A Privacy Index for mHealth Apps

Selected Issues in Information Systems
- Information Privacy -

Prof. Dr. Ali Sunyaev

Brüggemann, Thomas
Hansen, Joel

02/14/2016

A Privacy Index for mHealth Apps

In recent years, the growth of the consumer electronics market has seen a boost through the introduction of smartphones and tablet computers (Kim, Kang, Lee & Lee, 2015, p. 59). More and more users are able to install a variety of different applications (apps) on their mobile devices (Adhikari, Richards & Scott, 2014, p. 1). Among them are apps that provide information and consultation on medication and health related topics (de la Vega, Miró, 2014 p. 1), making mobile and portable health care possible (Kim, Kang, Lee & Lee, 2015, p. 59; Kotz, 2011, p. 1) and allowing users to monitor their health related issues, comprehend specific medical conditions and achieve fitness goals (Adhikari, Richards & Scott, 2014, p. 1). By entering private information, for example, medication intake, disease symptoms or blood value data, users expose vulnerable information (He, Naveed, Gunter and Nahrstedt, 2014, p. 1; Kotz, 2011, p. 2; Kumar et al., 2013, p. 235), but in return receive a tailored application experience that provides relevant information. (Dehling, Gao, Schneider & Sunyaev, 2015, p. 2) This can potentially reduce the health care costs and improve well-being in many ways (Kumar et al., 2013, p. 228). In the past, personal health data was managed and stored solely in hospitals, but in recent years, it is collected and managed by mHealth apps as well. Therefore, it is critical to protect users' personal data as well as possible, in order to minimize privacy risks and make mHealth apps most beneficial to users (Kim, Kang, Lee & Lee, 2015, p. 59; Kotz, 2011, p. 2-4). Research is required to analyze and assess the potentials of mHealth apps and their challenges (Kumar et al., 2013, p. 228).

The risk to users is that the personal health related data can be misused. An app service provider could potentially sell the users' data to health insurance companies or job agencies. Insurance companies are interested in the customer's health data, in order to tailor the insurance policies to the individual health status of the customers. An insurance company will most likely prefer healthy people, in offering good insurance rates, over not so healthy people (Zubayd, Saleh, Aloul, & Sagahyroon, 2015, p. 3). It is unclear or only vaguely stated in the privacy policies of app providers, how personal user information is treated or stored. Privacy policies lack a standardised format (Adhikari, Richards & Scott, 2014, p. 3), are typically written in formal legalese and are hard to understand by the majority of people. This leads to comparison issues between apps themselves and their app providers. It is challenging to evaluate the data processing of user information and compare different apps regarding data privacy, because the required information is not given, hidden in legal language or not processable for an average educated human. (Dehling, Gao & Sunyaev, 2014, p. 1-2)

In order to match the fast spread of mHealth apps, traditional research that usually has a long time lag from the initial empirical or theoretical study to final publication, needs to be adjusted, as the technology may be enhanced, before the publication process is completed. Developments of mHealth technologies and the adaption of the society may need to be continuously implemented into current mHealth research. Otherwise, they could produce outdated results, as the initial requirements have become obsolete. Updating the requirements or the data basis, while the research study is ongoing, is against scientific norms. The integration of assessment methods in mHealth holds the potential to continuously improve mHealth over time, adapt to changes, and even enhance mHealth design. (this section follows Kumar et al., 2013, p. 230-231) “Although these methodologic challenges present exciting new opportunities for scientific innovation, the marketplace and consumers are not waiting for scientific validation.” (Kumar et al., 2013, p. 235)

While the usage of mHealth apps is widely spreading and sales are increasing, there is a wide gap between the scientific and commercial perspective on mHealth apps (de la Vega, Miró, 2014 p. 1-2; He, Naveed, Gunter and Nahrstedt, 2014, p. 1). Even though mHealth has a potential to extend and improve health care, most of the apps have not been assessed and validated with the needed degree of accuracy, regarding their privacy risks. A scientific knowledge base needs to be established, by analyzing the benefits of mHealth apps and elaborating solutions for arising concerns, such as security and privacy challenges, as well as ethical issues. The fast growth of the mHealth app market is forcing science to recover lost ground and provide improvement ideas for mHealth apps, based on research findings. (de la Vega, Miró, 2014 p. 1) First steps in this direction have already been accomplished by several institutions, but the connection between commercial-practice and research is weak. (de la Vega, Miró, 2014 p. 2) De la Vega, Miró (2014) mention one initiative, *PatientView* is the host of the web page *myhealthapps.net*, which is recommended by the European Commission. People can review health related apps on a zero to five Likert-scale and share their experiences, making the benefits and concerns of individual apps more transparent. (p. 1-2) In a further step the collected information need to be evaluated and guidelines, or even tools, for improved assessments should be developed. As a result of our elaboration we intend to focus on the path of developing a tool for increased transparency of privacy risks.

As the “mHealth research community is now challenged to develop methods to preserve participant privacy and confidentiality while satisfying research needs” (Kumar et al., 2013, p. 235), we identified a weighted-sum equation privacy risk index as a potential element for our research. It can be used for decision making purposes and pairwise comparisons (Triantaphyllou, Shu, Sanchez & Ray, 1998, p. 1). In order for users to identify the privacy risks and to make apps more comparable, an advanced assessment of each individual app is necessary. Therefore, we will apply the method of weighted-sum equations in our elaboration on the study field of information systems, by assessing and comparing privacy risks of mHealth apps.

Since the Apple AppStore and Google PlayStore do not offer privacy risk assessments or comparison displays and increased efforts in mHealth research are essential to realize the full extent of benefits from mHealth apps (Kumar et al., 2013, p. 235), the research question of this paper will be: How can users be provided with an easy-to-use privacy risk assessment for mHealth apps?

Objectives

Our main objective is the development of an algorithm that generates a privacy risk index, which brings more transparency into the assessment of apps regarding their privacy risks (Tavani, 2007, pp. 1-2). Users benefit from our research by receiving processed information concerning the use of their health related data. This will be a vital improvement to the current state of research that assesses information about privacy policies. Oetzel and Spiekermann (2013) describe a methodology for a step-by-step privacy impact assessment that systematically addresses privacy issues and gives companies and app providers a guideline for designing their privacy practices accurately (pp. 126-131). It is important to provide users with the right information about privacy practices to support decision making. Therefore, we focus on providing processed information about privacy practices for users, by displaying all information on an intuitive graphical interface.

