Strengthen Digital Sovereignty of Smartphone Users: Evaluation Results of a Tailored Analysis Tool for App Behavior

Susen Döbelt[1] and Dominik Lange[2]

Abstract: A usable analysis tool that provides information on risky app behavior and offers options for action, can contribute to strengthen the digital sovereignty of smartphone app users. To this end, it should be tailored and meet the requirements of a human-centered design. Therefore, we conducted a lab test with $N = 38$ participants. They evaluated a prototype of our analysis tool in terms of its usability, transparency and potential to increase self-efficacy for data protection and privacy preservation. Furthermore, we investigated the effects of the tailoring by providing a congruent and an incongruent variant for behavioral stages. Both, usability and transparency evaluations differed positively from the average. Moreover, the interaction with the tool significantly increased the participants' self-efficacy and thus strengthened their digital sovereignty. Our tailoring of texts had a positive impact at least on the efficiency evaluation. This could be further developed by extended tailoring of e.g., the GUI.

Keywords: Usable Privacy, User Research, Human-Centered Design, Tailoring, Digital Sovereignty, Smartphone Apps.

1 Introduction

The right to informational self-determination [FLR17] also applies in the digital space, and not just since the introduction of the European GDPR. However, realizing and establishing digital sovereignty [Go17] is difficult for any individual in everyday life. Smartphones and apps, for example, are daily companions that on the one hand provide easily accessible information, but on the other hand also access users' data and pass it on to third parties. There is little transparency here, but this is a prerequisite for the valued preservation of one's privacy [TM15] [Ca09] and taking appropriate - often demanding - measures. Therefore, the goal of our research project was to create a usable tool that provides transparent information on risky app behavior and offers options for action to strengthen the digital sovereignty of smartphone app users. The purpose of this study was to review the human-centered and –tailored design of this tool.

[1] Professorship of Cognitive Psychology and Human Factors, Chemnitz University of Technology, Wilhelm-Raabe-Str. 43, 09120 Chemnitz, Germany, susen.doebelt@psychologie.tu-chemnitz.de
[2] Professorship of Cognitive Psychology and Human Factors, Chemnitz University of Technology, Wilhelm-Raabe-Str. 43, 09120 Chemnitz, Germany

doi:10.18420/OID2024_12

Several analysis methods that create transparency on the behavior of smartphone apps and provide information regarding the associated risk have been available for some years [En14]. Results can be requested by interested users from specific platforms (e.g., App-Checker [In22], Appicator [Fr24]) or provided via an app directly to their smartphone (e.g., Androlyzer [Te20]). Very few of these tools have been developed systematically in terms of user- or human-centeredness [In19], empirically evaluated concerning usability, user experience (UX), or even behavioral implications.

In the field of research, Gerber et al. [Ge18] investigated the effects of the app *FoxIT*, which entailed a static permission analysis method and feedback about the risk of installed apps. Results of a field trial showed that privacy awareness and knowledge of participants increased. Furthermore, changes in smartphone settings were reported. In a lab study with a provided smartphone, Bal et al. [BRH14] were able to show that their tool *Styx* based on dynamic monitoring of information flows via *TaintDroid* [En14], was perceived as user-friendly. The usage contributed to increasing user confidence due to the transparent information on app information flows. Van Kleek et al. [Kl17] suggested using Data Controller Indicators to disclose the transfer of information to third parties. The user evaluation revealed that more transparent information leads to decisions for apps with fewer organizations receiving data and could support users to make confident decisions.

In addition, Döbelt et al. [Dö20] were also able to highlight the importance of transparency and usability based on two user studies (online and lab) and five guidelines for tools that analyze app behavior. However, transparency played a specific role here on two dimensions: 1.) In terms of information delivery about apps, and 2.) In terms of the design of the tool itself: it should meet the requirements of transparency on its data handling [Dö20]. If tools are to be developed to further strengthen the digital sovereignty of users, additional options for action should be offered [Dö20]. Here, however, a 'one size fits all' solution is not sufficient [DH23] to encourage behavior change. Instead, a comprehensive alignment between the users and the tool may further increase its effectiveness [Li16]. Therefore, a tailored approach has been proposed [Kn15]. The *User-Tailored Privacy by Design* model [Wi17] suggests personalized nudges. In this context, the classification of behavioral stages and tailored interventions, like those for changing pro-environmental behavior [Ba13b], could serve as a template [DG21].

