these insights, both Nguyen at al. [11] and Koch et al. [5] proposed an extended approach to automatically interact with consent dialogues of mobile apps and observe the effect of different choices. We, in contrast, do not interact further with the apps after opening them. Thus, the observable behavior is limited, compared with the aforementioned studies. For example, we do not examine if an app sends requests to third parties after the user explicitly disallowed data-sharing with third parties.

The effectiveness of notification methods in prompting remediation has been explored in various contexts. Stock et al. [17] notified the owners of websites with security issues, examining if the tone of the mail, the sender name, or the inclusion of tracking resources influence the response or fix rate. They identified two roadblocks for remediation: Many mails did not reach the receiver because they bounced or were filtered as spam. The other roadblock were human factors: Less than 26% of domain operators who read the notification mail opened the attached report.

Maass et al. [9] conducted a similar study, examining the impact of framing and notification channel. They contacted the owners of websites that incorporated Google Analytics in a way that is non-compliant with the GDPR. They varied the channel between mail and letter and the sender between a private individual, a university's computer science group, and a university's law group. The framing was varied between a focus on the privacy violation, the GDPR violation, and possibly resulting fines. The medium with the largest impact on remediation was the letter. The most effective sender was the university law group, with the most effective framing being the mention of the possible fines.

We therefore hypothesize that the content of the notification impacts the remediation rate and send three different notifications. Our generic mail is similar to the generic mail by Stock et al. Our legal mail highlighting the rules of GDPR is similar to the message by Maass et al. that highlighted the GDPR violation. Maass et al. added a link to a self-check tool to their notifications. The tool's website also contains a guide that outlines the correct configuration of Google Analytics. We hypothesize that such guidance helps the app provider to fix the app. We test this hypothesis by sending technical guides on correct SDK configuration in the third mail.

Nguyen et al. [10] also notified the developers of the apps which shared user information without explicit consent. The notification was sent via mail to the Play Store contact mail address. They asked the developers if they were aware of the non-conformity and if the developers were planning to take action. The developers of 15.7% of the of apps visited the attached report URL. 3.8% replied to the notification and answered the questions. Similar to Nguyen et al., we also notify the developers of Android apps and obtain the contact mail address from the Play Store.

3 Methodology

In this section, we describe the execution of our notification study. First, we cover the selection and analysis of the city apps. The second part covers design and evaluation methodology of the notification study.

3.1 App Analysis

In this section, we describe the selection and subsequent scanning of apps, resulting in the apps comprising the sample for our notification study. The process of notifying the developers of these apps is described in Section 3.2.

3.1.1 App Selection. Our selection process for apps provided by cities in Germany is depicted in Figure 1: We manually searched the Google Play Store using the query "*<City Name>* App" for the 711 German cities with populations exceeding 20,000 inhabitants, as listed in [19] on 2022-09-20. Many of the apps we found provide