As a secondary objective we determine privacy risk factors by surveying current apps concerning their privacy risks. This will generate valuable information for a more accurate app assessment. The achievement of this goal is important, in order to determine a privacy risk index of mHealth apps. The privacy index of a mHealth app will be based on several categories concerning the input of personal data, a secure data connection and the ability to login to an app (Ackerman, 2013, pp. 12-16, 20-21). In addition to that, the derived privacy risk factors will be weighted with respect to their importance for a preserved protection of personal data. As a result we will develop an easy-to-use graphical interface for the comparison of mHealth app's privacy risks (Palmer, 2002, pp. 151-153).

Method

Dataset. We conduct our study by analyzing a given dataset that initially contains a list of 600 apps. The dataset includes 300 apps from the Google PlayStore and 300 apps from the Apple AppStore. The apps were selected from the app store categories "Medical" and "Health & Fitness". The app name, description, price, app store id, discovery date and the permissions[1] the app requires to run on a device, as well as the average rating and the total rating count, are included in the dataset. Since our research procedure requires the installation of apps on our mobile devices, we limit the original dataset to those 476 apps that are listed as free of charge. We argue that these free apps are potentially more prone to a higher privacy violation risk than paid apps, since their revenue model is usually built around displaying advertisements to the user and therefore gather user related data to personalize the advertisements (Ackerman, 2013, pp. 18-21). Moreover, we want to ensure that each app in our dataset has the same chance of being reviewed in equal levels of detail. Therefore, our goal is to be able to download every app, as is feasible, to our smart phones. Based on this dataset, we identify potential factors in regard to their privacy risks.

Four step procedure of identifying privacy risk factors. In order to rate every single app from our reduced dataset, we used the following procedure. With the intention of identifying possible privacy risks, we first read the description of the app inside the app store. We analyze it regarding indicators of privacy related information input that users possibly have to expose, in order for the app to display tailored information.

For the next step, we inspect the screenshots provided by the app store entry. The screenshots could potentially indicate user information that is requested from the user, by showing text input fields for user information, such as medication intakes or diseases. If neither the screenshots nor the app description provides sufficient information on personal

[1] Detailed app permissions listing is an Android only feature

data input, we download the app to our own smartphones and try it out. After downloading the app, we checked the data transfer with the web debugging proxy application *Charles Proxy*[2]. This proxy application allows us to visualize the HTTP connections the app opens. In an optional final step, we read the privacy policy or terms of service to obtain information about the personal data usage, but only if a data transfer displayed in the web debugging proxy application remains uncertain.

Categorization of privacy risk factors. At first, to ensure a balanced quality in our app ratings, each of our team members examined the same set of 10 apps from the Google PlayStore individually. We chose the Google PlayStore, since Android smartphones were available among all team members. These 10 sample apps were randomly selected from our initial dataset and we used them to explore the first categories of personal data input (which we abbreviated as *P* in our final equation). After assessing this test sample, we compared our results among the team members. If there were any differences in the rating of apps among our team members, we discussed the individual ratings and agreed on a consistent procedure.

Early in our research, we defined potential personal data targets (abbreviated as *T* in our final equation). We refer to personal data targets as the target or host destination the app potentially sends the user's personal data to. Personal data could be sent directly to the app provider, as well as research projects, analytics tools and marketing agencies (Adhikari, Richards & Scott, 2014, p. 1). Some apps may offer data storage and syncing on the app providers' remote servers, which leads to a potential privacy risk for users, since, from the user perspective, the data vanishes on a non-traceable and non-retrievable remote server. (Adhikari, Richards & Scott, 2014, p. 4, 7; He, Naveed, Gunter and Nahrstedt, 2014, p. 1-2)

Additionally, we tracked with our proxy application, if the apps uses click tracking analytics tools or contacted advertisement servers to display advertisement banners. We listed

[2] https://www.charlesproxy.com, visited 02/09/2016

those findings in the category "unspecific data targets", since it is unclear to us, if personal data is being exchanged with these target hosts and they might pose a threat to information privacy. In our final equation we defined the abbreviation *U* for the unspecific data traffic to analytics or advertising services.

We differentiated the data transfers between an open and an encrypted HTTP data transmission, which we abbreviated as *S* in our final equation. In case of an encrypted connection, we could only suspect, which data is actually being transferred. A secure data transmission should always be used by mHealth apps, in order to guarantee the confidentiality and integrity of personal data against any misuses during the transmission. (He, Naveed, Gunter and Nahrstedt, 2014, p. 1-2; Kim, Kang, Lee & Lee, 2015, p. 60)

As we proceeded, we extended the categories of personal data input (*P*) continuously as needed. Throughout our research, we identified 13 types of categories concerning privacy risks, as displayed in Table 1. For the sake of this paper we will only be referencing the more critical categories below. We identified the category *"life status specs"* among the user inputs, that we defined to include, e. g. data about pregnancy or smoking habits. *"Medication intake"* contains the amount and kind of medication consumed by the user. The category *"vital values"* represents information such as blood metrics or heart rate. "*Diseases"* and "*Symptoms"* are each assigned to single self-explanatory categories that represent the input of disease and symptom information. (Adhikari, Richards & Scott, 2014, p. 1; He, Naveed, Gunter and Nahrstedt, 2014, p. 4) We identified a distinct set of categories that users have to supply to the apps, which are listed in our dataset. We would like to note, that the categories broadly follow along with the domains of mHealth apps Kumar et al. (2013) extracted in their study. (p. 228-229)

For each personal data input we asked the binary question of reasonableness of personal data collection and assessed accordingly with yes or no. It is possible that apps ask for personal data that is not noticeably used by the app and seems fraudulent to be collected by app providers. In our final equation we use the abbreviation *D* for the reasonableness of the data collection.

Furthermore, we differentiated between two kinds of login categories (abbreviated as *L* in our final equation). If a login is required (Adhikari, Richards & Scott, 2014, p. 4), a user either has to register via a username or an email address, or otherwise via a social media login such as Facebook. In the case that there is no login required we assign those apps the value *"none"*.

Deriving a factor-weight-equation for a privacy risk index algorithm. As soon as we rated all of the apps from our dataset, we used the ratings to derive an algorithm for calculating a privacy risk index that assigns each app with a value on a scale between 0.0 and 1.0. A privacy risk index of 0.0 would indicate that the app poses no privacy risk, due to the data we were able to collect. The value 1.0 on the other hand expresses a strong privacy risk. The privacy risk index is a factor-weight-equation based on the six factors that we surveyed for each app.

