Platform Privacy - Research for a
self-determined life in the digital world

Plattform Privatheit

Michael Friedewald & Murat Karaboga (Eds.)

FREEDOM IN DIGITAL INFRASTRUCTURES

Poster Proceedings

Imprint

Research Papers of the Platform Privacy, No. 4

Editors

Michael Friedewald, Murat Karaboga
Fraunhofer Institute for Systems and Innovation Research ISI

Series

ISSN (Print)	2942-8874
ISSN (Online)	2942-8882
DOI	https://doi.org/10.24406/publica-3685

Publication

October 2024, 1. Edition
Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe

Suggested citation

Friedewald, M. & Karaboga, M. (Eds.): Freedom in digital infrastructures. Poster-Proceedings. Research Papers of the Platform Privacy, No. 4. Karlsruhe: Fraunhofer ISI, 2024. https://doi.org/10.24406/publica-3685

Table of contents

1 Preface

The 9th interdisciplinary annual conference of the Privacy Platform, held on October 17 and 18, 2024, at Villa Elisabeth in Berlin, has become a leading conference on privacy, data protection, and self-determined life in the digital world for the German-speaking research community. It serves as a vital forum for exploring the interplay between digital infrastructures and the safeguarding of individual and collective freedoms. The conference aims to highlight the importance of informed consent in healthcare research, raise awareness about privacy issues through engaging methods, and examine the implications of smart technologies for citizens.

The Privacy Platform connects interdisciplinary scientific projects funded by the BMBF under the initiative "Privacy Platform – Supporting Citizens in Exercising the Fundamental Right to Informational Self-Determination." It plays a crucial role in fostering dialogue among academia, civil society, politics, and industry. By generating knowledge on ethical, legal, and social aspects of privacy and data protection, the platform empowers citizens in their digital interactions.

The proceedings include three contributions that address key themes in privacy and data protection. These papers explore how technical solutions can enhance consent mechanisms in healthcare, the effectiveness of serious games for privacy awareness, and the societal implications of smart home and smart city technologies. Collectively, they represent an important step toward strengthening individual rights and promoting informed engagement in the digital age.

2 Toward Technically Enforceable Consent in Healthcare Research

Johannes Lohmöller, Jan Pennekamp und Klaus Wehrle[1]

2.1 Motivation

Digital health technology has, over the past years, become an integral part of healthcare research, for instance, to gain insight into clinical decision-making (Wang et al., 2022) and its impact on patient trajectories or for large-scale studies, e.g., in cardiovascular research (Denaxas & Morley, 2015). The increasing digitization of healthcare facilities has led to a broad availability of electronic health record (EHR) data and a recent political and public surge to utilize this data. The COVID-19 pandemic further accelerated this trend (Dron et al., 2022). Ethics committees are starting to acknowledge the potential for so-called *secondary use* in medical research, hinting at a broad utilization of EHR data in the (near) future. Currently, there are ongoing state-level initiatives, e.g., in Germany, to centralize EHR data for research purposes (Rau et al., 2024).

Since protecting private medical data is a crucial personal right, a fundamental principle in medical research is giving consent, i.e., individual data subjects agree to a specific use of their data, similar to how they would agree to participate in a clinical trial. For secondary data use, however, the individuals subject to the data can hardly give explicit consent, e.g., as specific research questions have yet to be formulated at the time of data collection. Likewise, individuals often are not reachable anymore once research questions have been fixed for data collected in the past. Here, state-of-the-art practices are general consent or broad consent forms, although these are criticized for being too vague and not specific enough (Barazzetti et al., 2020). Generally, data subjects are willing to consent to secondary use of their data, as highlighted by a recent meta-review (Baines et al., 2024). However, the key to such consent is that individuals retain control over their data and that benefits are clear to them. We thus argue that technical means to enforce consent in healthcare research are highly beneficial. To this end, we investigate such technical means to implement consent in healthcare research that render central data collection and collecting broad consent in advance, as currently discussed (Rau et al., 2024), redundant.

Our work complements distributed data analysis tools, including MedCo (Froelicher, 2020; Raisaro et al., 2019), UnLynx (Froelicher et al., 2017), PCORnet (Yuan et al., 2017), or PHT (Mou et al., 2023), showing the feasibility and need for privacy-preserving decentralized analysis of health records and medical research data. None of them, however, incorporates consent on a research project or even query level. Likewise, related data ecosystems fail to reliably and transparently provide (technical) guarantees to manage patients' consent decisions (Geisler et al., 2022; Lohmöller et al., 2024). In this work, we thus sketch a system design for consent-aware distributed data analysis. Our goal is to provide technical enforcement of consent on a per-query basis while still allowing for large-scale studies. Thereby, we aim to empower users to reliably, transparently, and privacy preservingly handle their consent affinity.

2.2 Sketching Technical Consent Enforcement

Figure 1 introduces the high-level protocol flow, as described in the following. We consider data-holding institutions (e.g., clinics) and data subjects (e.g., patients) each (part-time) active parties in the system such that researchers can only access the data if both parties agree. Thereby, data

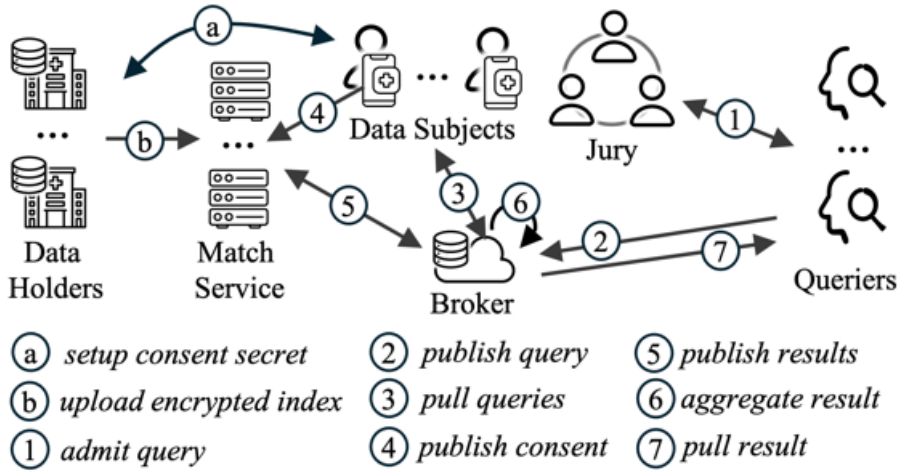
[1] RWTH Aachen University, Aachen, Germany, {lohmoeller,pennekamp,wehrle}@comsys.rwth-aachen.de

subjects remain in control over their data and give consent on a per-query basis. To enforce consent cryptographically, we utilize a public-key searchable encryption scheme (PEKS) that allows defining trapdoors to be matched against an encrypted distributed index (Boneh et al., 2004; Froelicher et al., 2017).