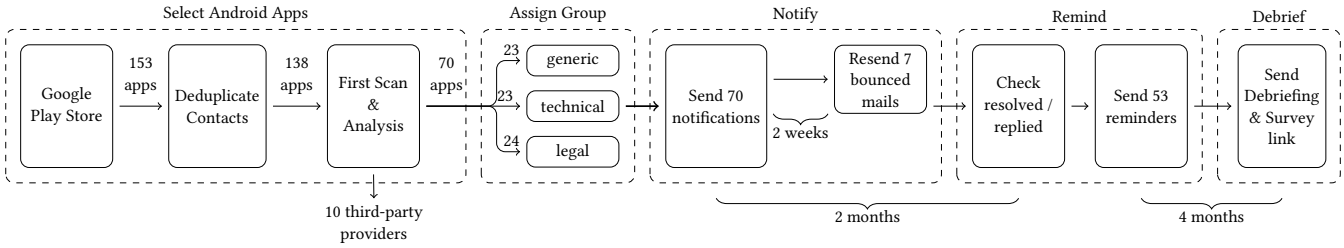

Figure 1: Notification study design

tourism or municipal services, or solicit feedback from citizens. We deemed the required affiliation to the city to be fulfilled if either the privacy policy or the linked website's imprint cites the city as responsible. For two cities we found multiple apps for the same city, in which case we selected the app with the highest number of downloads. Additionally, a handful of companies managing apps for multiple cities were identified based on the mail addresses provided in the Google Play Store. To prevent potential bias in the results of this study due to multiple notifications reaching the same entity, we selected only the app connected to the city with the most inhabitants for each company. This means, we removed another 15 apps from the dataset. Based on the provided mail address, we distinguished apps where we directly contact city representatives from apps where we instead contact a third-party agency. This information is used to equally distribute these two classes of apps over the three groups receiving different notifications as discussed in Section 3.2. These selection criteria resulted in a total of 138 city apps to be scanned.

3.1.2 Scanning. We used the analysis pipeline by Nguyen et al. [10] to collect the HTTPS traffic generated by every app directly after starting it, without human interaction. We scraped the contact information and downloaded the app from the Google Play Store. As part of the pipeline, each app was automatically installed and run on a rooted Pixel 6a using Frida[1] and objection[2]. We intercepted traffic using mitmproxy[3], for which it was necessary to disable certificate pinning. We ultimately decided to use the traffic dumps generated by mitmproxy only to extract a list of hosts the app sent requests to, because every HTTP(S) request necessarily contains the client's IP address, which constitutes identifiable information[4] [4]. Because no user interaction takes place when the app is run, no explicit consent to such a transmission was ever granted.

After the initial scan, we filtered the list of contacted hosts for those associated with third parties, whose main offices are located outside the European Union (to which a transfer of personal information without giving consent was deemed a GDPR compliance violation at the time of our scan [12]). We identified ten such third-party services and their corresponding domains in transmitted HTTP(S) requests: *Facebook, Pushwoosh, Google Firebase Crashlytics, Google Firebase, OneSignal, Visual Studio App Center, Google Analytics, Google Fonts, Google Maps, YouTube.* 70 out of the 138 scanned apps sent HTTP(S) requests to one of these third parties outside the European Union without user consent.

The proxy we used intercepts all traffic, including background traffic by the operating system Android. This might cause false positives, for example, when an Android background service sent a request to a Google host while an app was open. To avoid false positives, we ran the pipeline several times and only considered third-party services contacted in all runs. To prevent misattribution of HTTP requests, we install and run one app at a time. We consider an app as fixed if no requests to tracker domains are sent before any user interaction takes place. Thereby, we also treat apps that were removed from the Play Store as fixed, because no requests can be observed. We regularly re-scanned the apps once per week to detect fixes.

3.2 Notification Study Design

We conducted a between-subject study with three conditions. The design of this notification study, as depicted in Figure 1, is explained in this section.

3.2.1 Group Assignment and Notification Design. To investigate what style of mail notification is most effective, we split the 70 apps into three groups containing 23, 23, and 24 apps. The groups received a generic, technically- or legally-focused notification mail, respectively. Figure 2 displays the composition of the different notification messages.

The first group received a generic notification mail similar to the plain notification sent by Stock et al. [17]. The second group received technical guides on how to disable sending requests to third parties. To create these guides, we searched each third party's documentation for guides on obtaining user consent. Specifically, we looked for configuration options to disable third-party communication until consent is obtained. For Android SDKs, this usually came down to disabling automatic initialization. If the app displayed a website using an Android WebView and the website embedded third party resources, we looked for guides on configuring the website to be GDPR-compliant. If such configuration options were provided, we prepared a short snippet and appended it to the technical mail. The snippet contains step-by-step instructions on how to disable the SDK and later re-enable it after consent was granted. We assume that this configuration is practicable even for SDKs that are relevant for the functionality of the app, such as push messaging. If no sufficient documentation or APIs existed, we recommended removing the SDK.

[1] https://frida.re/
[2] https://github.com/sensepost/objection
[3] https://mitmproxy.org/
[4] While this interpretation may be subject to differing opinions, compare a recent ruling [13] by the European Court of Justice regarding Vehicle Identification Numbers.