Based on this literature, we developed a human-centered prototype of an analysis tool with behavioral stage-specific tailoring. The tool was supposed to make the handling of app data accessible, visible and assessable. In addition to usability and transparency, we also investigated the impact on perceived self-efficacy. This describes the individual's assessment of being able to learn an action, even if they have not currently been able to master it [SJ02]. It is an important factor in predicting whether a person will consider or perform an action [Ba00]. Therefore, it presents an interesting variable to study the impact of our prototype on the digital sovereignty of smartphone app users.

3 Research questions and hypotheses

In our laby study, we focused on two research questions for our prototype: **(RQ1)** How usable and transparent is it evaluated? and **(RQ2)** Can it increase self-efficacy regarding data protection and privacy preservation of smartphone app users? Based on the presented content as well as the underlying user-centered design guidelines, we assume:

H1: The prototype is evaluated as above-average usable and transparent.

H2: Self-efficacy regarding data protection and privacy preservation is higher after interaction with the prototype than before.

In addition, a third exploratory question arose regarding the tailoring of the prototype: **(RQ3)** On which aspects of UX does the congruent tailoring to the behavioral stage have a positive impact?

4 Method

4.1 Study Design

Independent variables
For RQ1, the independent variable represents the use of our prototype; for RQ2, the time of measurement before (T1) and after (T2) the use. For RQ3, a distinction was made about the variant of the prototype. Based on the behavioral stage assessed during the recruitment, the congruent or incongruent variant was presented randomly and counterbalanced. This self-assessment of behavioral stage was adapted from the domain of pro-environmental behavior [Ba13b]. Participants who assigned themselves to the stage *predecision* or *preaction* were grouped into the early behavioral stage group, all others (*action* or *postaction*) into the late one.

Dependent variables
To evaluate our prototype, several established questionnaires were used: For testing H1, the System Usability Scale (SUS; [Br96] [RRR13]). It captures the usability of a system using 10 items answered via a 5-point rating scale ("*strongly disagree*"-"*strongly agree*"). Answers are aggregated to the SUS score (0-100), which allows results to be classified in grades from "*A+*" to "*F*" [LS18]. We chose 68.00 ("*C*") as a benchmark, which is at the center of the curved grading scale. Additionally, the SIPAS questionnaire [Sc21] was used to assess transparency. It measures the ability to perceive, understand and predict the information processing of a system [Sc21]. The 6 items are answered using a 6-point rating scale ("*not at all*"-"*completely*" true). The suggested unidimensional scale [SF22] is used here. Since the questionnaire was developed for a different field of application, the mean value (3.5) of the response scale serves as a benchmark.

was adapted to our application context (privacy preservation and data protection). The 10 items are answered on a 4-point rating scale (from "*not agree*" to "*agree exactly*").

Further, the User Experience Questionnaire (UEQ; [LHS08]) was used to investigate the exploratory RQ3. It captures UX by 6 subscales [Sc23]: Attractiveness (valence and overall impression), Perspicuity (easy to learn), Efficiency (without unnecessary effort), Dependability (control of interaction), Stimulation (exciting and motivating use), and Novelty (innovative and creative design). The UEQ contains 26 items and is designed as a 7-level semantic differential (from "*-3*" to "*+3*") with contrasting adjectives. The benchmarks referenced in [STH17] were used to interpret mean evaluations.

