https://doi.org/10.2352/ISSN.2470-1173.2017.6.MOBMU-302

Privacy Issues in Mobile Health Applications - Assessment of Current Android Health Apps

Anett Hoppe [1]***, Jenny Knackmuß***[2]***, Maik Morgenstern*** [1]***, Reiner Creutzburg*** [2]
[1] ***AV-Test GmbH; Klewitzstr. 7, 39112 Magdeburg, Germany***
[2] ***Technische Hochschule Brandenburg - Brandenburg University of Applied Sciences, Department of Informatics and Media, P.O.Box 2132, D-14737 Brandenburg an der Havel, Germany***

Abstract

Mobile phones are constant companions in modern life. More and more users rely on an increasing variety of mobile applications for everyday tasks – an app offers distraction during a long wait at the doctor's, reminds to take an often forgotten medication or monitors current fitness values. While enabling a variety of tasks, every single app has potential access to a multitude of user information. Mobile phones contain an astonishing diversity of personal facts from contacts, call data, calendars to messaging contents or intimate health data. Despite the potential risks, users are reportedly negligent when it comes to the control of apps' access permissions and tend to grant wide access rights without further scrutiny.

Does this negligence cause personal information to be leaked to potentially malevolent actors? The presented assessment focuses on the privacy behavior of applications with a scope in user health and well-being, such as the above-mentioned pill reminder. These apps do not only require access to certain data on the mobile device, they also collect potentially sensitive data such as the frequency and type of medication the user wants to be reminded of. The paper at hand presents an analysis of mobile apps offering operational scope in the health sector. Covered elements are the apps' permission profile, their transmission behavior and their compliance with privacy regulation.

Introduction

There is a multitude of statistics about the spread and use cases of modern technology. A report from Gartner, for instance, states, that world-wide, smartphone sales amounted to 1,4 billion devices in 2015. A nearly three-fold increase compared to 2011 where a number of 472 million devices was estimated – and the market is growing. According to the same statistic, the market is dominated by Google's operating system Android and Apple's iOS, with 82,8% and 13,9% market share respectively.

It is safe to say that smartphones have become permanent companions for people living in the Western world. Through the day, they provide functionalities which satisfy a diversity of everyday needs: connection (phone calls, messages, social media), organization (calendars, reminders, lists), productivity (learn/work wherever you are), information and entertainment (games, news, magazines), just to name a few.

While supporting a wide set of features and activities, mobile devices handle diverse types of user data. They store contacts, appointments, notes, personal messages. In a world, where data is coined "the new oil" a multitude of actors is interested to tap into this valuable information source.

In today's online economy, personal information has become the currency. There is a lot of providers who legitimately ask for user information to provide their services. A calorie-counting app will need data about the user's meals, a menstrual calendar is designed to collect and process the respective, though personal, data. Anyhow, in an environment where more information means more revenue, even well-meaning app developers can be tempted to collect more than the absolutely necessary for their service. User's are reportedly negligent when it comes to reading security details of software (such as privacy policies and terms of service). Thus, they might not even notice that the tiny, helpful app on their phone knows more about them than it should.

The number of applications in the app stores is too vast to run detailed analyses on all of them. Therefore, this study thus focuses on a sub-class of mobile apps which handles specifically sensitive information. There is a increasing number of mobile software which targets applications in the health and medical sector: track your calorie intake, your steps, your sleep; set automatic reminders for medicines and healthy habits, get involved in self-help communities with people with similar ailments. For such a software to work properly, it *needs* access to user information which may reveal habits, preferences and weaknesses. If spread or misused, however, the same information may cause harm to the user's life situation and reputation. Actors which are specifically interested in such private information are, for instance advertising networks, insurance companies, and even governmental organizations.

High-quality data can be sold for a good price. Actually, for such a good price that it becomes profitable for a malevolent actor to produce and publish a running app with a health function. The users receive a factual counter value – but are often unaware about the trade done with their personal data in the background.

This study analyses a sample of popular health and fitness apps from the German Google Play Store. The focus here is the analysis of the legal documents which are to accompany software products which treat personal user data[1]. It extends a study presented in [12] which analyzed the apps' permission requests – and related it to the presented functionality.

[1] Indeed, German regulation forces service providers to host information about their terms and data practices in an accessible way.

Related Work

This paper focuses on the analysis of Android applications from the Google Play Store. Android is by far the most distributed mobile operating system and seems thus a reasonable starting point for a general analysis.

Of course, we are not the first to consider Android security. Given the vastness of the respective body of research, the follwoing section places emphasis on works which focus on privacy aspects. Subsequently, we give a survey of legal/guiding frameworks which serve the evaluation of privacy compliance in mobile applications.

For the sake of conciseness, we exclude two neighboring topic areas from this review: Firstly, wearable devices will not be considered – the sample contains only standalone apps (even though some of them can optionally be extended by wearable devices). The considerations of this article focus on Android applications. Articles concerned with e.g. iOS privacy can be found in [4, 14].