The third group received information on the legal status of the data transmission. The mail highlighted under which circumstances the GDPR permits data sharing. It outlines that no explicit consent was given and the other conditions for lawful processing are also not met (especially mentioning that legitimate interest does not apply for public bodies). The mail further mentions that according to the "Schrems II" judgment, the transfer of personal data (e.g., IP addresses) outside the European Union (i.e., to the U.S.) was ruled unlawful.

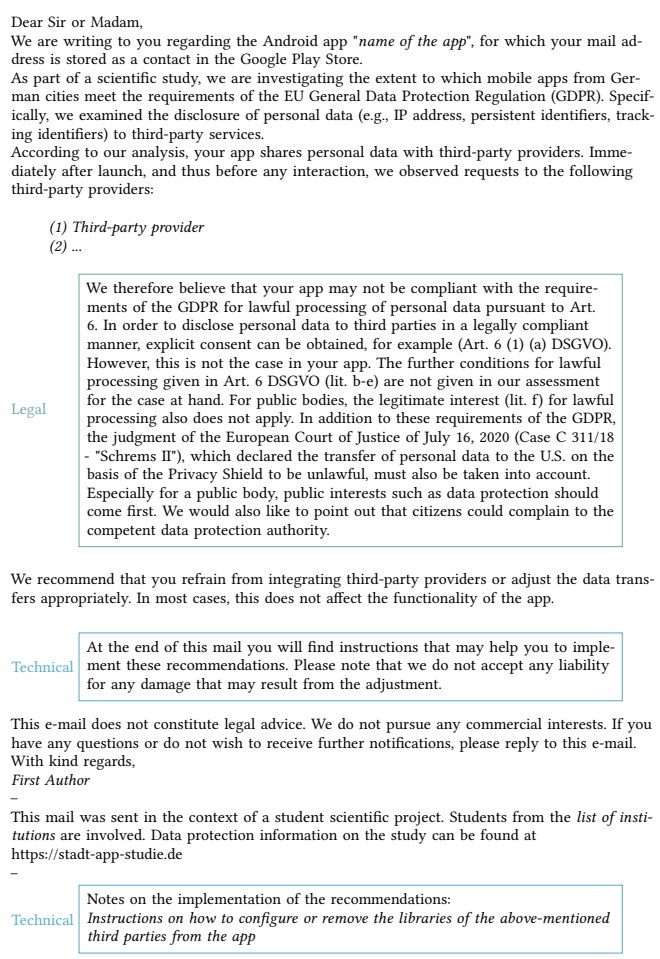

Figure 2: Composition of the notification messages

We sent the mails using the address *kontakt@stadt-app-studie.de* (English: *contact@city-app-study.de*). The domain was chosen to appear neutral and unaffiliated with any particular research institution, yet professional. We introduced ourselves as student researchers.

The first batch of 70 notifications was sent on 2023-03-23. Due to technical problems, e.g. mails being rejected by spam filters, seven of these notifications were not accepted by the receiving mail server. We waited for two weeks (2023-04-07) and sent the mails again. This time, all were accepted. After two months (2023-05-16 and 2023-05-31 for the second batch), we sent reminders to all contacts who neither replied to our previous mail nor improved their app to prevent data sharing without user consent. We replied to every response we received during the study with a generic acknowledgment stating we would only reply again after the study had ended (i.e., for the debriefing). We replied individually to questions regarding our study or GDPR-compliant app configurations.

3.2.2 Debriefing and Survey. We sent a debriefing mail, including a short survey, as done in previous work [9, 17]. This was done on 2023-09-07, six months after the start of the study and four months after the reminder, together with a final scan. The debriefing contained a brief explanation of the study, as well as both the technical and legal information texts. The attached survey was designed to gain qualitative insights about our mail's usefulness and trustworthiness to improve future notifications. The participants were asked whether they read the received mails, why they did or did not fix their apps and what parts of the mail made it seem more or less trustworthy. The full survey can be found in Appendix A. We publish the original mail templates in German, as well was the technical guides, as part of the artifacts of this paper[5].