Triantaphyllou, Shu, Sanchez & Ray (1998) promote the use of a factor-weight-equation[3] as a decision making support tool. (p. 4) We decided to use the factor-weight-equation, since the privacy risk index functions as a decision support tool for the app users.

[3] The factor-weight-equation, as we call it, is often referred to as the weighted-sum model. We sticked to our factor-weight-equation term since our algorithm distinguishes between factor and weight variables and we want to link the text and algorithm closer together to improve understanding.

The overall equation can be expressed as:

$$PrivacyRiskIndex_{App} = S_{App} * w(S) + T_{App} * w(T) + P_{App} * w(P) + L_{App} * w(L) + U_{App} * w(U) + D_{App} * w(D) \quad (1)$$

where:
S = Security of data connection to personal data targets available?
T = Target destination of personal data
P = Personal data categories as listed in Table 1
L = Is a login required to use the app?
U = Does the app use any analytics or advertising service?
D = Data collection reasonable and not fraudulent?
w() = The weight assigned to the factors above, $w(S) + w(T) + w(P) + w(U) + w(R) + w(D) = 1$

We assigned a set of default weights for the privacy risk factors based on our own perception of the importance that these factors have in contributing to the overall privacy risk. Usually the reliability and validity of measures (such as the weight values in our equation) are established in researches under controlled laboratory conditions. Reliability is achieved when a measure is consistent, meaning it produces the same result under consistent conditions. Validity describes the aspect of a measure to be, what it claims to be. Yet, mHealth apps are used in the real world, by a variety of people with different preferences. Therefore, mHealth assessment methods have to consider several factors, such as diverging user groups, data collection models and information flow. However, in the context of mHealth apps, common measures have to be set which apply across various populations and environments. (Glasgow & Riley, 2013, p. 238; Kumar et al., 2013, p. 230-231)

In the case we would set the privacy risk factor weights ourselves, we simply could not meet all users weight preferences. Not only would our weights lack reliability and validity, but it would also make our graphical interface less beneficial for users. Therefore, we decided that we let the users set their own assumed measures by weighting the influent factors themselves. (Glasgow & Riley, 2013, p. 237-238)

We derived our default factor weights from the risk assessment weights that Ackerman (2013) proposed, and found that we overall adhere with their weighting system overall (p. 20). Again, it is important to note that the users of our user interface can set the weight for each privacy risk factor to suit their own preferences, upon initialization. Our individual default weights for each factor are listed in *Table 2*. We argue that the *personal data target* destination and the quantity and quality of personal information are most important in influencing the privacy risk, since the target host (e.g. remote server) is the point in which the users lose track and control of their data. In relation to these two top factors, we set the weight for the secure data connection lower, because we think that the likelihood of unencrypted data being tracked or sniffed for an individual person is reasonably low in comparison to the immediate risk by the factors *personal data target (T)* and *personal data categories (P).*

After setting the weights for each factor, we took a look at the scoring model of each individual factor. Since the factors *S, U* and *D* are binary coded, no further scoring is necessary. We code a false with 0.0 and a true with 1.0 and as a special case, we only set the factor *S* to 1.0 if the connection was unencrypted and also a *personal data target* was identified. Otherwise we assume that no personal data is being transmitted, the encryption of connections is of no relevance and the score value can be set to the non-risk-bearing case (score value = 0.0). (following Kumar et al., 2013, p. 235)

For the factor *T,* we assigned default scoring values based on our discussion of relative importance in contribution to the privacy risk, as seen in *Table 3*. As soon as it was unknown to us, during our app surveys, where the personal data is actually being sent, we set the score of those apps to the highest privacy risk of 1.0. This is mainly because obscure host names were listed in our HTTP proxy application, shortly after entering personal information was supplied. On the other hand, if we were able to identify the data traffic, we argue that personal information transmitted to *advertising or marketing* companies poses the highest privacy risk and scored the factor with a 0.4 value. Sending data to *Facebook* is scored with a 0.3 value and data transferals to the *app provider* with a 0.2 score, since we believe that those two targets at least aim at providing a tailored user experience value to the user. This leaves 0.1 for the category *research projects.* We assume that the personal data usage in research projects is generally less of a privacy risk, since the first goal of researchers should not be to abuse the data. We still want to model a remaining risk with the factor 0.1.

Coding a scoring model for the factor *P* left us with a decision. A single app can ask for multiple categories of personal data input and our scoring model would need to add the scores for each existing category up to calculate the final score for *P*. We defined 13 categories of personal data input, but the maximum number of categories identified in a single app was 5. This would lead to a single app never reaching the maximum score of 1.0, since only the sum of all 13 category scores could equal to 1.0. It would have been possible to multiply each score with a correction-factor that would set the sum of the 5 highest scoring categories equal to 1.0. This way the apps in our dataset would be able to score the 1.0 for the factor *P.* But instead of correcting the individual factors *P, T* and *R*, we decided to apply a correction-factor to the final privacy risk index as follows:

$$Correction_{PrivacyRiskIndex} = \frac{1.0}{max(PrivacyRiskIndex)} \qquad (2)$$

This way, the app with the highest risk index will always score 1.0 and the lowest always 0.0. The application of the correction score is only feasible if the dataset is large enough and follows the normal distribution. The correction score would distort the relative margins between the privacy risk index values. For our study we would need a larger app dataset size in order to effectively apply the correction score.

Lastly, we implemented a separate score that indicates how well we inspected (or were able to inspect) a given app. The assigned weights are listed in *Table 4*. We state that just inspecting the description text or the screenshots is far less thorough in assessing the privacy risk, than actually downloading the app, inspecting the HTTP connections via a proxy application or even reading over the privacy policy. So we designed our algorithm to always take the minimum weighted source that is present for a given app to use as the factor score. If the rating source field of an app says “App downloaded, HTTP Proxy” the resulting value according to *Table 4* would be 0.2, since “HTTP Proxy” with 0.2 is lower than “App downloaded” with 0.4 and we can assume that we rated the app more thoroughly.

The implementation of the complete algorithm is available on the source code management platform Github[4] under the open source MIT license.