Android privacy analyses

The examinations targeting the privacy compliance of Android applications can be roughly classified in the following categories (equivalent to general security analyses, as proposed in [11]):

1. **Static analysis:** Evaluation of the application's program code, without actually executing it;
2. **Permission analysis:** Type of static analysis, concerned with the assessment of the access permissions reserved by the application and their necessity for the offered functionality;
3. **Dynamic analysis:** Observation of the program behavior on execution, its processes, created files, and network activities.

Some authors proposed the evaluation of applications outside the laboratory environment, as for instance through crowdsourcing solutions [1].

Static analysis

IccTA [16, 15] analyses the data flows between the mobile applications' components. It uses reverse engineering techniques to detect code snippets which leak user information to external sinks. Using the tool ApkCombiner, the tool is able to detect even privacy leaks which arise from a combination of several apps. Flowdroid [2] uses similar techniques, and provides the platform FlowDroid for the performance comparison of different taint analysis tools. Other tools and analyses have been presented in [8] and [19].

A general literature review of Android apps' static analysis can be found in [17].

Also a static analysis method in the privacy context is the examination of the legal documents accompanying the mobile application. Privacy legislation often demands the communication of certain basic information about privacy practices – usually compiled in documents such as the terms of use of a software or a privacy policy. A survey similar to the one presented in this article has been, for instance, provided in reports of the Future of Privacy Forum (FPF) [6].

Permission analysis

Permission analysis is an Android-specific type of static analysis. The operating system runs every application in a distinct sandbox. Access to resources outside its own sandbox have to be explicitly requested. These "permissions" can be granted by the system automatically (in non-critical cases) or will be explicitly asked for to the user[2].

An early tool for the automatic analysis of permission requests was presented in [27]. [24] uses Bayesian statistics to classify apps into permissible instances and potential privacy risks. Other works examine the human factor in permission granting – [9] performs a user study which shows that even expert users may grant information access too lightheartedly; [18] examine the influence of graphical, color-coded interfaces on the users' willingness to grant permission requests.

Dynamic analysis

Dynamic analysis tools monitor app behavior – locally and in network traffic – for a classification into permissible and data-leaking instances. Examples have been presented in [29, 22]. Some approaches focus on network transmissions, such as AppIntent [28]. [5] presents an approach for dynamic taint analysis.

Some systems combine features of static and dynamic analysis, e.g. [7, 26]. For a more detailed review of Android privacy analysis, please refer to [20].

Evaluative and regulatory frameworks

The human right to privacy is a concept which is mirrored in over 150 national constitutions worldwide[3]. This may allow the assumption that privacy is a basic human need, transcending cultural bias.

Anyhow, while relying on similar principles, national legislative texts mirror cultural differences in the conception of privacy. The resulting differences can create unknown complexities when is comes to online services. Data transfers on the Web regularly cross national borders, data storage facilities are seldom situated in the same region as the service provider and/or the end user.

It was thus necessary to base the evaluation of the mobile apps on the conditions in one specific legal context. For the present moment, we decided to settle the analysis close to home – and used the principles fixed in the German Data Protection Act as a baseline. The first of the following paragraphs will give a short introduction in its core requirements.

Legislation specifies the formal minimum requisites which a manufacturer has to provide concerning user privacy. There are other interest groups which provide guidance to this respect: Governmental and non-governmental organizations may publish recommendations, standards which detail and augment the legal baseline. Certification schemes aim to facilitate transparency for consumers. The second paragraph gives a short introduction to existing privacy guidelines.

[2] "Working with System Permissions" `https://developer.android.com/training/permissions/index.html`

[3] `https://en.wikipedia.org/wiki/Right_to_privacy`, last edited 2016-10-08, retrieved 2016-10-13

Legislation:

As outlined above, this analysis focuses on the provisions of the Germany's Federal Data Protection Act (GDPA). The German data privacy regulations are considered as among the strictest in the world. Anyhow, the following data protection principles are also part of the General Data Protection Regulation which will take effect in Europe in May 2018.

Prohibition with reservation of permission: It is strictly prohibited to collect, process and exploit personal data, unless it is explicitly permitted by the law or the data subject consented to it.

Immediacy: Personal data is to be collected directly from the data subjects, unless there is an explicit legal exception or the collection would involve disproportionate effort.

Proportionality: The principle of proportionality accounts for the complexity of competing interest in privacy regulation. It aims to balance the needs of the involved actors – privacy, technical feasibility, freedom of expression [13].

Data avoidance and data economy: The data collection is to be limited to the minimal extent necessary to achieve the defined objective of the data processing. Personal data should only be stored if this is indispensable. The manufacturer should make use of anonymization and pseudonymization techniques to further reduce the amount of personal information in the collected datasets.

Earmarking: The data collection is related to a specific purpose to which the manufacturer has the legal permission, or the data subject consents to. Re-purposing data for another objective is not allowed, unless the individual explicitly consents to the new data processing.