4.2 Sample

A total of N = 38 (26 female) subjects participated in the lab study. They were 24 years old on average (M = 23.95, SD = 5.03, min = 18.00, max = 41.00), and 71% reported a high school diploma as their highest educational qualification. The majority (92%) were third-semester students (M = 2.57, SD = 2.33) of psychology (69%). Participants rated themselves as significantly less tech-savvy (M = 3.76, SD = .93; t(37) = -2.51; p = .017, d = -0.41) than a comparable sample (N = 300; M = 4.14; [FAW19]). With regard to smartphone competence, the ratings (M = 3.82, SD = .70) were significantly (t(37) = 3.13, p = .003, d = 0.51) higher than those of a 2009 comparison sample (N = 460, M = 3.47; [Ka09]). This is only marginally true for their competence with apps (M = 3.64, SD = .53; t(37) = 1.96, p = .057, d = .32). The participants mainly (53%) used a smartphone with an Android operating system (OS) v12.

4.3 Procedure

The study started at the end of 2022 in Chemnitz, Germany, was conducted under pandemic-related restrictions, and approved in advance by the university's ethics committee. The procedure here is shortened and originally included additional tasks, variables, and qualitative data collection, which are omitted in favor of a reduced presentation.

Recruitment and selection of participants
The study announcement contained the rough goal (understand apps in more detail and evaluate a prototype) as well as the different remuneration options. It was distributed via various channels online and offline. Interested persons accessed the recruitment questionnaire via a QR code or link, implemented with *LimeSurvey* (v3.28.29+220920). After a brief welcome, respondents were first asked to generate an individual subject code for a pseudonymization of the data. This was followed by demographic information (age, gender, education, employment) and on the smartphone in use, i.e. OS and version. In the following, the behavioral stage was assessed (adapted from [Ba13a]; incl. n/a option). Lastly, contact details and compensation preferences were entered separately. To complete the questionnaire took a mean of 13 min (SD = 5.60). Participants selected

for the lab test used a smartphone with an Android OS > 6.0, to ensure seamless handling of the test smartphone. Furthermore, as of this version, runtime modification of permissions was possible. Individuals for whom behavioral stages were also available received an invitation.

Equipment and laboratory study
The laboratory room was equipped with two tables one for the participant and one for the experimenter. They were aligned at right angles to each other to avoid direct observation. Additional hardware (*ASUS* UX32V notebook, *LG* Flatron E2411 monitor, *Logitech* B110 mouse, *Cherry* G230 keyboard and the test smartphone (*Samsung Galaxy A33 5G*, OS v12) was placed on the participants' table to conduct the study.

After the room and equipment had been prepared in accordance with the hygiene policy, the participants were welcomed, signed the information on participation and data protection, and were presented with the experimental materials. They were informed that they could ask the experimenter questions at any time and that the app was not yet fully developed. The test started with the collection of self-descriptive variables (technology affinity, smartphone competence). Participants were then asked to rate their self-efficacy expectations in advance. They then had 10 minutes to freely explore the prototype. This was followed by two tasks of 5 minutes each: 1.) Inform yourself about the app *FitnessPro* with the help of the prototype; 2.) Take action using the prototype to minimize the risk of the app *FitnessPro*. Next, the participants were asked to fill out the evaluation questionnaires and the repeated self-efficacy assessments. In the end, the participants were remunerated. The test lasted a mean of 74 minutes (SD = 7.23).

4.4 Material: Prototype user interface

The prototype was implemented as smartphone app on the device it provides information about. The content was presented in German, but anonymized and translated here. The user interface (UI) was structured into three main areas: 1.) App risk, 2.) Device security, and 3.) Third-party providers. The examined area is described in Figure 1 and below (for more information see [Ch24]).

The start page of the app displayed a pie chart that included the number as well as the average risk score of the apps analyzed (only one for the lab test). The app list entries below were color- and numerically-coded according to their risk value. When selected, a drop-down at the very top of the subsequent screen explained this value. Below, the user could: a) Access analysis details: app permissions and third-party providers, and b) Take options for action: Change permissions, install alternatives or delete the app. The *Analyzed Permissions* page contained results from dynamic app analysis and highlighted whether a permission is currently withdrawn (green hand) or granted (orange exclamation mark). The *Analyzed Third-Party Providers* section entailed an overview of those providers to which data is forwarded other than the app provider. Here, also their risk value and the corresponding description were displayed. In the lower part *Options* installed from a store to be defined. The final option was to uninstall the app.