Development of an intuitive graphical user interface for app privacy risk assessment. The next step was to provide the user with a user interface, that allows an easy assessment of the privacy risks, which an app poses and compare the privacy risk factors among multiple apps. Users will be able to select one or more apps and compare each privacy risk factor, in order to gain a fast overview and a quick usage decision, without reading complicated privacy policies. We find that an inter-comparability between apps can most easily be achieved by listing the app rating results in a table view next to each other, as seen in *Figure 1*. Our user interface is available to the public and serves as a first step towards providing a comparison view on apps from the app stores, in regard to their privacy risks. The privacy risk factors are listed on the left side, as vertical table headers. The user is able to enter the name of an app into a search box, as seen in *Figure 2.* While typing, a selection list with matching apps appears below the search box. The list is displayed in *Figure 3*. The user can select an app from the list and the website immediately adds the selected app to the comparison view. This creates an easy to use, fast and responsive user interface, allowing users to customize the view with instantaneous reaction times (Palmer, 2002, p. 154) and tailor the user interface to their needs (Germonprez, Hovorka & Collopy, 2007, p. 355).

As seen in *Figure 1,* we added another user-friendly feature to the user interface, by color coding the privacy risk index with the colors green, orange and red. This way the user

[4] https://github.com/thomasbrueggemann/privacy-risk-mhealth

can, in addition to understanding the number value of the privacy risk index, compare the selected apps with just the glimpse of an eye, by looking at the colors. The color green indicates very little privacy risk, orange intermediate and red stands for a potentially high privacy risk. Dehling, Gao, Schneider & Sunyaev (2015) proposed the idea of clustering apps by "assessments of potential damage through information security and privacy infringements" (p. 1) into archetypes. We used the same dataset as Dehling, Gao, Schneider & Sunyaev (2015) for our study therefore we could easily access the archetype clustering information of the apps (p. 1, 3-4, 9, 11-12). If an app of our dataset is clustered within an archetype, the privacy-risk-index cell also hosts the numbers of the lowest and highest privacy risk index apps from this archetype. These numbers are clickable in order to add the highest and lowest privacy risk index apps to the table view.

We want to stress that the apps that are available for comparison are only those apps, which were included in our initial dataset and that we manually reviewed. It is not possible to compare any random app available on the app stores right away. Every app has to be manually added to our dataset and reviewed one at a time. Therefore, we propose our equation as a prototypical way to evaluate the privacy risk based on our dataset. The equation can be extended to other datasets in the future.

Results

During the assessment of all 476 apps from our initial dataset, we experienced that some apps are not available for download on the app stores anymore. They have been removed by their app providers. This leaves our processed dataset with 298 apps, 147 iOS and 151 Android apps.

We found that there is no direct data transfers to research project hosts (or host names that we could identify as belonging to research projects) and research data use was only described 3 times in the privacy policies. Equally low was a potential data connection to

Facebook with 2 cases. Personal data was sent to the app provider directly within 63 apps. We suspect that personal data was sent to advertisers or marketing companies within 27 apps. These data connections, potentially transferring personal data, established a secure and encrypted HTTP connection within 42 apps, while 28 apps did not encrypt the data connection at all. In 228 cases it was either unclear to us if the data connection was encrypted or the app did not send any data at all. It occurred within 28 apps that personal data was requested without noticeably being used by the app, which seems fraudulent to users. 105 apps requested personal data that was actually used to tailor the app experience to users' needs, and in 165 cases no data input was requested at all. 51 apps required a login via username and password or a social media (via Facebook, Twitter or Google), in order to either be able to use the application at all or just to tailor the experience to the individual user.

Figure 4 shows a histogram of the distribution of privacy risk indices, we calculated for all apps, multiplied by 100 and rounded to the next integer value on the x-axis. The y-axis shows the amount of privacy risk indices in a certain range cluster. The histogram clusters index-values in increments of 7 and clearly shows that the majority of privacy risk indices lay below10. With the exception of the cluster ranging from 28 to 70, the slope of values is constantly decreasing towards higher privacy risk index values.

User interface. The user interface consists out of two main views. Initially the users are presented a weighting view (*Figure 5*), in which the weights of all privacy risk index factors can be customized. The user's weights are stored in a client side cookie.

The second view is the main apps table view. Via a search bar, apps can be added to the table view, as seen in *Figure 1*. As soon as apps are added to the table view, the privacy risk data can be discovered. Hovering a table view cell will display a small, black popup area that provides detailed information on the privacy risk factor. A little yellow bolt icon in front of a table view cell indicates the factor that has the most influence on the privacy risk index of that column. The privacy risk index itself is the large, color coded (green, orange, red) number from 0 to 100. We made the user interface publicly available for everyone to use. It is currently hosted with a free plan on the cloud service provider Heroku[5] and can be accessed under the web address https://privacy-risk-mhealth.herokuapp.com. Due to the free plan, Heroku puts applications that have not been used for 60 minutes into a sleep state.[6] Accessing the web application after these 60 minutes results in a slower response time of the application, since it has to wake up from the sleep state. This is a limitation by the cloud service provider's free plan and should not be taken into account for the following speed assessment. Aside from the initial start after the sleeping period, we put our focus on making the web application response time as fast as possible to increase the usability (Palmer, 2002, p. 163). The response time of the web application is on average below 70 milliseconds (ms) as measured with the *Chrome Dev Tools*[7]. The download delay (Palmer, 2002, p. 155) of the complete application, including all resources such as images, scripts and stylesheets, is below 800ms on average.

Discussion

It is interesting to note, that research projects were identified only three times as potential data targets, during our privacy risk assessment. This underlines the thesis, that the link between the practice world and the research world is not developed enough. This is a dilemma, as the sharing of mHealth data to search projects might pose a privacy risk on one hand, but it also enables mHealth research to conduct studies on real world data.

Further interesting results are that one out of ten apps included in our dataset collects data without a perceived reasonableness for it. We suggest that apps only ask for data that is actually used by the apps to tailor the experience to users needs, otherwise the data usage

[5] https://www.heroku.com/, visited 12/14/2015
[6] https://blog.heroku.com/archives/2013/6/20/app_sleeping_on_heroku, visited 12/14/2015
[7] https://developer.chrome.com/devtools, visited 12/14/2015

appears fraudulent and leaves a negative overall impression of the app. Also, 10% of the apps transfer data without encryption, which Adhikari, Richards & Scott (2014) has already pointed out, by mentioning that “many free mHealth apps for mobile phones send data, connect to third-party sites, use unencrypted connections” (p. 3). Even though the usage of a secured and encrypted data transfer is not visible to users, app providers should implement a secured data transfer. Apps could even be labeled as using a secure connection, for marketing purposes to enhance the trustworthiness.