Transparency: Before the collection of personal data, the data subject has to be informed about the extent of the collection, its purpose, storage conditions and retention periods. The individual consents to this specific data processing and may revoke the consent at any time. The manufacturer has to provide means and mechanisms for the data subject to review, correct, block and delete the collected data.

The here presented analysis aims for an evaluation of compliance concerning the last principle *Transparency*. Privacy policies are the common way of communicating data practices to the users of a service – and receiving the necessary consent for data practices not covered by the law. In consequence, they have to present the necessary information for an informed decision: collected data types, transfer and storage conditions, processing purposes and techniques, potentially used anonymization/pseudonymization techniques, and the mechanisms which allow them to influence the data content and processing.

Guidelines

The analysis at hand focuses on a baseline evaluation – we aim to find out if mobile applications do at least respect the legal standards of their distribution area. Anyhow, there are several sets of guidelines and recommendations which aim to provide app developers with assistance with respect to privacy-oriented app design.

In Germany, the "Düsseldorfer Kreis, a committee of independent data protection officers, issued a document for this purpose [3]. It delimits the areas of applicability of the German DPA, transfers the data protection vocabulary to mobile applications and specifies the requirements to consider.

Of particular interest for our analysis is the first point: When is German data protection regulation relevant for the data processing of a mobile app? According to the document, there is two main cases when the German DPA applies:

1. The manufacturer has a seat or data processing entities on German soil.
2. The manufacturer has no seat in Germany, nor in the European Economic Area, but collects/uses data from people in Germany.

For the sake of completeness, it shall be stated that other organizations provided similar recommendation catalogs, e.g. the US Food and Drug Administration (FDA) [25], and the British National Health Service (here with a focus on health apps) [21].

Certification

Another set for the evaluation of privacy compliance comes from certification schemes. These are often aligned with legal standards, but may exceed the requirements defined by the law. Certifying entities usually motivate their effort with a simplification for the consumer. Comparative results and quality labels are meant to provide orientation in a confusing diversity of competing offers.

In the case of medical apps, there is no definite standard for certification. An early attempt was presented by Happtique in 2013, in form of the HACP label [10]. After some of the certified products showed problematic security issues in other tests, the company suspended the program[4].

The pioneering certification program is replaced by current efforts:

- **Certifications:** Other actors follow in Happtique's footsteps, such as the mHealth Label[5] in the French region.
- **Peer review:** Some academic sources provide a journal-like peer review process for medical applications, such as the Journal of Medical Internet Research[6] and the MIT spin-off Ranked Health[7].
- **Test platforms:** Online platforms such as iMedicalApps[8] provide test reports of medical apps, compiled by experts from different domains.

Notably, many of the presented certification schemes focus on the medical viability of the mobile application. While some technical details such as transfer encryption are touched, the main concern is the usage of correct and current scientific sources, reliable implementation and updates.

[4] `goo.gl/epaZLW`, retrieved 2016-10-23
[5] `http://www.mhealth-quality.eu`
[6] `https://mhealth.jmir.org/announcement/view/67`
[7] `http://www.rankedhealth.com`
[8] `http://imedicalapps.com/`

Analyses and results

The focus of this work is the analysis of the applications' privacy statements. The analysis of their permission requests and network activity has been presented in [12]. A connecting element of both papers is the used sample set which is described in a first section. Subsequently, we outline the results of the permission analysis as they relate to the survey at hand. Finally, we describe insights gained from the document analysis, summarizing the overall contents and some extreme examples of data collection.

The subsequent sections are dedicated to the main focus of this paper, the analysis of the formal documents which accompany a mobile app. A first section reviews formal requirements and their implementation in the sample set. Secondly, we describe the results of the document content analysis.

Sample set

The sample set preferably uses products which are offered by German app developers. Like this, it can be safely assumed that the respective manufacturers are aware of German privacy regulation, and, by the criteria given in Section , subject to its requirements.

The sample apps have been chosen with breadth in mind. The goal was to use products from a wide variety of sub-domains. The sub-categories have been chosen following the classification proposed by [23] and are as follows:

1. Support and reminders
2. Explanation and revision of diagnoses
3. Search for medical information
4. Search and comparison of medical institutions
5. Risk monitoring (allergies, diabetes etc.)
6. Fitness tracking
7. Calorie tracking
8. Recipes
9. Contraception/Fertility tracking
10. Baby diaries
11. Sleep tracking
12. Stress handling, mental health

The sample set contains five apps for each of the categories. They were chosen based on the origin of the developer (preferably German) and popularity (apps with higher download numbers were preferred). Only one category lacked examples from the German market, the sleep trackers. In this case, foreign samples have been used.

The document analysis has been performed some weeks after the laboratory tests presented in [12]. Two of the original apps were not available anymore at this point in time and had to be excluded from the sample set: "Meine Elternzeit", a baby and parenting app, and "Sleep Tracker, Version 1.2" (by Uevo LLC), a sleep tracking app.

Permission and network analysis

The first work package analyzed the permissions requested by the examined apps – and tried to draw a relationship to their offered functionality. This allows to qualify the requests into intelligible, unclear and downright unreasonable ones.