We propose to store the encrypted index under the control of data providers (e.g., hospitals), let researchers submit queries (translated to trapdoors) to the data providers, and cryptographically involve the data subjects in the trapdoor evaluation process, which prohibits query evaluation on data without the data subject's consent. Specifically, we include data subjects in the process by distributing a share of the evaluation key material to the data subject out of band, such as when the data is initially collected. This key share is cryptographically required to evaluate trapdoors, which implies that the data subject needs to actively contribute her share, i.e., consent to query evaluation.

As data subjects can hardly assess whether a specific query is in their interest or beneficial to them, we require an independent jury to review the queries and a short, to a non-expert audience, understandable summary of the research goals. Based on this summary, data subjects can make an informed decision about contributing their data. We employ threshold cryptography to ensure that the jury must agree on a joint query admission decision before a query can be executed. This admission ensures that queries are not directly harmful, such as being exploited for tracing specific individuals. Additionally, we aggregate results from multiple data providers before relaying them to the researchers via the broker shown in Figure 1. This aggregation unlinks the data from any individual and their physical origin, thus preserving the data subjects' anonymity.

Figure 1: Schematic protocol flow enabling technically enforced consent.

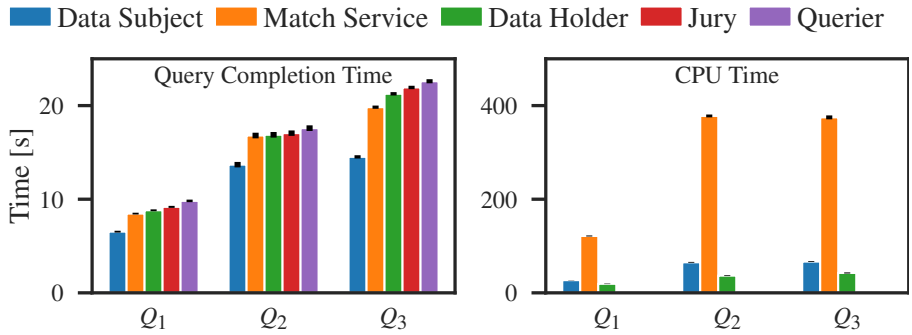


Queries and data are encrypted with a jury-managed public-private key pair. Steps a-b are one-time operations, Steps 1-7 execute per query. Non-collusion is required between data holders and the test servers. Besides, one jury member must be benign.

We evaluate our approach in a simulation study, i.e., we simulate a state-level network composed of German university hospitals. With the HCUP national inpatient sample (Khera et al., 2017), we employ a large dataset of ICD-10 coded diagnoses and evaluate the same real-world queries within the setting as related work before (Yuan et al., 2017). The (one-time) setup for generating keying material and encrypting individual data items consumes 290 ms per record in the dataset, which we consider feasible for practical deployment on a large scale. More interestingly, Figure 2 touches upon the query results of our simulated real-world evaluation, showing that in an idealized scenario (all entities are online and respond immediately), our three queries Q_1(Chung et al., 2015), Q_2 (Habermann et al., 2014), Q_3 (George et al., 2014), which, e.g., analyze the clinical treatment and outcome of a cancer subtype, all execute within 22s. In this setting, the most time-consuming part is matching trapdoors against the encrypted index, followed by the giving consent operation of the

data subjects. However, the latter will be spread among up to 6.3 million data subjects in the network, thus incurring a negligible overhead per individual. Practically, we expect the query runtimes to be dominated by user interaction, e.g., after notification via push messaging or regular digests of incoming requests. A smartphone app holding the data subject's keys would be a convenient way of interacting with data subjects. Other options, including delegating consent management to a general physician or relatives, are technically possible and can complement these efforts. Overall, these results show that the suggested approach scales to real-world scenarios.

Figure 2: Query performance for three realistic queries on 6.3 million records.



Left: time passed between retrieval and completion. Right: consumed CPU time (sum over all parallel computations). By subsampling our dataset, we find that the computationally intensive operations scale linearly with the number of records in the dataset.

2.3 Discussion and the Road Ahead

A frequent concern is that the self-determination and freedom of giving consent, as well as the duty to decide, might lead to consent fatigue. Then, users are overwhelmed by the number of consent requests, similar to Cookie consent banners (Kretschmer et al., 2021). One way to mitigate this issue would be to deploy a local consent agent, e.g., as part of the aforementioned consent management app, that accepts queries based on user-defined rules and criteria the jury has reviewed. Compared to the broad consent practice, data subjects would still be free to review and reconsider their choices (Matzutt et al., 2017). Besides, such an agent would keep the subject's sovereignty high while allowing transparency and accountability.

Compared to a fully centralized system, our proposal at first seems to limit the query experience: While the formulated trapdoors support querying specific ranges and logical formulas for combining multiple matching criteria, a distributed approach that requires human interaction can hardly compete with the performance and interactiveness of a central database. However, efforts to implement centralized EHR databases at the state level or beyond suffer from various issues, including ethical and legal concerns, among others (Baines et al., 2024). Here, data governance strategies are scarce: They need to (1) comply with regulatory frameworks, (2) be perceived as trustworthy by the public, and (3) must not become an overly complicated access barrier for researchers (Rau et al., 2024). We argue that collecting consent on a per-query basis shifts the burden of finding a universal governance strategy to multiple individual choices with clear consequences and limited scope, thereby reducing decision and governance complexity. With this work, we thus sketch a technical basis for this paradigm shift and call for future work to study the data subject's perception of privacy and the willingness to provide their data in more detail.

Besides analyzing implementing the consent agent and studying individual data subjects' perception of privacy, future work should investigate the system's usability with real users, e.g., as part of a clinical trial, and assess its impact on the patient's overall willingness to participate in secondary data use. From a technical perspective, we are confident that the system can be deployed in a real-world scenario, as the evaluation shows good scalability regarding participating entities and data volume.

2.4 Conclusion

Our work complements existing decentralized data analysis tools with enforceable consent, as users are cryptographically involved in the query evaluation. Thereby, it enables data subjects to decide participation freely and on a per-query basis while still allowing for large-scale studies. Thus, our work empowers users to reliably and transparently handle their consent affinity. Our evaluation shows that the computational overhead is reasonable and that the system scales to the demands of real-world scenarios. This work thus contributes a cryptographic option for giving consent to the ongoing discussion on how to open up healthcare data for research (Baines et al., 2024).