A reason for the high quantity of low privacy risk indices, as seen in *Figure 4*, is the amount of apps that do not track the vulnerable health data, but rather provide meditation sounds or medical dictionary functionality. These apps pose low privacy risks to users. We tried to find a connection between these apps and their app store categories “Medical” and “Health & Fitness”. An example of a meditation sound app in the category “Medical” with a low privacy risk index is “Control Your Mind BrainMassage”, on the other hand “Calming Music to Tranquilize” with an equally low privacy index, is listed in the category “Health & Fitness”. This may be due to the fact that app provider can select the app store categories by themselves. An exclusion of these apps by app store category is not possible.

Our publicly available user interface offers the chance to enrich the information experience of discovering privacy risks that mHealth apps pose. We believe that research can make use of this data and develop tools and frameworks to address privacy risk drivers in more detail.

Conclusion. With regards to our research question, users are provided with a graphical interface, allowing them to assess and compare privacy risks of mHealth apps in an easy way. The process of selecting the right app out of the overload of available mHealth apps (van Velsen, Beaujean & van Gemert-Pijnen, 2013, p. 1), has been enhanced by our contribution. Users receive processed information about their privacy risks, by using our interface, leading to increased transparency (Zubayd, Saleh, Aloul, & Sagahyroon, 2015, p. 4). Therefore, on one hand the uncertainty about risks concerning the app data usage has decreased and on the other hand individual users are empowered in in making decisions regarded their choice and usage of mHealth apps, by considering our provided privacy risk index. In this way, our contribution supports the narrowing of the gap between the simple existence of information about providers' privacy practices and the actual usage by users.

Future research. In order to scale our idea of creating a unified privacy risk index, we think that future research should focus on automating the process of reviewing the privacy risk factors. An automated process could be able to download apps from the app store and automatically decompile the source code, as well as trace user input fields and the way the input is handled within the source code. This would only be possible to Android apps, due to strict download regulation of the Apple AppStore. In order to circumvent these issues, the app survey process possibly has to be integrated into the app stores by the store providers themselves. Future research could address the need for app store providers to deeper survey and score the way information is treated.

Limitations. As we finish our explanations we want to mention some limitations regarding our research. We were limited mainly in the tracking of personal data transfer. Even though we could track that data was transferred from our smartphones to the app provider or third parties, we could not always track which data was sent. If we would actually have had a possibility to track which data was transferred, our assessment of a privacy risk would be more precise.

Furthermore, we set the default weights of the factors to the best of our knowledge, but a group of experts, such as data protection officers, could possibly assess them more accurately. Users could be surveyed about their perceived privacy risks and implications could be derived for setting accurate weights. We recommend the inclusion of lookalike app names for future research, as a further privacy risk factor (Chia, Yamamoto & Asokan, 2012, pp. 317-319). Moreover, as the data from the dataset was collected in 2013, many apps were not available in the app stores anymore. Also, the dataset included several apps of one app provider that only differed in their names, but not in their functionality, such as meditation sound apps. Even though we still examined a large amount of apps successfully, a current dataset without the redundant apps would be more fitting for future research. Lastly we only examined free apps due to budget restrictions and we suggest that future researchers should rate paid apps as well.

References

Ackerman, L. (2013). Mobile Health and Fitness Applications and Information Privacy. *Privacy Rights Clearinghouse, San Diego, CA*.

Adhikari, R., Richards, D., & Scott, K. (2014). Security and privacy issues related to the use of mobile health apps. ACIS.

Chia, P. H., Yamamoto, Y., & Asokan, N. (2012). Is this App Safe?: A Large Scale Study on Application Permissions and Risk Signals. *Proceedings of the 21st International Conference on World Wide Web*, 311-320.

de la Vega, R., & Miró, J. (2014). mHealth: a strategic field without a solid scientific soul. A systematic review of pain-related apps. *PloS one*, *9*(7), e101312.

Dehling, T., Gao, F., & Sunyaev, A. (2014). Assessment Instrument for Privacy Policy Content: Design and Evaluation of PPC. *WISP 2014 Proceedings*, Paper 2.

Dehling, T., Gao, F., Schneider, S., & Sunyaev, A. (2015). Exploring the Far Side of Mobile Health: Information Security and Privacy of Mobile Health Apps on iOS and Android. *JMIR mHealth and uHealth*, *3*(1), 1-17.

Germonprez, M., Hovorka, D., & Collopy, F. (2007). A Theory of Tailorable Technology Design. *Journal of the Association for Information Systems*, *8*(6), 351-367.

Glasgow, R. E., & Riley, W. T. (2013). Pragmatic measures: what they are and why we need them. *American journal of preventive medicine*, *45*(2), 237-243.

He, D., Naveed, M., Gunter, C. A., & Nahrstedt, K. (2014). Security concerns in Android mHealth apps. In *AMIA Annual Symposium Proceedings* (Vol. 2014, p. 645). American Medical Informatics Association.

Kim, J. T., Kang, U. G., Lee, Y. H., & Lee, B. M. (2015). Security of Personal Bio Data in Mobile Health Applications for the Elderly. *International Journal of Security and Its Applications*, *9*(10), 59-70.

Kotz, D. (2011). A threat taxonomy for mHealth privacy. In *COMSNETS* (pp. 1-6).

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., ... & Hedeker, D. (2013). Mobile health technology evaluation: the mHealth evidence workshop. *American journal of preventive medicine*, *45*(2), 228-236.

Oetzel, M. C., & Spiekermann, S. (2014). A Systematic Methodology for Privacy Impact Assessments: A Design Science Approach. *European Journal of Information Systems*, *23*(2), 126-150.

Palmer, J. W. (2002). Web Site Usability, Design, and Performance Metrics. *Information Systems Research*, *13*(2), 151-167.

Tavani, H. T. (2007). Philosophical Theories of Privacy: Implications for an Adequate Online Privacy Policy. *Metaphilosophy*, *38*(1), 1-22.

Triantaphyllou, E., Shu, B., Sanchez, S. N., & Ray, T. (1998). Multi-Criteria Decision Making: An Operations Research Approach. *Encyclopedia of Electrical and Electronics Engineering*, *15*, 175-186.

van Velsen, L., Beaujean, D. J., & van Gemert-Pijnen, J. E. (2013). Why mobile health app overload drives us crazy, and how to restore the sanity. *BMC medical informatics and decision making*, *13*(1), 1.

Zubaydi, F., Saleh, A., Aloul, F., & Sagahyroon, A. (2015). Security of mobile health (mHealth) systems. In *Bioinformatics and Bioengineering (BIBE), 2015 IEEE 15th International Conference on* (pp. 1-5). IEEE.