Overall, it can be stated that there is a tendency to request more than the absolutely necessary access. In the test set, only ten applications did *not* demand for any of the permissions Android defines as "dangerous". 25 abstain from permission of the "normal" category. Only 8 samples do not ask any of the regular permissions, but none of them also avoids permissions from the "Other" permission category.

Especially popular is the request for the user's location, derived from the device's GPS sensors. 26 of the 60 applications ask for the respective permission. Looking at the app functionalities, however, the request is only reasonable in 9 cases. 5 more can be considered borderline cases – they may include functionalities for the localization of nearby facilities such as laboratory offices and diaper changing tables (in baby apps). The 12 remaining apps do not mirror the permission in their functionalities.

Similar conditions could be found for the requests for camera access, 12 in total. Only 2 apps have matching functionalities, 5 more could use the camera for some side functionality. For 5 apps, however, it is hard to find a reasonable explanation based on the apps' core functionalities e.g. when it comes to health magazines ("Vigo Gesundheitsmagazin"), pharmacy searches ("Apotheke unterwegs" and "Apotheke vor Ort"), step counters ("Schrittzhler-App BG Verkehr") and pain diaries ("CatchMyPain").

Policy analysis

Formal requirements

Availability of privacy statement: The Google Play Store offers a dedicated space in the app description to link manufacturer information and terms of use. This is the first logical place to search for information about the manufacturers' identities and their data practices.

However, this intuitive positioning is only used adequately by a fraction of the app providers. In 32 of 58 cases, the app description features a link labeled "privacy policy" or the like. Only 22 of them lead directly to the description of the app's data practices. 5 examples link the webpage of the app developer, the other 5 raise an HTTP error.

To extend the sample set, we actively searched the manufacturer's web presence for privacy-related documents. In some cases, we installed the app on a test device and retrieved the information during the installation process or from the running app. Through active search, we could retrieve some kind of data-related document for 42 of the 58 app samples. Anyhow, only 19 of these policies referred specifically to the mobile application, in contrast to a multitude of documents which seem to specify data collection at the company website or by the manufacturing company in general. Most of the document samples could be retrieved in German language (38); if no German version was available, the analysis was based on the English alternative.

German regulation requires clear information about the privacy policy's validity, and contact data which allow questions and demands. Only 15 of the found 42 documents include a date of validity; in two cases the statement was as old as from 2011, about half of the policies have recent validity dates (7 examples, validity from 2015/16). Basic contact information were provided by 28 app providers, usually in form of an email address. Less often, the user is provided with an explicit name of the contact person (10 cases), a postal address (17 cases) or a telephone number (7 cases). This leaves 30 cases in which a user is left to search by herself for the appropriate contact information.

For informational purposes, we classified the manufacturers based on their general involvement in health-related business. The majority of developers were individuals or small software companies (20/58). 17 of the apps were provided by bigger firms with some relation to health, e.g. publishers of medical magazines. In 10 cases, the company had strong ties to the medical business, e.g. insurances, pharmaceutical companies, pharmacies. One app was provided by a university within the framework of a research project.

Content

In a second working step, the content of the found privacy policies has been analyzed. To the best of our abilities, we limit the considerations to the information collection happening in the mobile applications. Collection types which clearly relate to the company's web presence have been excluded.

Collected user information The policies in the sample set referenced over one hundred different data types overall. It is important to note that the following passages refer to information collection the manufacturer asks the user to consent to. No technical analysis has been performed to confirm the mentioned information types are, indeed, collected and submitted. Hence, the analysis outlines *potential* data collection based on the granted consent.

The variety of collected information depends on the application's purpose. Based on the app's functionality, certain data collection procedures can be clearly necessary. Anyhow, it is not always obvious, why user information is not processed locally, but also submitted to the manufacturer's servers.

Most applications collect some *basic information about the user* herself – 21 wish to store the person's name, 22 an email address. In many cases, the companies desire extended contact data – such as a phone number (12/42), a fax number (3/42) and even a home address (14/42). Popular is also the question for the user's gender (7/42), age (7/42) and birth date (7/42). Less often, we see the question for weight (3/42) and size (4/42).

Apart from this generic information, some applications also want to collect highly sensitive facts from their users. Two apps from the category "Search for medical professionals" potentially store the user's type of health insurance, namely "Weisse Liste" and "Jameda". The former furthermore adds the user's patient ID and the insurance company ID. The latter additionally sends the patients stated reason for the appointment and symptoms to the company server. The site and app "DocCheck" stores medical specialty and the respective proof document of registered medical professionals.

Across certain app categories the information demand varies importantly – even though the products offer equivalent functionalities. The "FDDB Calorie Counter" demands information such as the user's movements, changes in altitude, taken meals and not further specified context data. However, two other examples of calorie counters offer similar functions while claiming to not collect any user data.