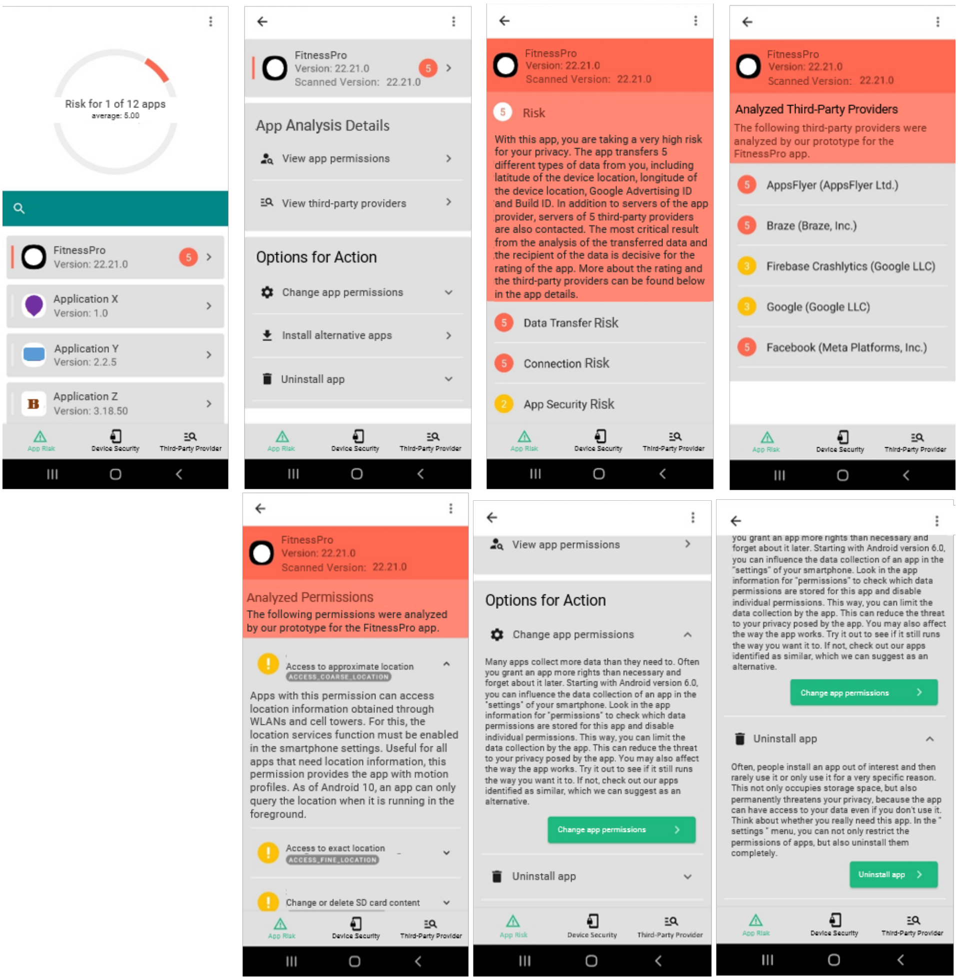

Fig. 1: UI of the prototypes risk area at the time of the laboratory test, texts here are translated into English and were originally presented in German, names and icons of apps have been anonymized

Prototype tailoring: As theoretical work (see section 2) implies that the adaption of a system to the user offers advantages, two prototype variants were implemented. These are tailored to the early or late behavioral stages. Different texts have been applied in the three areas under *Options for Action* to provide specific guidance to users. For the early behavioral stage, the texts began with a problem description to build problem awareness. The description of the options for action is also more detailed than that of the late ones. Additional feedback for behavior was added here instead.

5 Results

The raw data was first recoded according to the respective analysis instructions, and mean values or subscales were calculated. These are first reported descriptively and analyzed with respect to their distribution. Afterwards, a non- or parametric test was applied to examine the hypotheses.