3.3 Ethical Considerations and Data Protection

None of the involved organizations required IRB approval for this kind of study. We conducted this study openly and identified ourselves as researchers in the first notification mail. However, we withheld the technical or legal information that we provided to some groups from the other groups. We limited the impact of withholding that information by releasing all relevant information during debriefing.

Because cities as public bodies are unable to be sued under GDPR, we believe sending a legally-focused notification should not cause them harm, unlike with private companies where a legally-focused mail might cause harm (e.g., in the form of attorney fees, as apparently happened in response to the Princeton-Radboud study [18]).

We publish the data set of the dynamic app analysis as part of the artifacts[5]. We only publish this data in pseudonymous form to protect the privacy of the individuals and organizations involved. We published a privacy policy with a short description of the study on the project website, `stadt-app-studie.de`.

3.4 Qualitative Analysis of Responses

To analyze the responses to our notifications and identify differences between the three groups, we categorized the mail responses using an open-coding-alike approach [2]. During the analysis, we treated all mails received from the same entity as one document. Two authors analyzed the mails independently to identify codes that described how the receivers of our mails perceived the notification, the tone of their reaction, and what future actions they planned. A common code book containing the following six codes was derived from a discussion between these two authors:

Thankful Explicit expression of gratitude for the notification.
Review Acknowledgment of the notification with an investigation on the issue ongoing.
Fix Promised Claim that the issue was already fixed or will be fixed in the near future.

[5] https://zenodo.org/records/13140707

Policy Reference to updates in the privacy policy or forwarded to the city's data protection officer.

Rejected Claim that the issue we were notifying on is legally not a problem.

Query The response contains a question.

Using this code book, two authors independently tagged the mails. The inter-author agreement, quantified through Krippendorff's alpha [7], was found to be $\alpha = 0.85$. Disagreements were resolved by discussion.

4 Results

Time passed	Generic	Legal	Technical
1 week	0	2	1
1 month	0	2	1
3 months	0	2	2
6 months	0	4	4

Table 1: Fixed apps

In this section, we present the response and fix rates of our notification study as well as an initial analysis of the replies.

During our initial scan, 70 out of 138 apps of German cities, i.e., 51%, sent HTTP(S) requests to third parties outside the European Union. This result is in line with the work of Samarasinghe et al. [16] who discovered trackers in 37% of government apps. We also searched for the transmission of user data, advertising IDs, or device identifiers using the pipeline by Nguyen et al. [10]. None of the 70 apps that contacted third parties displayed this kind of behavior. Of the top ten ad-domains which received user data as observed by Nguyen et al. [11] only `facebook.com` was present in the traffic that we observed.

Table 1 lists the number of remediated apps over the time of our notification study. In the group that received the generic notification, none of the 23 recipients fixed the app. Of the 23 recipients of the mail with technical guidance, we observed four apps no longer sending third-party requests. Of the 24 recipients that received the legal mail, four apps no longer sent removed to third parties. The improvements were implemented in the weeks after our notifications and reminders. We noticed that 3 of these 8 cities opted to remove the app from the Play Store completely, rather than update a probably abandoned app. Two of these cities belonged to the legal notification group, one received the technical notification. We observed a fix rate of approximately 17% for both the technical and legal notification.

We applied Barnard's exact test [1] to determine if the observed difference in remediation is statistically significant. Our null hypothesis is "*The odds of the developers fixing the app are the same for the technical and the generic group*". Our alternative hypothesis is "*The odds of the developers fixing the app are larger for the technical group than for the generic group*". Using a significance level of $p < .05$ ($p = .0209$), we can reject the null hypothesis in the one-sided test in favor of the alternative hypothesis. Analogous, we can also reject the null hypothesis in favor of the alternative hypothesis for the one-sided test of the difference in remediation between the legal and the generic group with $p < .05$ ($p = .0231$). Thus, we observed a statistically significant impact of the technical guidance or legal information.

Several city apps were built as simple web views displaying a mobile-optimized version of the city website. Interestingly enough, one city had already removed Google Analytics from their main website, but still used it in the version of the website displayed to app users. While we can only speculate about the reasons for this, we find it very likely that this was simply an oversight. We provide the pseudonymized results of our app analysis as part of the artifacts.