Some of the apps offer social components. The above mentioned app "Jameda" mainly offers a platform for the search and scoring of medical professionals. Anyhow, it also provides a social networking possibility – therein, the symptoms stated by the user are used to suggest other users with similar health problems. While supporting the formation of self-help groups, this may also leak sensitive symptom information to other users. Connections to external social networks may entail further leakage of personal data.

The situation gets even more uncomfortable when the user is urged to share information of third parties. This is mainly an issue in apps surrounding family planning. Especially baby and parenting apps collect extended information about the user's partnership and the offspring's name, birthday, and development. Especially the application "Meine Elternzeit" motivates users to share maximum detail.

Collected device information User-generated contents are in general extended by collecting additional information directly from the user's device. This data comprise technical data, such as the device model and firmware, but also sensor inputs, such as GPS location and sounds perceived over the microphone.

For the app developer, device data have the advantage that they can be collected automatically. This means less interference with the user's normal usage patterns – and less user awareness about the amount of collected data. One core task is to identify the individual user over time – and hardware information can be very helpful to do so. Thus, identifiers are a very popular data point:

- 26/42 collect the device IP address;
- 9/42 collect the device ID (IMEI);

Few products also collect the device's MAC address and SIM-attributed phone number. Interesting here is, however, a discrepancy. From the 42 application with a privacy policy, 21 demand the permission access to the phone's device data, but only nine mention this in the respective document.

A similar discrepancy stands out when it comes to the collection of location information. 26 of the 42 apps request the permission to access the device's GPS functionality. In the privacy policies, only 8 samples mention this fact.

Apart from factual device information, behavioral data is a second big interest. With the right permissions, a mobile application can monitor the user's web surfing, app installation and usage and interactions with media. 6 of the reviewed samples ask for the permission to access the device's app history. This enables to collect the list of applications which are installed on the device. Anyhow, this is hardly ever mentioned in the privacy policies. The documents rather state limited user monitoring in the own application (8/42), and in newsletters and other communications (4/42).

Some of the apps state information collection which stands out with respect to the rest of the samples. The "Chefkoch" recipe app claims to follow the user's online activity and to collect detailed click streams. The fertility app "MyDays X" gathers a list of the installed apps on the device, and claims to monitor the user's shopping behavior (without specification how this is done).

The Adidas fitness app catches attention in two dimensions: On the one hand, the product collects a wide range of user information. As stated, this comprises contact histories, product ratings, participation in loyalty programs and in-app acquisitions. On the other hand, the vendor states to link this information with other information. One of the possible sources being your activity in the company's real-world stores.

Context of data collection Full disclosure comprises information about when, how and for what purpose the user information is collected. In comparison to the description of the collected data types, however, this part of the privacy policies appears rather fragmentary.

The most covered aspect in this context is the collection purpose. Most commonly, however, with the objective to attain the user's consent to data processing beyond the app's main function. The statements include the permission of

- statistical analysis, without specification of purpose (16/42);
- website and service optimization (15/42);
- marketing (12/42);
- profiling/personalization (10/42).

Furthermore, the manufacturers wish consent to contact their users on different communication channels. The most common way being email, in 15 of 42 cases, but also via traditional postal service (6/42) or phone calls (3/42). Noticeable here is the Adidas fitness app who claims to use SMS and "other technical means" (not further defined) besides all of the above to make contact.

Data storage, security and transfer Information about the manufacturers' practices with respect to data security are scarce, at least in the official documents. (Further information about actual storage conditions, as found under laboratory conditions can be found in [12].)

Most manufacturers (22/42) give some generic statement about the security of their data storage solutions, some mention access limitation and authentication procedures. In contrast to other product groups, none gives indication of security certifications of their data centers. The topic of data security at partnering companies is hardly broached.

Equally concealed are the manufacturers' retention practices. Only 9/42 even mention retention periods. Anyhow, none of them gives a full account. The statements commonly refer to user account data and their retention after account deletion. There is hardly any reference to automatically collected device and usage data and their aggregations.

Data transfers in third countries are the rule. Commonly, the user is informed that these third countries may have "lower privacy standards". The wide usage of Google Analytic products suggests that most usage data are at least transferred to the U.S. However, based on the privacy policies no detailed deductions are possible.

Data processing and sharing In today's business environment, it is common practice to outsource certain business processes. Most commonly, user data is shared with external partners for processing and analytics (12/42) and for storage (6/42). Commonly, the privacy policies do not discuss the security and privacy conditions at the premises of the chosen partners. In few cases, a limitation of the partner's access is mentioned.

In 5 cases, the users are asked to consent the distribution of their information within the company group. In all of them it is left to the user to find out which entities this entails. 6 samples mention the possibility of a company merger and/or sale – which would include the user data as an asset. 4 manufacturers maintain the right to share user data with research facilities.