5.1 **Usability and transparency** (RQ1)

Descriptive data revealed that the subjects rated the prototype with a mean SUS score of $M = 80.00$ ($SD = 13.90$). This corresponds to the grade "*A-*", i.e. a very good rating. The items assessing transparency were rated with "*somewhat agree*" ($M = 4.39$, $SD = 0.71$). The Shapiro-Wilk test for the SUS score showed a significant difference ($W(38) = .884$, $p < .001$) from the normal distribution. Therefore, we tested nonparametric and according to H1 one-tailed versus the mean SUS score of 68.00. The Wilcoxon test revealed a significantly higher SUS score with a high effect size compared to the selected benchmark ($W_s = 660.00$, $z = 4.02$, $p < .001$, $d = 1.86$). To further examine H1, the normal distributed transparency score ($W(38) = .946$, $p = .066$) was tested by a one-tailed t-test against the mean of the response scale (3.5). This also revealed a significant difference ($t(37) = 7.643$, $p < .001$, $d = 1.24$) with a large effect size.

5.2 **Self-efficacy** (RQ2)

Descriptive analysis of self-efficacy ratings revealed that at T1 the mean assessment was $M_{T1} = 2.38$ ($SD_{T1} = 0.54$, "*hardly true*") and at T2 $M_{T2} = 2.52$ ($SD_{T2} = 0.50$, "*somewhat true*"). At both times of measurement, the data was normally distributed ($W_{T1}(38) = .971$, $p_{T1} = .423$; $W_{T2}(38) = .984$, $p_{T2} = .852$). A one-tailed t-test for dependent samples was used to test the H2 and revealed a significant increase in self-efficacy ($t(37) = - 2.71$, $p = .005$, $d = - 0.44$; see Figure 2) with a small effect size.

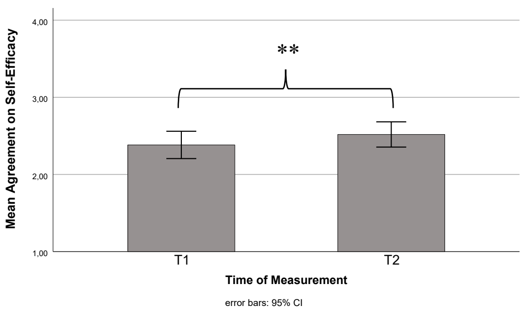

Fig. 2: Mean agreement of $N = 38$ participants for on self-efficacy, pre (T1) and post (T2) interaction with the prototype, ** marks a significant difference $p < .01$

The descriptive results of the UEQ (Table 1) showed that the mean ratings in the incongruent condition could be categorized largely as "*above average*" or "*good*", while in the congruent variant were mostly "*excellent*".

	Incongruent					Congruent				
UEQ Scale	*M*	*SD*	Min	Max	Category	*M*	*SD*	Min	Max	Category
Attr	1.41	0.78	-0.83	3.00	Above Average	1.57	0.67	0.50	2.83	Good
Persp	1.70	1.22	-2.25	3.00	Good	2.05	0.75	-0.25	3.00	Excellent
Eff	1.47	0.63	-0.25	2.25	Good	1.97	0.58	1.00	3.00	Excellent
Dep	1.78	0.56	0.75	2.75	Excellent	1.82	0.56	0.50	2.75	Excellent
Stim	0.96	0.92	-0.75	2.50	Below Average	1.29	0.55	0.00	2.25	Above Average
Nov	0.95	0.90	-1.50	2.50	Above Average	1.20	0.89	-0.50	2.75	Good

Tab. 1: UEQ evaluation for incongruent (n = 19) and congruent (n = 19) prototype variant, Attr = attractiveness, Persp = perspicuity, Eff = efficiency, Dep = dependability, Stim = stimulation, Nov = novelty; benchmark category for mean evaluation follows [STH17]

The normal distribution was violated for perspicuity ($W_{incon}(19) = .819$, $p_{incon} = .002$; $W_{con}(19) = .836$, $p_{con} = .004$) and was therefore analyzed nonparametrically. A one-tailed t-test for independent samples was calculated for all other subscales and a Bonferroni correction (p = .010) was applied. Only one statistically significant difference with a strong effect size emerged for the efficiency ($t(36) = 2.543$, $p = .008$, $d = 0.83$; Figure 3).