Acknowledgments

This research was funded within the VeSiTRUST project and supported under grant no. 02J24A030 by funds of the German Federal Ministry of Education and Research (BMBF). The authors thank Michael Herwig for his initial work on distributed searchable encryption.

Bibliography

Baines, R., Stevens, S., Austin, D., Anil, K., Bradwell, H., Cooper, L., Maramba, I. D., Chatterjee, A., & Leigh, S. (2024). Patient and Public Willingness to Share Personal Health Data for Third-Party or Secondary Uses: Systematic Review. Journal of Medical Internet Research, 26, e50421. https://doi.org/10.2196/50421

Barazzetti, G., Bosisio, F., Koutaissoff, D., & Spencer, B. (2020). Broad consent in practice: Lessons learned from a hospital-based biobank for prospective research on genomic and medical data. European Journal of Human Genetics, 28(7), 915–924. https://doi.org/10.1038/s41431-020-0585-0

Boneh, D., Di Crescenzo, G., Ostrovsky, R., & Persiano, G. (2004). Public Key Encryption with Keyword Search. In C. Cachin & J. L. Camenisch (Eds.), Advances in Cryptology—EUROCRYPT 2004 (pp. 506–522). Springer. https://doi.org/10.1007/978-3-540-24676-3_30

Chung, T. K., Rosenthal, E. L., Magnuson, J. S., & Carroll, W. R. (2015). Transoral robotic surgery for oropharyngeal and tongue cancer in the U nited S tates. The Laryngoscope, 125(1), 140–145. https://doi.org/10.1002/lary.24870

Denaxas, S. C., & Morley, K. I. (2015). Big biomedical data and cardiovascular disease research: Opportunities and challenges. European Heart Journal - Quality of Care and Clinical Outcomes, 1(1), 9–16. https://doi.org/10.1093/ehjqcco/qcv005

Dron, L., Kalatharan, V., Gupta, A., Haggstrom, J., Zariffa, N., Morris, A. D., Arora, P., & Park, J. (2022). Data capture and sharing in the COVID-19 pandemic: A cause for concern. The Lancet Digital Health, 4(10), e748–e756. https://doi.org/10.1016/S2589-7500(22)00147-9

Froelicher, D. (2020). MedCo2: Privacy-Preserving Cohort Exploration and Analysis. Digital Personalized Health and Medicine. https://doi.org/10.3233/SHTI200174

Froelicher, D., Egger, P., Sousa, J. S., Raisaro, J. L., Huang, Z., Mouchet, C., Ford, B., & Hubaux, J.-P. (2017). UnLynx: A Decentralized System for Privacy-Conscious Data Sharing. Proceedings on Privacy Enhancing Technologies. https://petsymposium.org/popets/2017/popets-2017-0047.php

Geisler, S., Vidal, M.-E., Cappiello, C., Lóscio, B. F., Gal, A., Jarke, M., Lenzerini, M., Missier, P., Otto, B., Paja, E., Pernici, B., & Rehof, J. (2022). Knowledge-Driven Data Ecosystems Toward Data

Transparency. Journal of Data and Information Quality, 14(1), 1–12. https://doi.org/10.1145/3467022

George, E. M., Tergas, A. I., Ananth, C. V., Burke, W. M., Lewin, S. N., Prendergast, E., Neugut, A. I., Hershman, D. L., & Wright, J. D. (2014). Safety and Tolerance of Radical Hysterectomy for Cervical Cancer in the Elderly. Gynecologic Oncology, 134(1), 36–41. https://doi.org/10.1016/j.ygyno.2014.04.010

Habermann, E. B., Thomsen, K. M., Hieken, T. J., & Boughey, J. C. (2014). Impact of Availability of Immediate Breast Reconstruction on Bilateral Mastectomy Rates for Breast Cancer across the United States: Data from the Nationwide Inpatient Sample. Annals of Surgical Oncology, 21(10), 3290–3296. https://doi.org/10.1245/s10434-014-3924-y

Khera, R., Angraal, S., Couch, T., Welsh, J. W., Nallamothu, B. K., Girotra, S., Chan, P. S., & Krumholz, H. M. (2017). Adherence to Methodological Standards in Research Using the National Inpatient Sample. JAMA, 318(20), 2011. https://doi.org/10.1001/jama.2017.17653

Kretschmer, M., Pennekamp, J., & Wehrle, K. (2021). Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web. ACM Transactions on the Web, 15(4), 1–42. https://doi.org/10.1145/3466722

Lohmöller, J., Pennekamp, J., Matzutt, R., Schneider, C. V., Vlad, E., Trautwein, C., & Wehrle, K. (2024). The unresolved need for dependable guarantees on security, sovereignty, and trust in data ecosystems. Data & Knowledge Engineering, 102301. https://doi.org/10.1016/j.datak.2024.102301

Matzutt, R., Müllmann, D., Zeissig, E.-M., Horst, C., Kasugai, K., Lidynia, S., Wieninger, S., Ziegeldorf, J. H., Gudergan, G., gen. Döhmann, I. S., Wehrle, K., & Ziefle, M. (2017). myneData: Towards a Trusted and User-controlled Ecosystem for Sharing Personal Data. https://doi.org/10.18420/IN2017_109

Mou, Y., Li, F., Weber, S., Haneef, S., Meine, H., Caldeira, L., Jaberansary, M., Welten, S., Yediel Ucer, Y., Prause, G., Decker, S., Beyan, O., & Kirsten, T. (2023). Distributed Privacy-Preserving Data Analysis in NFDI4Health With the Personal Health Train. Proceedings of the Conference on Research Data Infrastructure, 1. https://doi.org/10.52825/cordi.v1i.282

Raisaro, J. L., Troncoso-Pastoriza, J. R., Misbach, M., Sousa, J. S., Pradervand, S., Missiaglia, E., Michielin, O., Ford, B., & Hubaux, J.-P. (2019). MedCo: Enabling Secure and Privacy-Preserving Exploration of Distributed Clinical and Genomic Data. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 16(4), 1328–1341. https://doi.org/10.1109/TCBB.2018.2854776

Rau, E., Tischendorf, T., & Mitzscherlich, B. (2024). Implementation of the electronic health record in the German healthcare system: An assessment of the current status and future development perspectives considering the potentials of health data utilisation by representatives of different stakeholder groups. Frontiers in Health Services, 4, 1370759. https://doi.org/10.3389/frhs.2024.1370759