An concerning example of information sharing is the app "CatchMyPain". It serves as a symptom diary, collecting data about a user's ailments, symptoms and medication. All these information are compiled to a user profile in a kind of "medical social network". The objective is to relate similar users and enable the formation of self-help groups. However, the policy does not state how a user can control the information flows towards other users. There is thus a possibility of unwanted leakage of personal diary information. Apart from that, the company states to share "profile and health data" with healthcare professionals, researchers and manufacturers of medical products. While this excludes direct transmission of a user's name and email address, it includes socio-demographic information (gender, age, birthday), journal entries, pain profiles, diagnoses, treatments and other contents.

Cookies are a popular means to track a user's online itinerary. 24 of the 42 manufacturers mention to them in their products. Not always the type and origin of the cookies is detailed. Overall, the privacy policies reference a totality of 54 different trackers. Especially popular are google-derived ad trackers (analytics (13), adwords (3), adsense (5)), and social networks (facebook plugin (11), google+ (8), twitter (5), youtube (1)). Two manufacturers claim to use a solution based on Piwik[9] – which can be hosted locally to avoid data dissemination and ensure user privacy. Anyhow, none of the examples state if their own facilities or a cloud-based solution.

Conclusion

For many users, mobile devices are permanent life companions – always close, always active, always online. They do support traditional tasks, such as phone calls and messaging. Flexible mobile applications enable them to branch out into a diversity of domains – using the internal sensors and connecting to additional wearable devices. This entails that the individual devices also send a continuous stream of information. On the one hand to the device manufacturer, on the other hand to the developers of apps with the corresponding permissions.

The pervasiveness of mobile devices enables them to exert wide influence on the users' habits. Health apps have thus become a support tool for the user to monitor health-related behaviors and to nudge healthier habits. However, they also enable their manufacturers to capture a multitude of real-life data. In a context where more data means more revenue, the temptation rises to collect more than absolutely necessary – and to disrespect the users' privacy.

This study reviewed the privacy policies of 58 popular mobile apps from the German Google Play Store. Transparent communication of privacy practices means to provide the necessary information in an (a) accessible, (b) comprehensible and (c) complete way. Our examination could show that none of the samples fulfilled these criteria perfectly.

Less than half of the apps (22/58) provide a correct direct link to their privacy statements in the app description page. Enhanced search allowed us to retrieve documents for 42 of the samples. Anyhow, only 19/58 explicitly referred to the data collection in the mobile app – as opposed to generic privacy policies cover-

[9] `https://piwik.org/`

ing all services offered by the company.[10]

In 90% of the cases, the privacy policy could be retrieved in German language. Four apps developers only made an English version of the document available. Bigger companies show a tendency to use generic, legalese formulations, while individual developers rather provide copied standard texts.

None of the examined samples provides complete information in their privacy statements. There is a strong emphasis on naming the data types which are collected. Other topic areas, as the moment of collection, processing purpose and storage condition are sometimes mentioned, but never elaborated. It is common to demand the user's consent for generic processing practices, such as "storage in some other country with possibly lower data protection standards" or "transmission to external partners". This leaves the users without a clear idea where their data goes to and who potentially has access.

When comparing the stated information collection with the apps' functionalities, we found several discrepancies. From menstrual calendars who track the user's location and installed applications, to fitness apps who link the collected data to other information sources. The data hungriness is also mirrored in the apps' permission requests: From overall 186 permission requests in all samples, only 109 are directly related to the application's functionality. In contrast, 35 are clearly beyond the app's scope (in addition to 42 which could be clearly decided). Furthermore, the practices stated in the privacy policy often do not match the permission requests on installing the app. In many cases, the granted permissions allowed access to a wider set of user information than was explained in the policy documents.

The goal of this study is not to assign blame. Especially the "extreme case" descriptions should be taken with a grain of salt. After all, these are examples of manufacturers who provide information about their intentions up-front – in contrast to several companies which offer no or very limited information about their data practices.

Large-scale health information from a multitude of different users are a big chance for medical research. In contrast to expensive studies, app-delivered data are cheap and, given self-interest of the user, mostly correct. They offer exciting possibilities to discover interactions between symptoms, medical treatments and patient behavior. However, it is important to openly communicate with the user how and by whom the data is used, how it is combined with other sources and, if the case may be, how her anonymity is protected. Furthermore, it should be allowed to consent to data usage for research purposes without automatically admitting to targeted marketing.

References

[1] Y. Agarwal and M. Hall, "Protectmyprivacy: detecting and mitigating privacy leaks on ios devices using crowdsourcing," in *The 11th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys'13, Taipei, Taiwan, June 25-28, 2013*, H. Chu, P. Huang, R. R. Choudhury, and F. Zhao, Eds. ACM, 2013, pp. 97–110. [Online]. Available: http://doi.acm.org/10.1145/2462456.2464460

[2] S. Arzt, S. Rasthofer, C. Fritz, E. Bodden, A. Bartel, J. Klein, Y. L. Traon, D. Octeau, and P. McDaniel, "Flowdroid: precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps," in *ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '14, Edinburgh, United Kingdom - June 09 - 11, 2014*, M. F. P. O'Boyle and K. Pingali, Eds. ACM, 2014, p. 29. [Online]. Available: http://doi.acm.org/10.1145/2594291.2594299