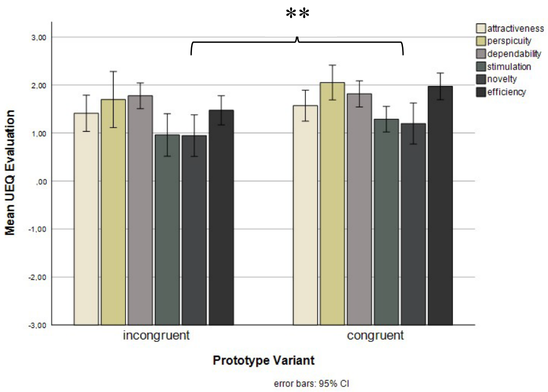

Fig. 3: Mean UEQ evaluations for subscales in the incongruent (n = 19) and congruent (n = 19) prototype condition, ** marks a significant difference $p < .01$

6 Discussion and limitations

The results of our laboratory test showed that our prototype was appreciated for its usability and transparency (H1 confirmed). Thus, the human-centered development based on specific design guidelines resulted in a positive evaluation of the tool. However, the evaluation could be positively biased, due to the notice that the app was a prototype developed as part of a research project. The participants may therefore have rated our app more leniently than they would have done with mature products. Additionally, the first positive impression might wear off long term. Least is opposed by the fact that usability evaluations seem to increase over time [Ku11]. In addition, the interaction with our prototype was predetermined by the tasks set. Thus, all participants got to know it equally and comprehensively. It remains unclear whether all the information would have been noticed in a self-determined interaction.

The H2 was also confirmed, as interaction with our prototype significantly increased participants' self-efficacy. Thus, our tool can contribute to strengthening users' digital sovereignty. Even though the prototype's objective was not described in advance, it was revealed during the lab test. A positive biasing of the evaluation due to social desirability is also possible here. In order to address this, it was pointed out that there are no right or wrong answers.

For our third exploratory question, the results show that our tailoring to the behavioral stages had a positive effect on the evaluation of efficiency. Our tool was perceived as more pragmatic and less cluttered in the congruent condition. However, differences remained limited to this aspect, probably because the prototype variants differed only in the texts. Different graphical designs might have been more crucial for further aspects.

The limitations of our lab study relate mainly to its external validity and transferability to real world settings. We provided a prepared smartphone to our participating sample of students, which differ in many respects from the overall population [An16]. Therefore, some results may be over- or underrepresented. Regarding the internal validity of the study, it can be noted that established instruments were slightly reformulated to adapt them to the present context. These changes could have an impact on construct validity, which cannot be followed up here.

6.1 Future work

Our tool will be further developed during the research project [Ch24]. The scope of smartphone security in particular will be expanded and evaluated again. This will include a follow-up survey examining the impact on behavioral change beyond the lab. Further research is also needed on the tailoring of analysis tools like ours. There are still many possibilities in the design of the graphical user interface (GUI) of information and options for action that can contribute, e.g., to increase awareness of the problem and the motivation to exercise digital sovereignty.

The usability of our prototype was rated as "*very good*" and its transparency was "*somewhat approved*". Both evaluations differ statistically significantly and positively from the average. Our prototype can therefore be considered as usable and transparent (RQ1). Additionally, it was encouraging that the interaction with our prototype significantly increased the self-efficacy of the participants (RQ2). Our prototype thus can contribute to the purpose of strengthening the digital sovereignty of smartphone app users. In addition, we found that our tailoring for the behavioral stages had a positive impact on the efficiency evaluation (RQ3). The behaviorally congruent app variant thus contributes to decreasing at least the resource investment of app users interested in protecting their privacy. This first result could be further developed by extended tailoring of other e.g., GUI elements.