Wang, J., Yang, L., Huang, X., & Li, J. (2022). Annotating Free-Texts in EHRs Towards a Reusable and Machine-Actionable Health Data Resource. In P. Otero, P. Scott, S. Z. Martin, & E. Huesing (Eds.), Studies in Health Technology and Informatics. IOS Press. https://doi.org/10.3233/SHTI220239

Yuan, J., Malin, B., Modave, F., Guo, Y., Hogan, W. R., Shenkman, E., & Bian, J. (2017). Towards a privacy preserving cohort discovery framework for clinical research networks. Journal of Biomedical Informatics, 66, 42–51. https://doi.org/10.1016/j.jbi.2016.12.008

3 Privacy Awareness Cards - A Serious Game for Privacy Awareness

Tom Lorenz, Michael Pleger, Tanja Böhm and Ina Schiering[2]

Abstract: To foster privacy awareness for juveniles and young adults a serious game approach is proposed. This approach is based on a set of cards allowing in the context of participatory workshops to explore privacy risks and to gain competencies. The set of cards and the workshops concept was iteratively developed based on a set of workshops. The results were promising. In the future an evaluation of the approach is intended and the transfer to other target groups.

3.1 Introduction and Background

Digitization is gaining importance in a broad range of aspects of daily life, education, health and professional. While people in general consider privacy as important, individual privacy concerns vary [7, 3]. Privacy incidents as e.g. data breaches happen nearly on a regular basis. Recently Ticketmaster, a company selling tickets in various categories, e.g. sport concerts, theatre, etc. has lost 560 million costumer data in a data breach.[3]

Incidents as data breaches as well as careless chosen privacy settings in social media, or a harmful environment present privacy risks. According to the GDPR the controller is responsible to address risks concerning the fundamental rights and freedoms of natural persons. The controller needs to consider a risk-based approach where risks are analysed from the point of view of the data subject [8]. Beside the accountability of the controller, it is important for data subjects to have a personal perception for their own privacy risks. To this end we present a workshop based on a serious game approach to raise the awareness concerning privacy risks. In the context of scenarios chosen by the participants from everyday life, our workshop contains a participatory, card-based approach in which participants actively acquire knowledge in groups. Also, the workshop intends to foster the self-efficacy of the participants by discussions about potential measurements. This card-based workshop concept was iteratively developed in the context of several workshop with juveniles and young adults from 2022 until 2024.

3.2 Related Work

Serious games in security and privacy were already investigated in several contexts. Denning et al. proposed a set of cards to describe mainly motivation and applied methods of the adversary, as well as the impact focusing on the adversary point of view [4]. Karagiannis et al. categorize existing game-based approaches towards security and privacy education [10]. Specific trainings are proposed to raise awareness towards phishing based on the so-called No-Phish cards [1]. Comparable approaches are also proposed in a business context [5] for a phishing awareness campaign. Recently a privacy memory game was employed to raise privacy awareness and especially the informed decision of young children [2].

3.3 Methodology

During several participatory workshops between 2022 and 2023 to raise privacy awareness of juveniles and young adults based on a slightly simplified version of privacy risk identification and analysis proposed in [6], participants expressed privacy risks in a general way. But they had difficulties

[2] Ostfalia University of Applied Sciences, Wolfenbüttel, Germany, {tom.lorenz1, mic.pleger, t.boehm, i.schiering}@ostfalia.de

[3] https://www.malwarebytes.com/blog/personal/2024/06/ticketmaster-confirms-customer-data-breach

to specify the risk itself and were unaware of measurements they could apply concerning these risks. Instead of teaching these lacking competencies a serious game based on a set of cards was developed which allowed juveniles and young adults to explore privacy risks themselves and to gain competencies about privacy risks and measurements. The aim of these workshops is to raise awareness and to gain self-efficacy.

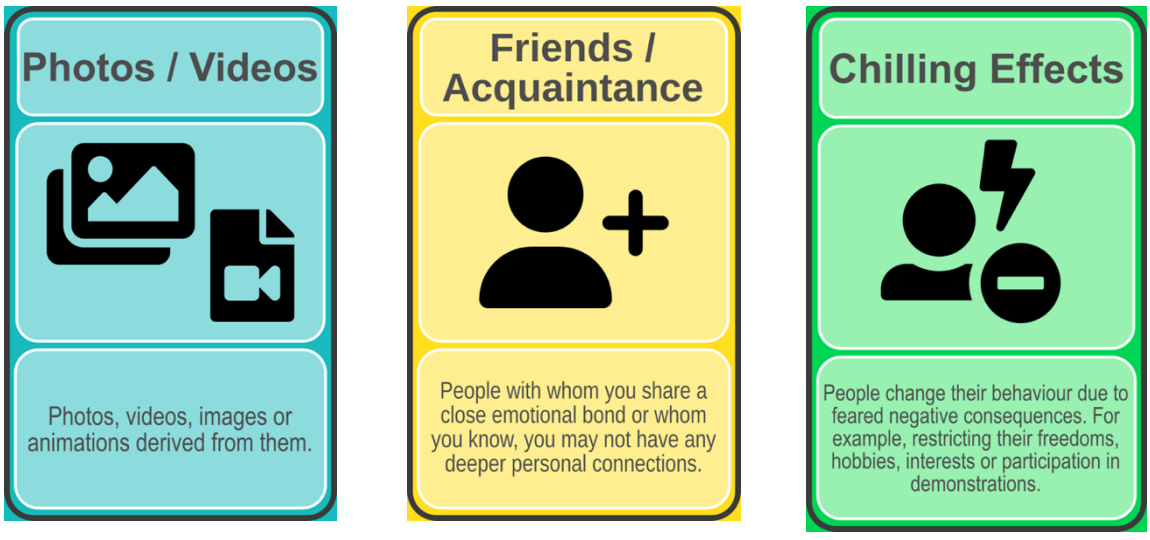
A first prototype of the set of cards and the workshop concept was iteratively developed based on the experience of these workshops. As a first validation this approach was tested in 2024 in five workshops with a total of about 140 participants. The feedback of the participants was very promising. In the first workshops with about 70 participants a digital set of cards was used, later in four workshops with between 15 and 20 participants a prototype of a paper-based set of cards was employed.