4.1 Responses

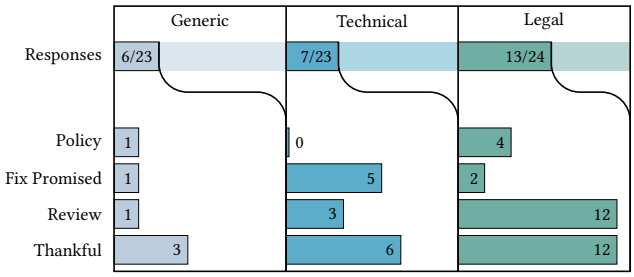

Figure 3: Distribution of the four most common codes in the responses to the three different types of emails.

Out of the 70 contacts notified during our study, 26 replied to at least one of our mails. On four occasions, multiple emails were received from the same entity. The response rate of 37% is still considerably higher than the rates observed during similar studies. Nguyen et al. [11] observed a response rate of 4% during their notification campaign that targeted popular and long-tail apps from the Play Store. Figure 3 shows how the four most common codes (*Thankful, Fix Promised, Review* and *Policy*) are distributed among the different email categories. As the two codes *Rejected* and *Query* were only rarely observed, and no significant difference between the notification types was observed, a closer inspection is omitted here:

A large proportion of the respondents expressed their gratitude: Specifically, 6 out of 7 receiving the technical and 12 out of 13 receiving the legal mail, while only 3 out of 6 respondents for the generic mail.

While 5 out of 7 responses to technical emails claimed to have fixed the issue already, only 1 out of 6 and 2 out of 13 of the responses to generic and legal mails, respectively, contained such claims. Interestingly, only one app where a fix was claimed during mail communication was actually fixed. This might be due to miscommunication between customer support and development teams, as we have observed changes in the apps after the notification, however, they were apparently insufficient to disable the third-party services.

With 12 out of 13 responses, almost every legal correspondent said they are reviewing or have reviewed the issue. At the same time, only 1 out of 6 responses to the generic and 3 out of 7 responses to the technical mail said this. Additionally, 5 responses directed us

toward the data protection officer or changes in the privacy policy, with 4 of them responding to a legal email from us. Although the number of responses is small and does not support quantitative conclusions, these findings suggest that the type of email, whether legal or technical, can affect how recipients approach the issue.

Furthermore, we also received an interesting response from one city, which noted the lack of a proper mail signature in our mail's footer. The sender of the reply expected a valid mail from a research group to contain a footer with an address or a logo.

4.2 Survey

After having sent 70 debriefings, we received 5 completed surveys, a response rate similar to Stock et al. [17]. While five responses forbid any form of quantitative analysis, some of the responses provide interesting insights.

Several recipients expressed skepticism towards emails from unknown senders: "An email from a sender we don't know indicating that we probably have a privacy problem in our app should be treated with caution these days". Further, the seriousness of our message was questioned: "When I received the first email, I immediately thought of advertising and that something really great was being sold here to make the app better".

Some recipients doubted the validity of our message, telling us that our allegations about the compliance violation were "fundamentally incorrect and therefore false and may have caused confusion among many addressees". The recipients speculated that "it seems like some students here are upset because they have found a paragraph only relevant to a small number of groups". Furthermore, we were told that "nowadays, every serious software provider should already be aware of [data protection requirements], and no external clarification is really needed."

We received positive feedback, especially from the group that got the technical email: "Concrete advice was given that made implementation easy." and "We would also be interested in having our app checked by independent third parties in the future in order to continue to be exemplary in terms of data protection law and to exclude deficiencies". However, not everyone shares this opinion: "Data protection can sometimes be a nuisance and hinder our work. Perhaps we could be left alone?"

5 Discussion and Limitations

In the following, we revisit our initial questions in light of the results.