Appendix A

Table 1. Personal data information that had to be entered in the apps in our survey clustered in categories and assigned with their weights for the privacy risk index equation

Category	Members of category	Weights
Medication intake	Pills / recipes, Medication dosage	0.147
Vital values	Blood pressure, Heart rate, Blood sugar, Blood values etc.	0.147
Diseases	Kind of disease	0.118
Symptoms	All acute, chronic, relapsing or remitting symptoms. For example: Mood changes, rash, swellings	0.118
Life status specs	Pregnancy, Lifestyle (activity), Smoking habits	0.106
Address	Country, State, Street	0.088
Body specs	Weight, Height, Body frame, Body fat, Temperature etc.	0.082
Family	Medical condition of children or ancestors, Family size	0.059
Medical appointments	Date, Doctor	0.053
Food intake	Calories, diet plan, drinks	0.035
Workout / Activities	Goals, Steps, Distance covered / GPS Tracking	0.029
Personality Test	Questions about own behavior in certain situations	0.012
Sleep Metrics	Sleep sound, dream description	0.006

Note: The sum of all weights is always 1.0

Table 2. Default weights assigned to each of the identified privacy risk factors that together form the factor-weight-equation of the privacy risk index

X	w(X)
T	0.4
P	0.3
L	0.1
S	0.1
U	0.05
D	0.05

Note:
S = Security of data connection to personal data targets available?
T = Target destination of personal data
P = Personal data categories as listed in Table 1
L = Is a login required to use the app?
U = Does the app use any analytics or advertising service?
D = Data collection reasonable and not fraudulent?
w() = The weight assigned to the factors above,
$w(S) + w(T) + w(P) + w(L) + w(U) + w(R) + w(D) = 1$

Table 3. Personal data target destinations identified from surveying the apps and their assigned weights

Category	Weight
Unknown target	1.0
or	
Advertisers / Marketeers	0.4
App provider	0.2
Facebook	0.3
Research projects	0.1

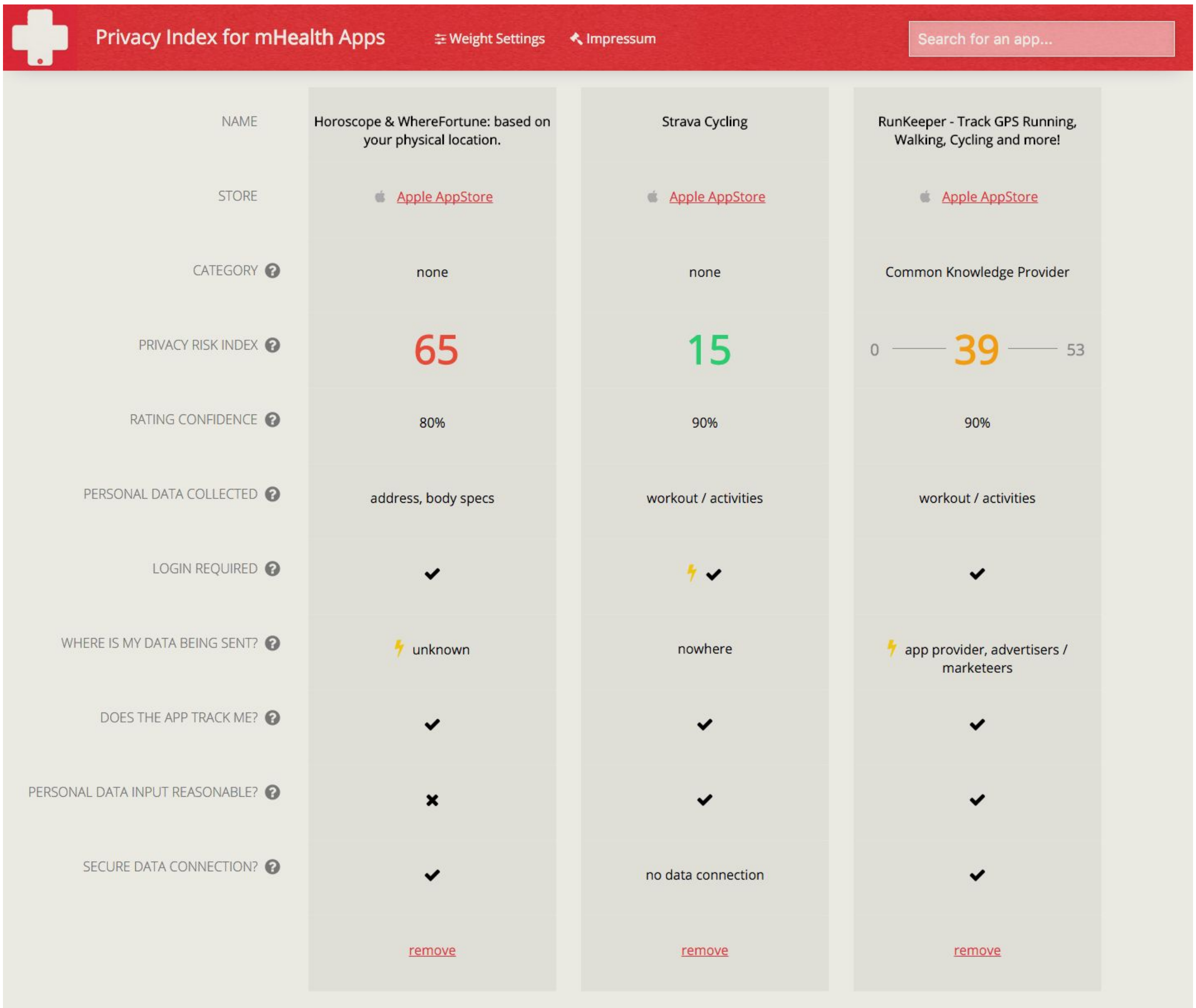
Table 4. The sources of information we used to rate each single app and their assigned weights

Source	Weight
Screenshots	0.8
Description	0.8
App downloaded	0.4
HTTP proxy	0.2
Privacy policy	0.1

Note: The sum of all weights does not add up to 1.0 since our algorithm picks the source with the lowest weight that is present at a given app survey as the resulting source weight.