3.4 Privacy Awareness Cards

The participatory workshop is based on a set of cards which are briefly sketched in the following. The workshop starts with a short introduction based on a recent privacy incident as set of cards e.g. a data breach which is used to describe the use of the cards and to foster the discussion of the participants about scenarios in their everyday life. The participants choose in groups of about four to eight people a scenario and analyse privacy risks. The following categories of cards are used to describe the privacy risks perceived in the scenario:

- **Data Cards:** Types of personal data which is processed in the scenario (figure 1a)
- **Stakeholder Cards:** Stakeholders which are involved in the scenario (figure 1b)
- **Consequence Cards:** Effect on the data subject when the risk occurs (figure 1c)

Figure 1: Involved Stakeholders, personal data and consequences are the first three categories[4]

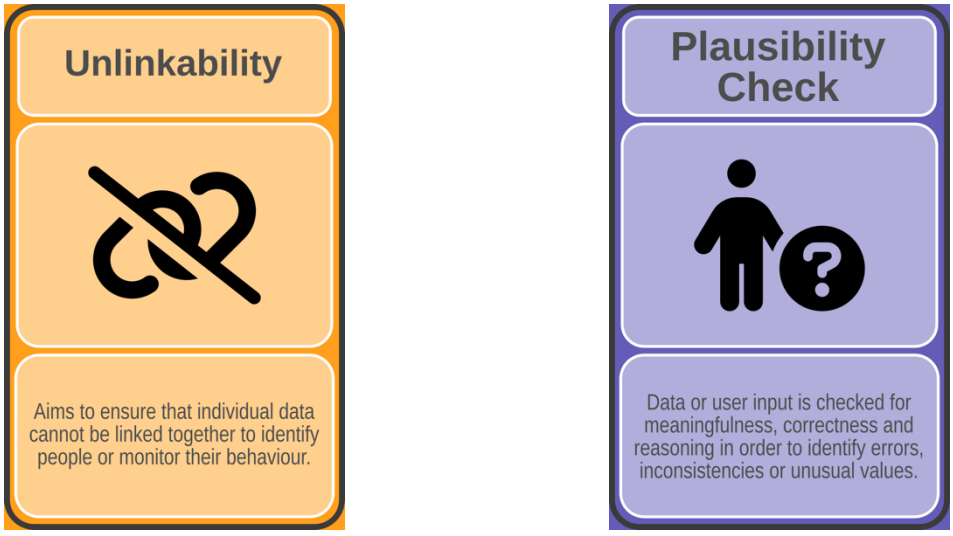


(a) Data Card (b) Stakeholder Card (c) Consequence Card

[4] Icons (Licensed under CC BY 4.0): Font Awesome Free 6.6.0 by @fontawesome - https://fontawesome.com License - https://fontawesome.com/license/free Copyright 2024 Fonticons, Inc.

The main focus is to increase awareness how privacy incidents happen, investigate the involved actors and describe potential damages. Based on this analysis of the scenario, in the next step the corresponding privacy risks are analysed from the point of view of the data subject based on cards employing the concept of privacy protection goals [9] (figure 2a). In the final step the groups discuss which actions or measurements they can employ themselves to address the risks (figure 2b). For all categories of cards also empty cards are provided to allow to add additional information which is lacking.

Figure 2: Privacy Goal Cards and Action Cards



(a) Privacy Goal Cards (b) Action Card

3.5 Discussion and Conclusion

The feedback for the participatory workshops and the set of cards is promising. In the following experiences from the workshops are summarized. While at first participants tend to be sceptical, an introduction based on relevant recent incidents, e.g. from the context of social media, leads typically to a vivid discussion.

The following phase, when the groups define a scenario which they want to analyse together, is crucial for the success of the workshop. Here sometimes guidance is needed to identify a suitable scenario from the daily life of the participants. The analysis of this scenario with the help of privacy awareness cards in groups leads to engaged discussions and to a thorough individual understanding of the identified risks and especially the consequences for the individual. Here the facilitators of the workshop need to pay attention that the scenario is described in sufficient detail for the analysis.

The participatory design of the workshop supports involvement and leads to positive feedback, making the topic privacy more tangible and fostering awareness. It is intended to optimize the design of the cards based on the current prototype. Also, measurements for guiding groups to describe their scenario in sufficient detail need to be developed. A thorough evaluation is planned and the transfer to further target groups.

Acknowledgment

This work was supported by the Federal Ministry of Education and Research (BMBF) as part of KI4All (16DHBKI056).

Bibliography

1. Aldag, L., Berens, B., Burgdorf, M., Lorenz, A., Thiery, M.C., Volkamer, M.: NoPhish-Challenge-Karten – Evaluation in der Praxis. Datenschutz und Datensicherheit - DuD 45(11), 721 (2021)
2. Appelt, D., Geissler, I.: Datenschutzsensibilisierung. Datenschutz und Datensicherheit - DuD 48(2), 125–127 (Feb 2024)
3. Coopamootoo, K., Gross, T.: Why Privacy Is All But Forgotten. Proceedings on Privacy Enhancing Technologies 2017 (Oct 2017)
4. Denning, T., Friedman, B., Kohno, T.: Poster—The Security Cards: A Security Threat Brainstorming Toolkit (2014)
5. Fox, D., Titze, C.: Phishing Awareness durch Gamification. Datenschutz und Datensicherheit - DuD 45(11), 727–732 (Nov 2021)
6. Friedewald, M., Schiering, I., Martin, N., Hallinan, D.: Data Protection Impact Assessments in Practice. In: Katsikas, S., Lambrinoudakis, C., Cuppens, N., Mylopoulos, J., Kalloniatis, C., Meng, W., Furnell, S., Pallas, F., Pohle, J., Sasse, M.A., Abie, H., Ranise, S., Verderame, L., Cambiaso, E., Maestre Vidal, J., Sotelo Monge, M.A. (eds.) Computer Security. ESORICS 2021 International Workshops. pp. 424–443. Springer International Publishing, Cham (2022)
7. Gerber, N., Gerber, P., Volkamer, M.: Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. Computers & Security 77, 226–261 (Aug 2018)
8. Hansen, M., Bieker, F., Bremert, B.: Datenschutz und Privatheitsschutz durch Gestaltung der Systeme. In: Roßnagel, A., Friedewald, M. (eds.) Die Zukunft von Privatheit und Selbstbestimmung: Analysen und Empfehlungen zum Schutz der Grundrechte in der digitalen Welt, pp. 259–300. Springer Fachmedien, Wiesbaden (2022)
9. Hansen, M., Jensen, M., Rost, M.: Protection Goals for Privacy Engineering. In: 2015 IEEE Security and Privacy Workshops. pp. 159–166 (May 2015)
10. Karagiannis, S., Papaioannou, T., Magkos, E., Tsohou, A.: Game-Based Information Security/Privacy Education and Awareness: Theory and Practice. In: Themistocleous, M., Papadaki, M., Kamal, M.M. (eds.) Information Systems, vol. 402, pp. 509–525. Springer International Publishing, Cham (2020)

4 On the way to data sensitization – What do smart home and smart city mean for citizens?

Alexander Rogalla, Björn Konopka, Manuel Wiesche and Simon Hensellek[5]

4.1 Introduction

In recent years, smart homes and smart cities have gained increasing attention as promising technologies to enhance citizens' quality of life by providing increased convenience, security, or energy efficiency (Marikyan et al., 2019). However, the rapid development of these technologies has also raised concerns among citizens, for instance, about privacy, data security, and the added value of these technologies (Balta-Ozkan et al., 2014). Since both smart home and smart city environments frequently handle sensitive user data, it is crucial to understand citizens' perceptions, knowledge gaps, and uncertainties regarding these technologies.