To what degree do city apps from Germany respect explicit consent before contacting third-party hosts? We found that 51% of the analyzed German city apps were likely non-compliant with the GDPR. While the situation is far from perfect, we also found that city apps demonstrate a considerably higher fix rate compared to an average "long-tail" app during the notification study by Nguyen et al. [10]. This trend reinforces the assumption that city app developers are more attuned to the importance of addressing privacy and data sharing concerns, despite the ongoing challenges that remain in this domain.

What impact does a technically- or legally-focused mail have on the response? Both technically-focused and legally-focused notifications seem to elicit more responses than generic notifications. Sending notifications with a technical focus has a slight tendency to lead to more promised fixes, whereas sending notifications with a legal focus has a slight tendency to lead to more replies stating the matter was under review. Furthermore, sending a legally-focused notification also led to more references to privacy policies, which we did not observe when sending technical notifications. Both types of focused notifications lead to a significantly higher fix-rate compared to the generic notification.

How can the responses in this study be used to optimize notifications about data protection violations in future studies? Some of the responses indicated that not all recipients deemed the mail trustworthy, for reasons that might not be immediately obvious to IT security professionals, namely a lack of mail signature (i.e. information about the sender in human-readable form in the footer of a mail) or the unprompted sending of mails (which is, of course, a property of notifications). This is in line with previous research [9] which found that distrust of unsolicited messages is strong and small factors like letterheads can increase trust.

Another problem is that functionality deemed relevant by the app vendors (e.g. push notifications, crash reporting etc.) often comes bundled with other functionality and tracking capabilities in SDKs. Evaluating the survey supports this interpretation: The respondents often mentioned they were already aware of data protection and did not need any notifications, however, they deemed it impractical to achieve full data protection compliance without negatively impacting functionality. Technical guides that outline how to disable the SDK until user consent is granted can therefore increase the remediation rate.

Furthermore, the low number of contacts that kept their promise to fix the app indicates issues implementing said promises. It is possible that the developers of the app made some changes that were, however, not sufficient to prevent unauthorized requests to third parties. Providing a self-check tool similar to the one provided by Maass et al. [9] could enable developers to evaluate their improvements. A subsequent study could, for instance, supply a guide on using the Android Studio network inspector to intercept the HTTP requests of the app.

Limitations. As we focused on apps from German cities, we only notified a rather small group of app providers and received only a limited number of responses. While our results allow for statistically significant conclusions, the small study size reduces the meaningfulness of our results. In addition, different app providers might react differently to our mails.

While IP addresses were considered personal data at the time of this study, recent rulings of the Court of Justice of the European Union [13] indicate, that this condition might change in the future. Furthermore, since the adequacy decision of the European Commission regarding the EU–US Data Privacy Framework on 2023-07-10, transfer of personal data to third parties in the United States of America is again permitted.

6 Conclusion

As part of our study we analyzed the web traffic of 138 apps provided by German cities. Our study revealed a concerning level of data protection inadequacies in city apps, with half of them communicating with third parties outside the EU without the user's consent. While this indicates a substantial gap in adherence to data protection standards, the fix and response rates, that we observed during our subsequent notification study, surpassed those described in previous work for average Android apps [10]. This suggests a greater awareness and sensitivity toward compliance and data protection issues among city app developers. Still, we observed a lack of trust into our notifications which is in line with previous studies.

Our findings indicate that the fix rate can be increased by including technical guidance or legal information in notifications. Based on these insights, we advocate to send more detailed notifications in similar studies. Since many app providers promised to remediate the issue but did not adjust the app sufficiently, we theorize that a self-check tool would have supported the providers to verify their changes.

Moreover, we encourage SDK vendors to prioritize the development of easy-to-deploy solutions to request consent from users before transmitting data as well as better documentation of GDPR-compliant SDK configurations. Furthermore, app developers should try to implement features like push notifications on Android using Google Play Services only, and reconsider how much tracking is actually necessary for purposes like crash reporting.

Acknowledgments

We would like to thank the Studienstiftung des deutschen Volkes which brought this research team together and supported and funded this research through their academic education and research program.

This work was supported by funding from the project "Engineering Secure Systems" of the Helmholtz Association (HGF) topic 46.23.01 Methods for Engineering Secure Systems and by KASTEL Security Research Lab.