Acknowledgments

This study was conducted as a part of the research project "*PANDERAM*" (funding code: 16SV8521, [Ch24]), funded by the German Federal Ministry of Education and Research. We thank our project partners Fraunhofer SIT, the Institut für Technik und Journalismus e.V., the DAI-Laboratory of TU Berlin and secuvera GmbH for their support in the prototype development. Furthermore, we want to thank Anna Graue, who was in charge of the experiments in the lab and supported the analysis of the data.

Bibliography

[An16] Andone, I. et.al.: How age and gender affect smartphone usage. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: adjunct, pp. 9-12. 2016.

[BRH14] Bal, G.; Rannenberg, K.; Hong, J.: Styx: Design and Evaluation of a New Privacy Risk Communication Method for Smartphones. In (Cuppens-Boulahia, N., Cuppens, F., Jajodia, S., El Kalam, A. A., Sans, T. eds.): ICT Systems Security and Privacy Protection. IFIP Advances in Information and Communication Technology, 428. Berlin, Heidelberg: Springer. 2014.

[Ba13a] Bamberg, S.: Applying the stage model of self-regulated behavioral change in a car use reduction intervention. Journal of Environmental Psychology, 33, 68–75. 2013.

[Ba13b] Bamberg, S.: Changing environmentally harmful behaviors: A stage model of self-regulated behavioral change. Journal of Environmental Psychology, 34, 151–159. 2013.

[Ba00] Bandura, A.: Social Cognitive Theory: An Agentic Perspective. Annual Review of Psychology, 52, pp. 1–26. 2000.

[Br96] Brooke, J.: SUS: A "Quick and Dirty" Usability Scale. Usability Evaluation. In Industry, 189(194), pp. 4-7. 1996.

[Ca09] Cavoukian, A.: Privacy by design: The 7 foundational principles. Information and

Privacy Commissioner of Ontario, Canada, 5. 2009.

[Ch24] TU Chemnitz, Professorship Cognitive Psychology and Human Factors: PANDERAM: https://www.tu-chemnitz.de/hsw/psychologie/professuren/allpsy1/forschung/panderam/index.html.en, accessed: 15/01/2024.

[Dö20] Döbelt, S. et.al.: Clearing the Hurdles: How to Design Privacy Nudges for Mobile Application Users. In (Moallem, A., ed.): Proceedings of the Second International Conference on HCI for Cybersecurity, Privacy and Trust, Held as Part of the 22nd HCI International Conference, Copenhagen, Denmark, pp. 326-353, Springer, Cham. 2020.

[DG21] Döbelt, S.; Günther, M.: Two Values Work Alike: Linking Proenvironmental and Privacy Preserving Behavior. In (Huckauf, A., Baumann, M., Ernst, M., Herbert, C., Kiefer, M., Sauter, M., eds.): TeaP 2021. Lengerich: Pabst Science Publishers. 2021.

[DH23] Döbelt, S.; Halama, J.: Clean up my Phone: Field Trial Results of a Privacy Tool Increasing Transparency about App Behavior. International Symposium on Technikpsychologie, Darmstadt, Germany. 2023.

[En14] Enck, W. et.al.: TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. ACM Transactions on Computer Systems, 32(2). 2014.

[FAW19] Franke, T.; Attig, C.; Wessel, D.: A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. International Journal of Human–Computer Interaction, 35(6), pp. 456-467. 2019.

[Fr24] Fraunhofer SIT. Appicator – Framework for App Securtiy Tests. https://www.sit.fraunhofer.de/en/appicaptor/?MP=iwllzb, accessed: 15/01/2024.

[FLR17] Friedewald, M.; Lamla, J.; Roßnagel, A.: Informationelle Selbstbestimmung im digitalen Wandel, Springer Vieweg, ISBN 978-3-658-17661-7. 2017.

[Ge18] Gerber, N. et.al.: FoxIT: enhancing mobile users' privacy behavior by increasing knowledge and awareness. In: Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust, pp. 53-63. ACM. 2018.

[Go17] Goldacker, G.: Digitale Souveränität. https://publica-rest.fraunhofer.de/server/api/core/bitstreams/71c726ab-133b-4cd9-8e4c-3d259453fcf8/content. 2017. accessed: 15/01/2024.