Our research aims to provide insights for developing targeted measures to raise awareness, provide education, and drive the creation of citizen-centric smart home and smart city solutions by bridging the gap between the theoretical concepts in literature and the practical understanding of citizens.

In our cross-sectional survey of 94 citizens, we found strong differences in the understanding as well as the knowledge levels related to smart homes and smart cities. Although there is a fundamental interest, there are also considerable knowledge gaps and uncertainties regarding data security and the practical use of smart services. The results emphasize the need to take citizens' concerns seriously and develop smart home and smart city solutions that are easy to use and understand, and that guarantee not only technological benefits and functionality but also privacy and data security.

4.2 Theoretical background: Smart services offer many advantages, but the practical concerns of citizens differ from the positive portrayal in the literature

Smart home and smart city services offer citizens valuable benefits in terms of convenience, security, energy efficiency, and quality of life (Marikyan et al., 2019). But what do citizens understand by smart home and smart city? The literature defines smart home as "a residence equipped with smart technologies aimed at providing tailored services for users. Smart technologies make it possible to monitor, control, and support residents, which can enhance the quality [of] life and promote independent living" (Marikyan et al., 2019, p. 139). Smart city refers to "promising communities that use intelligent technologies to connect people through internet devices to improve their quality of lives" (Zhou et al., 2023, p. 1).

While the literature emphasizes benefits, citizens often focus on potential risks related to the use of new technologies. The smart services offered in both environments often affect sensitive areas in the daily lives of users, raising concerns about issues such as privacy, data protection, and data security. Thus, Citizens understandably have concerns about the security and misuse of their personal data (Nehme & George, 2022), which may lead to reluctance among them to use smart services (Li et al., 2021). Moreover, citizens face a constant trade-off between functionality and privacy, as greater functionality often requires sharing more personal data, which increases privacy risks (Lenhart et al., 2023). Further, it appears that the positive portrayal of the smart home and smart

[5] Technische Universität Dortmund, Dortmund, Germany, {vorname.nachname}@tu-dortmund.de

city in the literature and research is not in line with the practical apprehensions and understanding of citizens (Marikyan et al., 2023).

This dichotomy between the optimistic academic perspective and general real-life understanding of citizens underscores the need for a more nuanced exploration of citizens' perceptions regarding smart home and smart city services. Understanding the citizens' perspective is crucial for developing user-centric smart technologies that address both the functional needs and the privacy requirements of potential adopters. The aim of this study is therefore to explore the understanding, acceptance, wishes, and concerns of citizens regarding smart home and smart city services. Further, the study addresses the gap between the existing concepts of smart homes and smart cities in the literature and the general real-life understanding of citizens about these concepts.

4.3 Methods: Explorative cross-sectional survey

To investigate citizens' perception and understanding of smart home and smart city technologies, we conducted an explorative cross-sectional survey at a city festival in Dortmund on May 4, 2024. This approach allowed us to capture a diverse sample of citizens in a natural setting, increasing the validity of our findings. Participants were invited to complete the survey on-site using provided devices or their own smartphones. The data was collected via a browser- based online questionnaire on LimeSurvey and analyzed with MaxQDA 24. A total of 94 respondents participated in the survey, of whom 39 (41.5%) were female, 49 (52.1%) male, and 2 (2.1%) were diverse. 4 people (4.3%) did not specify their gender. The exact descriptive statistics regarding age and gender are shown in Table 1. The participants were also highly diverse in terms of age (age range: 11-86 years, M = 35.6 years, SD = 16.76).

Table 1: Descriptive statistics for gender and age

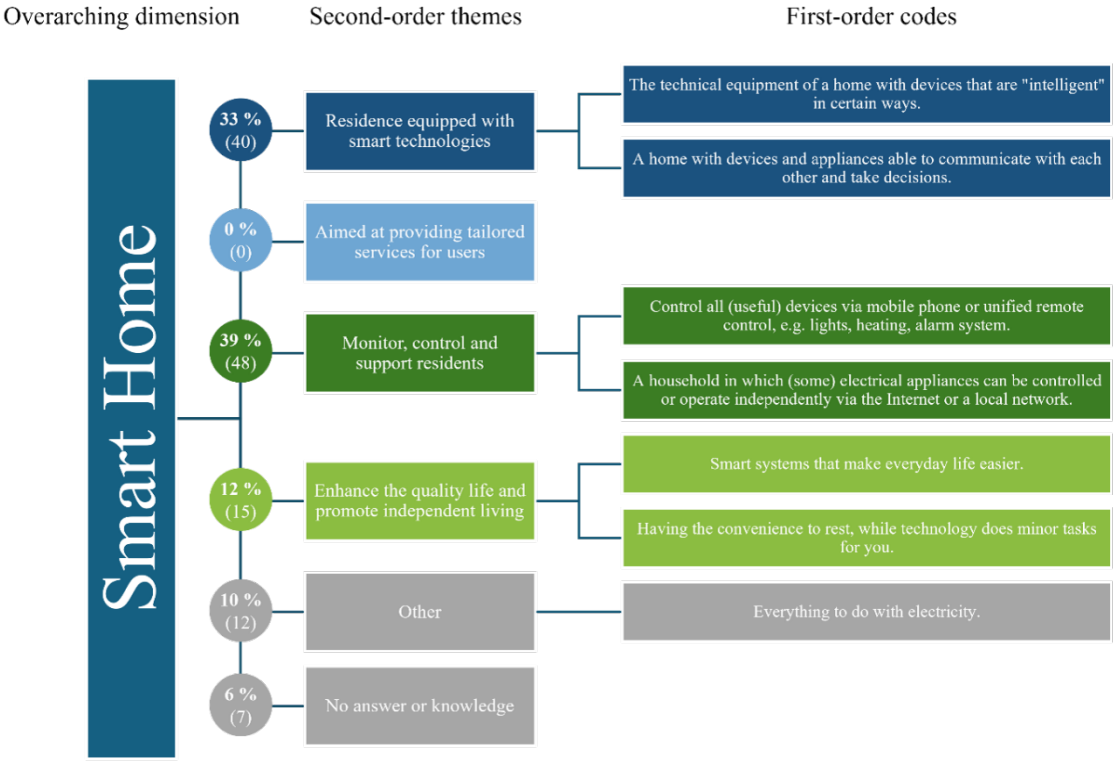
Variable	*n*	(%)	*M*	*SD*	Min	Max	No answer
Age			35.59	16.76	11	86	4
Male	39	41.5%					
Female	49	52.1%					
Diverse	2	2.1%					

Note: n = 94.