[3] *Orientierungshilfe zu den Datenschutzanforderungen an App-Entwickler und App-Anbieter*, Düsseldorfer Kreis - Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich, Bayerisches Landesamt für Datenschutzaufsicht, Promenade 27, 91522 Ansbach, Jun. 2014. [Online]. Available: https://www.lda.bayern.de/media/oh_apps.pdf

[4] M. Egele, C. Kruegel, E. Kirda, and G. Vigna, "Pios: Detecting privacy leaks in ios applications," in *Proceedings of the Network and Distributed System Security Symposium, NDSS 2011, San Diego, California, USA, 6th February - 9th February 2011*. The Internet Society, 2011. [Online]. Available: http://www.isoc.org/isoc/conferences/ndss/11/pdf/9_2.pdf

[5] W. Enck, P. Gilbert, B. Chun, L. P. Cox, J. Jung, P. D. McDaniel, and A. Sheth, "Taintdroid: An information-flow tracking system for realtime privacy monitoring on smartphones," in *9th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2010, October 4-6, 2010, Vancouver, BC, Canada, Proceedings*, R. H. Arpaci-Dusseau and B. Chen, Eds. USENIX Association, 2010, pp. 393–407. [Online]. Available: http://www.usenix.org/events/osdi10/tech/full_papers/Enck.pdf

[6] Future of Privacy Forum (FPF), "FPF mobile apps study 2016," Future of Privacy Forum (FPF), Tech. Rep., Aug. 2016.

[7] D. Geneiatakis, I. N. Fovino, I. Kounelis, and P. Stirparo, "A permission verification approach for android mobile applications," *Computers & Security*, vol. 49, pp. 192–205, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.cose.2014.10.005

[8] M. I. Gordon, D. Kim, J. H. Perkins, L. Gilham, N. Nguyen, and M. C. Rinard, "Information flow analysis of android applications in droidsafe," in *22nd Annual Network and Distributed System Security Symposium, NDSS 2015, San Diego, California, USA, February 8-11, 2015*. The Internet Society, 2015. [Online]. Available: http://www.internetsociety.org/doc/information-flow-analysis-android-applications-droidsafe

[9] J. Haggerty, T. Hughes-Roberts, and R. Hegarty, "Hobson's choice: Security and privacy permissions in android and ios devices," in *Human Aspects of Information Security, Privacy, and Trust - Third International Conference, HAS 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015. Proceedings*, ser. Lecture Notes in Computer Science, T. Tryfonas and I. G. Askoxylakis, Eds., vol. 9190. Springer, 2015, pp. 506–516. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-20376-8_45

[10] *Health App Certification Program (HACP) – Certification Standards*, Happtique, Feb. 2013. [Online]. Available: https://s3.amazonaws.com/cdn1.hubspot.com/hub/219577/HACP_Standards_FINAL_2.pdf

[11] H. Kang, J. Jang, A. Mohaisen, and H. K. Kim, "Detecting and classifying android malware using static analysis along with creator information," *IJDSN*, vol. 2015, pp. 479 174:1–479 174:9, 2015. [Online]. Available: http://dx.doi.org/10.1155/2015/479174

[12] J. Knackmuss, E. Clausing, and R. Creutzburg, "Investigation of three security relevant aspects of android ehealth apps - permissions,

[10] FiXme Note: include "install app" as a working step for search int he above section

storage properties and data transmission," in *Proceedings of the Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applications Workshop*, 2017.

[13] J.-M. Koch, *The Privacy, Data Protection And Cybersecurity Law Review*, 2nd ed. Law Business Research Ltd., 2015, ch. Germany, pp. 119–133.

[14] B. Li and Z. Feng, "A system for privacy information analysis and safety assessment of ios applications," in *International Conference on Security and Privacy in Communication Networks - 10th International ICST Conference, SecureComm 2014, Beijing, China, September 24-26, 2014, Revised Selected Papers, Part II*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, J. Tian, J. Jing, and M. Srivatsa, Eds., vol. 153. Springer, 2014, pp. 392–398. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-23802-9_31

[15] L. Li, A. Bartel, T. F. Bissyandé, J. Klein, Y. L. Traon, S. Arzt, S. Rasthofer, E. Bodden, D. Octeau, and P. McDaniel, "Iccta: Detecting inter-component privacy leaks in android apps," in *37th IEEE/ACM International Conference on Software Engineering, ICSE 2015, Florence, Italy, May 16-24, 2015, Volume 1*, A. Bertolino, G. Canfora, and S. G. Elbaum, Eds. IEEE Computer Society, 2015, pp. 280–291. [Online]. Available: http://dx.doi.org/10.1109/ICSE.2015.48

[16] L. Li, A. Bartel, J. Klein, Y. L. Traon, S. Arzt, S. Rasthofer, E. Bodden, D. Octeau, and P. McDaniel, "I know what leaked in your pocket: uncovering privacy leaks on android apps with static taint analysis," *CoRR*, vol. abs/1404.7431, 2014. [Online]. Available: http://arxiv.org/abs/1404.7431