References

[1] GA Barnard. 1947. Significance tests for 2× 2 tables. *Biometrika* 34, 1/2 (1947), 123–138.

[2] Juliet M. Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology* 13, 1 (1990), 3–21.

[3] Zakir Durumeric, Frank Li, James Kasten, Johanna Amann, Jethro Beekman, Mathias Payer, Nicolas Weaver, David Adrian, Vern Paxson, Michael Bailey, and J. Alex Halderman. 2014. The Matter of Heartbleed. In *Proceedings of the 2014 Conference on Internet Measurement Conference* (Vancouver, BC, Canada) *(IMC '14)*. Association for Computing Machinery, New York, NY, USA, 475–488. https://doi.org/10.1145/2663716.2663755

[4] European Commission. 2022. *What is personal data?* European Commission. https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_en

[5] Simon Koch, Benjamin Altpeter, and Martin Johns. 2023. The OK Is Not Enough: A Large Scale Study of Consent Dialogs in Smartphone Applications. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 5467–5484. https://www.usenix.org/conference/usenixsecurity23/presentation/koch

[6] Konrad Kollnig, Pierre Dewitte, Max Van Kleek, Ge Wang, Daniel Omeiza, Helena Webb, and Nigel Shadbolt. 2021. A Fait Accompli? An Empirical Study into the Absence of Consent to Third-Party Tracking in Android Apps. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, Online, 181–196. https://www.usenix.org/conference/soups2021/presentation/kollnig

[7] Klaus Krippendorff. 1970. Estimating the Reliability, Systematic Error and Random Error of Interval Data. *Educational and Psychological Measurement* 30 (1970), 61 – 70. https://api.semanticscholar.org/CorpusID:144036366

[8] Frank Li, Grant Ho, Eric Kuan, Yuan Niu, Lucas Ballard, Kurt Thomas, Elie Bursztein, and Vern Paxson. 2016. Remedying Web Hijacking: Notification Effectiveness and Webmaster Comprehension. In *Proceedings of the 25th International Conference on World Wide Web* (Montréal, Québec, Canada) *(WWW '16)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1009–1019. https://doi.org/10.1145/2872427.2883039

[9] Max Maass, Alina Stöver, Henning Pridöhl, Sebastian Bretthauer, Dominik Herrmann, Matthias Hollick, and Indra Spiecker. 2021. Effective Notification Campaigns on the Web: A Matter of Trust, Framing, and Support. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Online, 2489–2506. https://www.usenix.org/conference/usenixsecurity21/presentation/maass

[10] Trung Tin Nguyen, Michael Backes, Ninja Marnau, and Ben Stock. 2021. Share First, Ask Later (or Never?) Studying Violations of GDPR's Explicit Consent in Android Apps. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Online, 3667–3684. https://www.usenix.org/conference/usenixsecurity21/presentation/nguyen

[11] Trung Tin Nguyen, Michael Backes, and Ben Stock. 2022. Freely Given Consent? Studying Consent Notice of Third-Party Tracking and Its Violations of GDPR in Android Apps. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security* (Los Angeles, CA, USA) *(CCS '22)*. Association for Computing Machinery, New York, NY, USA, 2369–2383. https://doi.org/10.1145/3548606.3560564

[12] Court of Justice of the European Union. 2020. Judgment of the Court of Justice of the European Union in Case C-311/18 - Data Protection Commissioner v Facebook Ireland Ltd and Maximillian Schrems. https://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=11388612

[13] Court of Justice of the European Union. 2023. Judgment of the Court of Justice of the European Union in Case C-319/22. https://curia.europa.eu/juris/document/document.jsf?text=&docid=279492&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=568188

[14] European Parliament. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://data.europa.eu/eli/reg/2016/679/2016-05-04/eng Legislative Body: OP_DATPRO.