4.4 Methods: Explorative cross-sectional survey

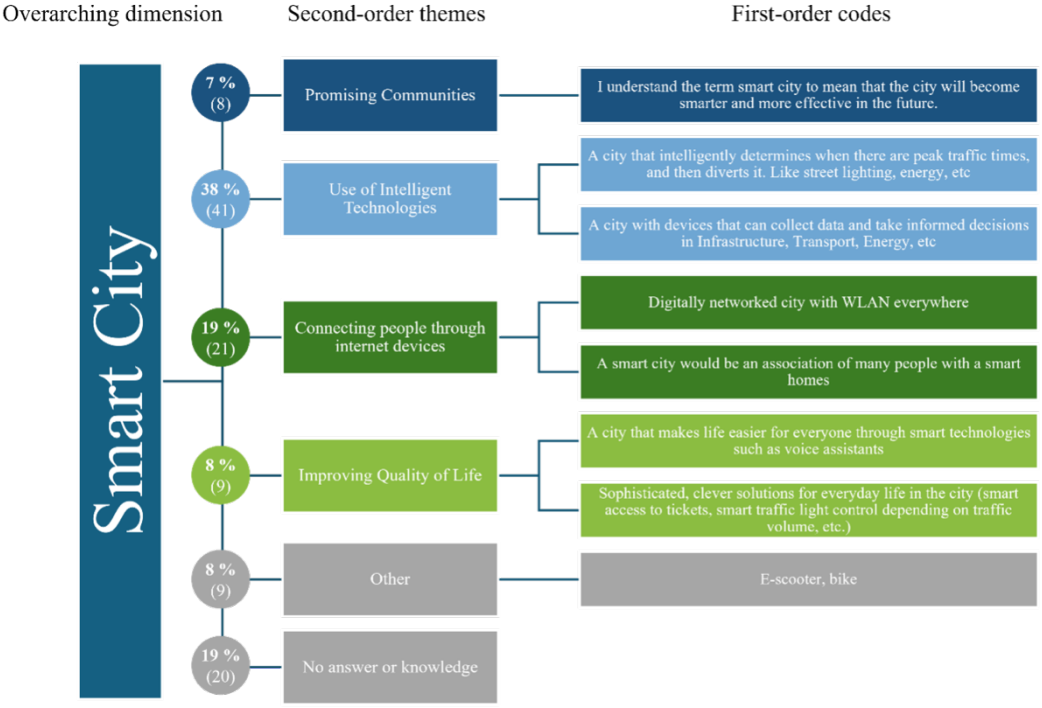
The survey provides data on the level of awareness and understanding of smart home and smart city among citizens. Citizens were asked to describe their understanding of smart home and smart city and to define the terms (as shown in Figure 1 and Figure 2). There were strong differences in the understanding of smart city, with some citizens not having a clear idea (20), others understanding it primarily as the use of smart technologies (41), and the networking of smart systems in urban infrastructures (21). Regarding smart home, most citizens understand smart home primarily as smart services for monitoring, control, and support (48), and as residences equipped with smart technologies (40). Interestingly, no one saw smart homes as providing customized services.

Figure 1: Classification of citizens' responses to the four dimensions of the smart home definition based on Marikyan et al. (2019).



Note. n = 94. Citizens' responses were classified based on the four dimensions outlined in the definition (Marikyan et al., 2019). Percentages of the coded segments are displayed in circles. Frequencies are indicated in parentheses.

Figure 2: Classification of citizens' responses to the four dimensions of the smart city definition based on Zhou et al. (2023).



Note. n = 94. Citizens' responses were classified based on the four dimensions outlined in the definition (Zhou et al. (2023). Percentages of the coded segments are displayed in circles. Frequencies are indicated in parentheses.

Moreover, the results indicate that citizens have different perceptions and levels of knowledge about well-known smart home products and brands. For example, brands such as Amazon Alexa (54), Apple Home (17), Google Home (12), and Phillips Hue (11), and products such as smart lighting systems (44), smart energy systems (26), and autonomous (vacuum cleaning) robots (18) were frequently mentioned. However, knowledge about the exact functionality and data security aspects varied considerably. A detailed list of all the brands mentioned is shown in Figure 3.

4.5 Conclusion: Considerable knowledge gaps and uncertainties regarding data security, smart home, and smart city

The results of the study are multifaceted. Whilst they demonstrate that there is a fundamental interest and a certain level of awareness of smart home and smart city services among citizens, there are also considerable gaps in knowledge and uncertainties regarding data security and practical application. These findings highlight the need for targeted education, awareness, and information campaigns to increase acceptance and understanding of these technologies among the general population. In addition, the results emphasize the need to take the concerns of citizens seriously and to develop smart home and smart city solutions that ensure both technological advantages and functionality, as well as the protection of privacy and data security. Furthermore, many citizens primarily associate smart home and smart city with the touchpoints offered by smart services, such as app interfaces on their smartphones and tablets, which are used to monitor and control their smart environment. This study contributes to the literature by bridging the gap between the theoretical concepts and citizens' practical understanding of smart homes and smart cities. Further, the results can inform the development of targeted measures to raise awareness, provide education, and drive the creation of technologies that address the needs and concerns of citizens.

4.6 Outlook: Towards a data-secure future

Future studies should first contrast, and then integrate the different perspectives and understandings of researchers, practitioners, and citizens regarding smart homes and smart cities. This could involve investigating similarities and differences between smart homes and cities as well as examining how citizens' attitudes, understandings, and knowledge levels evolve over time, and whether their preferences related smart services change accordingly. Incorporating qualitative methods in a mixed-methods approach, such as in-depth interviews or focus groups, could also lead to a more nuanced understanding of citizens' concerns and expectations.