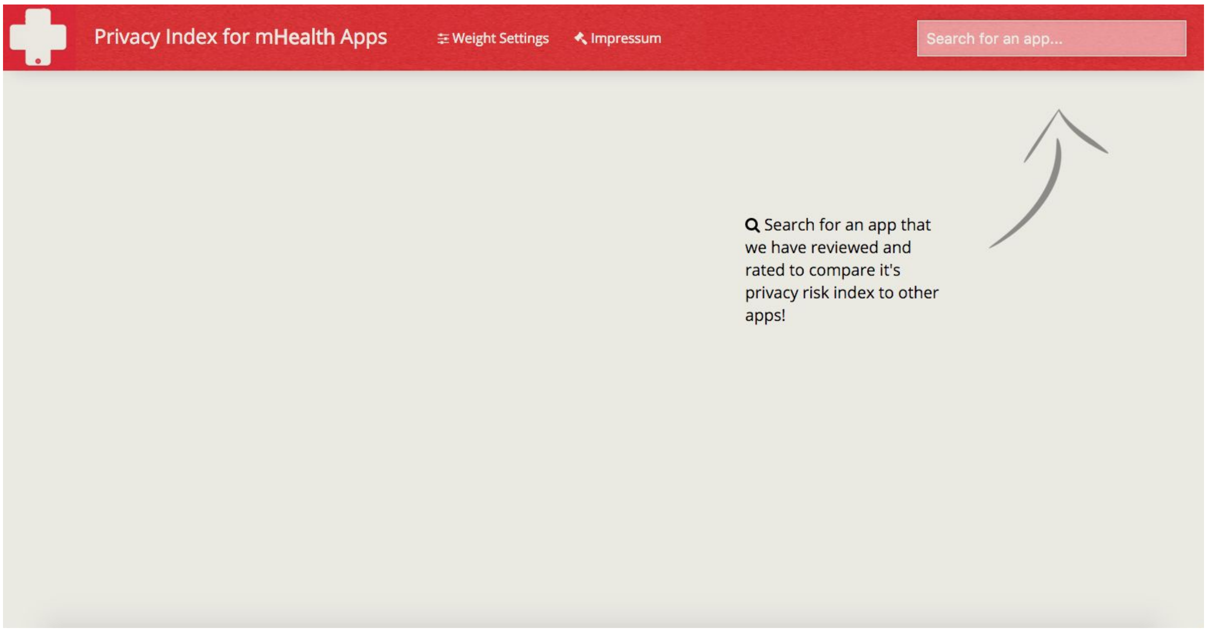
Appendix B

Figure 1. *Three apps have been selected and are listed in the comparison table view*



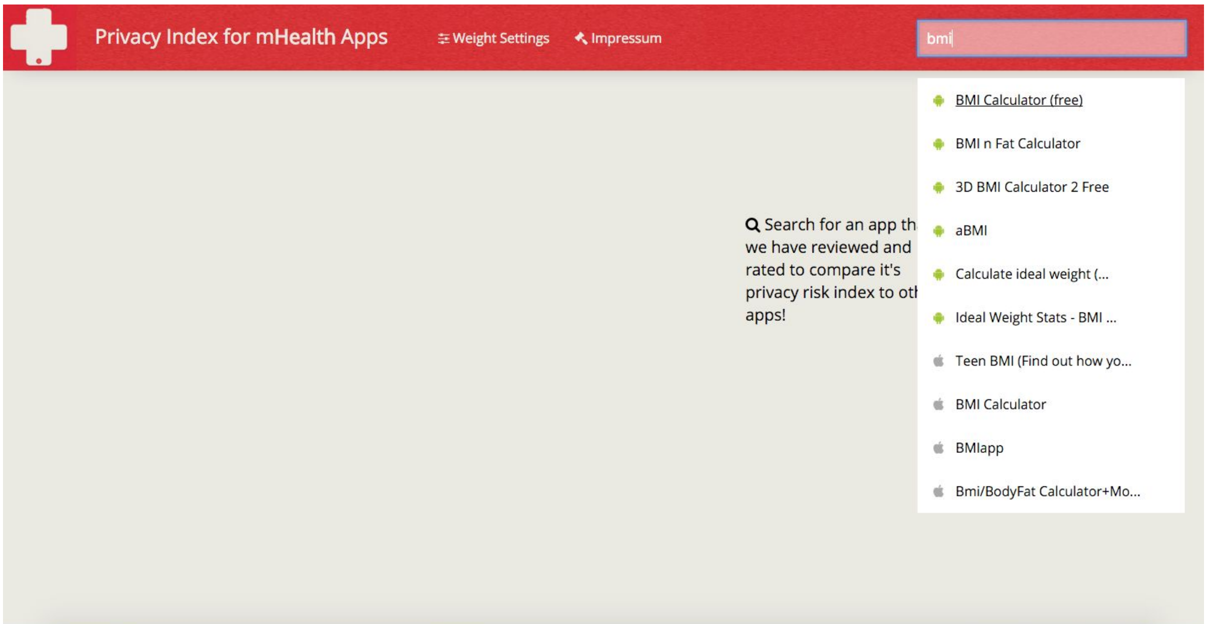
NAME	Horoscope & WhereFortune: based on your physical location.	Strava Cycling	RunKeeper - Track GPS Running, Walking, Cycling and more!
STORE	Apple AppStore	Apple AppStore	Apple AppStore
CATEGORY ?	none	none	Common Knowledge Provider
PRIVACY RISK INDEX ?	65	15	0 — 39 — 53
RATING CONFIDENCE ?	80%	90%	90%
PERSONAL DATA COLLECTED ?	address, body specs	workout / activities	workout / activities
LOGIN REQUIRED ?	✓	⚡ ✓	✓
WHERE IS MY DATA BEING SENT? ?	⚡ unknown	nowhere	⚡ app provider, advertisers / marketeers
DOES THE APP TRACK ME? ?	✓	✓	✓
PERSONAL DATA INPUT REASONABLE? ?	✗	✓	✓
SECURE DATA CONNECTION? ?	✓	no data connection	✓
	remove	remove	remove

Figure 2. Initial screen of developed GUI with promotion of first app search



Search for an app that we have reviewed and rated to compare it's privacy risk index to other apps!