[17] L. Li, T. F. D. A. Bissyande, M. Papadakis, S. Rasthofer, A. Bartel, D. Octeau, J. Klein, and Y. Le Traon, "Static analysis of android apps: A systematic literature review," SnT, Reports : Internal report 978-2-87971-150-8, 2016. [Online]. Available: http://orbilu.uni.lu/handle/10993/26879

[18] J. K. MacDuffie and P. A. Morreale, "Comparing android app permissions," in *Design, User Experience, and Usability: Technological Contexts - 5th International Conference, DUXU 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016, Proceedings, Part III*, ser. Lecture Notes in Computer Science, A. Marcus, Ed., vol. 9748. Springer, 2016, pp. 57–64. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-40406-6_6

[19] W. Meng, R. Ding, S. P. Chung, S. Han, and W. Lee, "The price of free: Privacy leakage in personalized mobile in-apps ads," in *23nd Annual Network and Distributed System Security Symposium, NDSS 2016, San Diego, California, USA, February 21-24, 2016*. The Internet Society, 2016. [Online]. Available: http://www.internetsociety.org/sites/default/files/blogs-media/price-of-free-privacy-leakage-personalized-mobile-in-app-ads.pdf

[20] M. H. Mughees, H. Haddadi, and P. Hui, "Privacy leakage in mobile computing: Tools, methods, and characteristics," *CoRR*, vol. abs/1410.4978, 2014. [Online]. Available: http://arxiv.org/abs/1410.4978

[21] NHS Innovations South East, *App Development: An NHS Guide for Developing Mobile Healthcare Applications*, National Health Service, May 2014.

[22] S. T. A. Rumee and D. Liu, "Droidtest: Testing android applications for leakage of private information," in *Information Security, 16th International Conference, ISC 2013, Dallas, Texas, USA, November 13-15, 2013, Proceedings*, ser. Lecture Notes in Computer Science, Y. Desmedt, Ed., vol. 7807. Springer, 2013, pp. 341–353. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-27659-5_24

[23] Statista, "Nutzung ausgewählter applikationen und -services in deutschland nach funktionsbereich 2015," Statista, Tech. Rep., 2015. [Online]. Available: http://de.statista.com.ezproxy.fh-brandenburg.de:2048/statistik/daten/studie/454390/umfrage/nutzung-von-digital-health-applikationen-und-services-nach-funktionsbereich

[24] O. Tripp and J. Rubin, "A bayesian approach to privacy enforcement in smartphones," in *Proceedings of the 23rd USENIX Security Symposium, San Diego, CA, USA, August 20-22, 2014.*, K. Fu and J. Jung, Eds. USENIX Association, 2014, pp. 175–190. [Online]. Available: https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/tripp

[25] *Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff*, U.S. Food and Drug Administration, Feb. 2015.

[26] M. Xia, L. Gong, Y. Lyu, Z. Qi, and X. Liu, "Effective real-time android application auditing," in *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*. IEEE Computer Society, 2015, pp. 899–914. [Online]. Available: http://dx.doi.org/10.1109/SP.2015.60

[27] W. Xu, F. Zhang, and S. Zhu, "Permlyzer: Analyzing permission usage in android applications," in *IEEE 24th International Symposium on Software Reliability Engineering, ISSRE 2013, Pasadena, CA, USA, November 4-7, 2013*. IEEE Computer Society, 2013, pp. 400–410. [Online]. Available: http://dx.doi.org/10.1109/ISSRE.2013.6698893

[28] Z. Yang, M. Yang, Y. Zhang, G. Gu, P. Ning, and X. S. Wang, "Appintent: analyzing sensitive data transmission in android for privacy leakage detection," in *2013 ACM SIGSAC Conference on Computer and Communications Security, CCS'13, Berlin, Germany, November 4-8, 2013*, A. Sadeghi, V. D. Gligor, and M. Yung, Eds. ACM, 2013, pp. 1043–1054. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516676

[29] Y. Zhou, X. Zhang, X. Jiang, and V. W. Freeh, "Taming information-stealing smartphone applications (on android)," in *Trust and Trustworthy Computing - 4th International Conference, TRUST 2011, Pittsburgh, PA, USA, June 22-24, 2011. Proceedings*, ser. Lecture Notes in Computer Science, J. M. McCune, B. Balacheff, A. Perrig, A. Sadeghi, M. A. Sasse, and Y. Beres, Eds., vol. 6740. Springer, 2011, pp. 93–107. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-21599-5_7

Author Biography

Anett Hoppe received a PhD in computer science from the University of Burgundy, France (2016). Since then she has worked as an IT Security Expert at the AV-Test GmbH in Magdeburg, Germany. Her working focus is on data privacy in modern software applications.

Maik Morgenstern received a diploma degree in Engineering from the University of Magdeburg and is a CEO and the Technical Director of AV-TEST GmbH. He manages the planning and implementation of new test scenarios, our technical innovations and our continuous reaction to new threats.