[In19] International Organization for Standardization: Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. (ISO 9241-210:2019). https://www.iso.org/standard/77520.html. 2019. accessed: 15/01/2024.

[In22] Institut für Technik und Journalismus e.V.: Mobil Sicher – App Checker. https://appcheck.mobilsicher.de/. 2022. accessed: 15/01/2024

[Ka09] Karrer, K. et.al.: Technikaffinität erfassen–der Fragebogen TA-EG. Der Mensch im Mittelpunkt technischer Systeme, 8, pp. 196-201. 2009.

[Kl17] Kleek, M. v. et.al.: Better the devil you know: Exposing the data sharing practices of smartphone apps. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 5208-5220. ACM. 2017.

[Kn15] Knijnenburg, B. P.: A user-tailored approach to privacy decision support, Ph.D. Thesis, University of California, Irvine, CA. 2015.

[LHS08] Laugwitz, B.; Held, T.; Schrepp, M.: Construction and Evaluation of a User Experience Questionnaire. In (Holzinger, A., ed.): HCI and Usability for Education and Work, pp. 63–76. Springer. 2008.

[LS18] Lewis, J. R.; Sauro, J.: Item benchmarks for the system usability scale. Journal of Usability Studies, 13(3). 2018.

[Li16] Liu, B. et.al.: Follow my recommendations: A personalized privacy assistant for mobile app permissions. In: SOUPS'16 Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security, pp. 27-41. 2016.

[RRR13] Rummel, B.; Ruegenhagen, E.; Reinhardt, W.: System Usability Scale (Translated into German). https://blogs.sap.com/2016/02/01/system-usability-scale-jetzt-auch-auf-deutsch/. 2013. accessed: 15/01/2024.

[Sc23] Schrepp, M.: User ExperiencevQuestionnaire Handbook. https://www.ueq-online.org/Material/Handbook.pdf. 2023. Accessed: 05/05/2024

[STH17] Schrepp, M.; Thomaschewski, J.; Hinderks, A.: Construction of a Benchmark for the User Experience Questionnaire (UEQ). In: International Journal of Interactive Multimedia and Artificial Intelligence, 4(4), 40. Universidad Internacional de La Rioja. 2017.

[Sc21] Schrills, T. et.al.: Are Users in the Loop? Development of the Subjective Information Processing Awareness Scale to Assess XAI. 2021.

[SF22] Schrills, T.; Franke, T.: How do Users Experience Traceability of AI Systems? Examining Subjective Information Processing Awareness in Automated Insulin Delivery Systems. 2022.

[SJ02] Schwarzer, R.; Jerusalem, M.: Das Konzept der Selbstwirksamkeit. In (Jerusalem, M., Hopf, D., eds.): Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, pp. 28–53. Beltz. https://www.pedocs.de/frontdoor.php?source_opus=3930. 2002. accessed: 15/01/2024.

[SJ03] Schwarzer, R.; Jerusalem, M.: SWE. Skala zur Allgemeinen Selbstwirksamkeitserwartung. In (Leibniz-Institut für Psychologie, ed.): Open Test Archive. Trier: ZPID. 2003.

[Te20] TU Berlin, DAI-Labor.: Androlyzer – Privacy Scanner for Android: https://download.cnet.com/Androlyzer-Privacy-Scanner/3000-20432_4-78524108.html. 2020, accessed: 15/01/2024.

[TM15] Trepte, S.; Masur, P. K.: Privatheit im Wandel. Eine repräsentative Umfrage zur Wahrnehmung und Beurteilung von Privatheit. Stuttgart: Universität Hohenheim. https://www.uni-hohenheim.de/fileadmin/einrichtungen/psych/Team_MP/Berichte/Bericht_-_Privatheit_im_Wandel_2014-06-18.pdf. 2015.

[Wi17] Wilkinson, D. et.al.: Work in Progress: User-Tailored Privacy by Design. Paper presented at 24th Annual Network and Distributed System Security Symposium, San Diego. 2017.