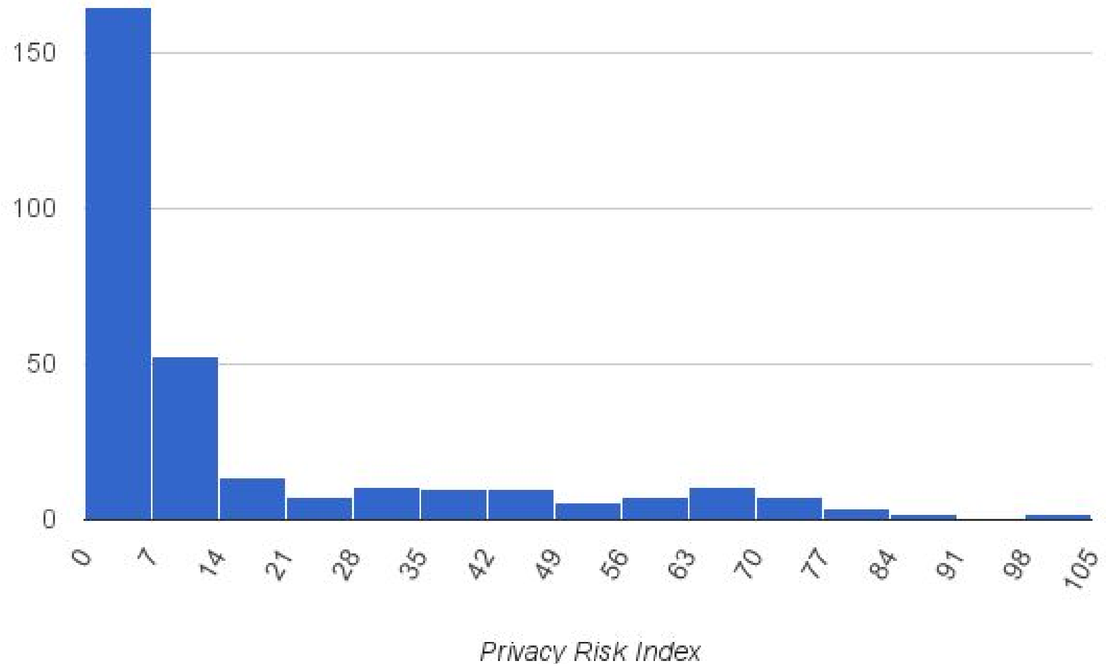
Figure 3. *Selection list after user typed in an app name to add to the comparison table view*



Search for an app th
we have reviewed and
rated to compare it's
privacy risk index to oth
apps!

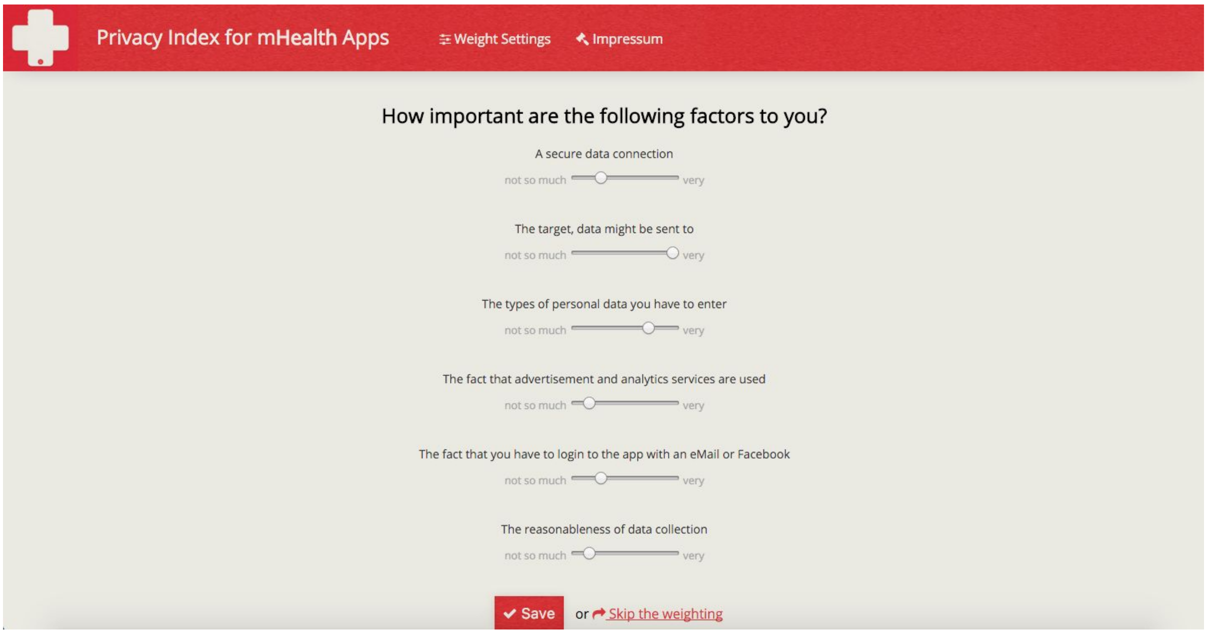
- BMI Calculator (free)
- BMI n Fat Calculator
- 3D BMI Calculator 2 Free
- aBMI
- Calculate ideal weight (...
- Ideal Weight Stats - BMI ...
- Teen BMI (Find out how yo...
- BMI Calculator
- BMIapp
- Bmi/BodyFat Calculator+Mo...

Figure 4. Histogram of privacy risk indices distribution



Note: On the x-axis we see the possible privacy risk indices multiplied by 100 and rounded to the nearest integer number for easier visibility.
On the y-axis we see the amount of apps in a certain privacy risk index range in increments of 7.

Figure 5. The weighting view where users can customize the privacy-risk-index factor weights to their own perception of weights



How important are the following factors to you?

A secure data connection
not so much — very

The target, data might be sent to
not so much — very

The types of personal data you have to enter
not so much — very

The fact that advertisement and analytics services are used
not so much — very

The fact that you have to login to the app with an eMail or Facebook
not so much — very

The reasonableness of data collection
not so much — very

Appendix C

Datasets

The initial dataset and our survey data as separate sheets
https://docs.google.com/spreadsheets/d/1zlF-aiR7qLRm1-qaCkLcCMTk-t-VXtj9R71MOlxT1Jo/edit?usp=sharing

The computed privacy risk indices for each app:
https://docs.google.com/spreadsheets/d/1canWMtrLDiADwyZw6Kh0C1XwN2mwoRBRv5BKQQPvbNs/edit?usp=sharing

Declaration of Good Scientific Conduct

German version

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

Hansen

Köln, ______________________ ______________________
Joel Hansen Thomas Brüggemann

English version

I hereby attest that I completed this work on my own and that I did not employ any tools other than those specified. All texts literally or semantically copied from other works are attributed with proper citations. This work has not been submitted in identical or similar form for any other exam, assessment, or assignment.

Hansen

Köln, ______________________ ______________________
Joel Hansen Thomas Brüggemann