[15] Hananel Rosenberg and Azi Lev-On. 2021. Mobile applications in local government. *Electronic Government an International Journal* 17 (01 2021), 1. https://doi.org/10.1504/EG.2021.10034454

[16] Nayanamana Samarasinghe, Aashish Adhikari, Mohammad Mannan, and Amr Youssef. 2022. Et tu, Brute? Privacy Analysis of Government Websites and Mobile Apps. In *Proceedings of the ACM Web Conference 2022* (Virtual Event, Lyon, France) *(WWW '22)*. Association for Computing Machinery, New York, NY, USA, 564–575. https://doi.org/10.1145/3485447.3512223

[17] Ben Stock, Giancarlo Pellegrino, Frank Li, Michael Backes, and Christian Rossow. 2018. Didn't You Hear Me? - Towards More Successful Web Vulnerability Notifications. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018*. The Internet Society, San Diego, California, USA. https://www.ndss-symposium.org/wp-content/uploads/2018/02/ndss2018_01B-1_Stock_paper.pdf

[18] Ross Teixeira, Jonathan Mayer, and Gunes Acar. 2021. *Princeton-Radboud Study on Privacy Law Implementation.* Princeton University and Radboud University. https://privacystudy.cs.princeton.edu/?update

[19] Wikipedia, The Free Encyclopedia" 2022. *Liste der Groß- und Mittelstädte in Deutschland.* Wikipedia, The Free Encyclopedia". https://de.wikipedia.org/wiki/Liste_der_Gro%C3%9F-_und_Mittelst%C3%A4dte_in_Deutschland

A Survey

The survey contained the following questions and answer options (all translated from the original German version into English).

Study on data protection in local authority apps

Dear participant,

thank you for your willingness to take part in this survey.

By answering the following questions, you are supporting a research project by students at *list of institutions*.

The following survey form is provided by Otto-Friedrich-Universität Bamberg, which is accompanying the study. Data protection information on the study can be found at https://stadt-app-studie.de

Responses are anonymous. An asterisk (*) indicates mandatory questions.

Have you received and read the first mail?

At the end of March / beginning of April, we sent you a mail with the sender 'kontakt@stadt-app-studie.de' regarding a data protection problem in your Android app. Did you receive and read it?

The mail had something like the following text: *link* (text only appears after a few seconds; we recommend right-clicking to open the link in a new tab so as not to close the survey).

- Yes
- No

Have you received and read the second mail?

We may have sent you a second mail, also with the sender 'kontakt@stadt-app-studie.de', in mid-May / end of May regarding the same problem in your Android app. Have you received and read it?

- Yes
- No

The mail seemed trustworthy.

- Agree wholeheartedly
- Agree
- Neutral
- Do not agree
- Do not agree at all

Why did the mail have this effect?

Please briefly describe which aspects of the mail led you to assess it as trustworthy or untrustworthy.

[Free text answer]

Based on the mail, I was able to understand the problem of sharing data with external providers.

- Agree wholeheartedly
- Agree
- Neutral
- Do not agree
- Do not agree at all

Why?

Please briefly describe which aspects of the mail helped you to understand the problem or why you were unable to understand the problem on the basis of the mail.

[Free text answer]

The mail seemed urgent.

- Agree wholeheartedly
- Agree
- Neutral
- Do not agree
- Do not agree at all

Why did the mail have this effect?

Please briefly describe which aspects of the mail led you to rate it as urgent or not urgent.

[Free text answer]

Have you made any changes to your app as a result of the mail?

- Yes
- No

Why?

Please briefly describe the reasons why you have or have not made any changes to your app.

[Free text answer]

If you have made changes to your app, when?

- After receiving the first mail
- After receiving the second mail

Why?

Please briefly describe the reasons why you made the changes to your app at the time indicated, if applicable.

[Free text answer]

If you have made changes to your app, how?

- I made the changes myself.
- I made the changes with help.
- I have forwarded the task to a colleague within the company.
- I have instructed the responsible external service provider to make the changes.
- I have commissioned a new external service provider to make the changes.

If you have not made any changes, why not?

- Problem was not known
- Solution to the problem was unclear
- Lack of time
- Problem was not a priority
- Notification did not seem urgent
- Not considered a problem

Is there anything else you would like to tell us?

[Free text answer]