Subsequently, the decision-making processes of citizens for specific smart products and services should be examined in depth, for instance, by using conjoint analysis to depict realistic decision-making situations. Further research could explore how enhanced privacy solutions can increase the acceptance of smart home and smart city services. Moreover, future studies could investigate the impact of targeted education and awareness campaigns on the understanding and acceptance of smart home and smart city services and validate the findings of this study in a larger sample.

Building on this study, the BMBF-funded research project Optimization of Informational Sustainability for Citizens in Data Ecosystems (Opt-IN) is conducting research in data protection and digital self-determination within the smart home and smart city context. The project aims to develop and research data protection-compliant ecosystems and innovative digital smart home and city solutions ("made in Germany"), based on real-world use cases to maximize customer benefit. These solutions will enable citizens to transparently control the use of their data and exercise their fundamental right to data sovereignty, while also providing high customer benefits.

Acknowledgments

This research was funded by the German Federal Ministry for Education and Research as part of the project Opt-IN (Ref. 16KIS1935K).

Bibliography

Balta-Ozkan, N., Boteler, B., & Amerighi, O. (2014). European smart home market development: Public views on technical and economic aspects across the United Kingdom, Germany and Italy. Energy Research and Social Science, 3(C), 65–77. https://doi.org/10.1016/j.erss.2014.07.007

Lenhart, A., Park, S., Zimmer, M., & Vitak, J. (2023). "You Shouldn't Need to Share Your Data ": Perceived Privacy Risks and Mitigation Strategies Among Privacy-Conscious Smart Home Power Users. Proceedings of the ACM on Human-Computer Interaction, 7(October), 34. https://doi.org/10.1145/3610038

Li, W., Yigitcanlar, T., Erol, I., & Liu, A. (2021). Motivations, barriers and risks of smart home adoption: From systematic literature review to conceptual framework. Energy Research and Social Science, 80(March 2021), 102211. https://doi.org/10.1016/j.erss.2021.102211

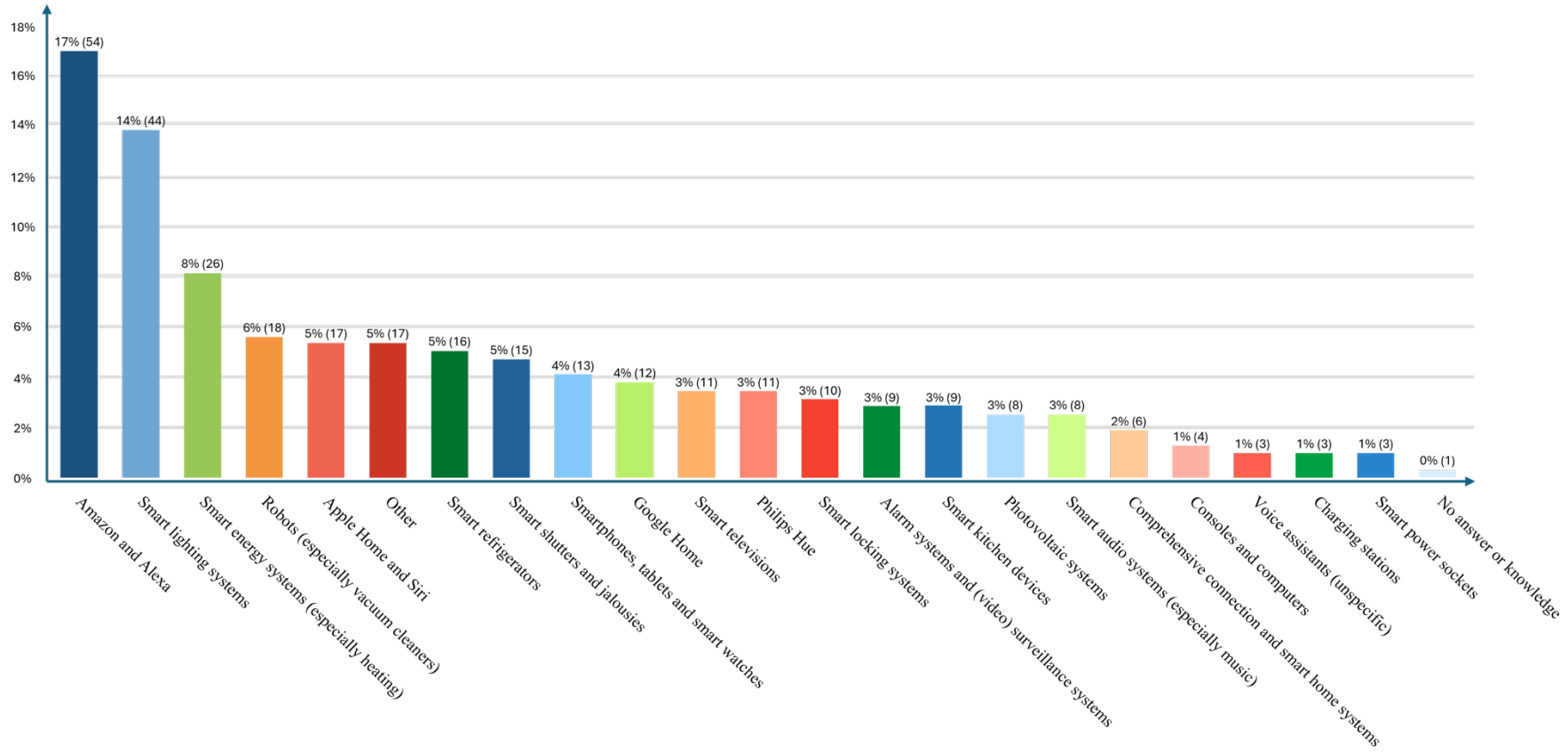
Marikyan, D., Papagiannidis, S., & Alamanos, E. (2019). A systematic review of the smart home literature: A user perspective. Technological Forecasting and Social Change, 138(August 2018), 139–154. https://doi.org/10.1016/j.techfore.2018.08.015

Marikyan, D., Papagiannidis, S., & Alamanos, E. (2023). Cognitive Dissonance in Technology Adoption: A Study of Smart Home Users. Information Systems Frontiers, 25(3), 1101–1123. https://doi.org/10.1007/s10796-020-10042-3

Nehme, A., & George, J. F. (2022). Approaching IT Security & Avoiding Threats in the Smart Home Context. Journal of Management Information Systems, 39(4), 1184–1214. https://doi.org/10.1080/07421222.2022.2127449

Zhou, S., Loiacono, E. T., & Kordzadeh, N. (2023). Smart cities for people with disabilities: a systematic literature review and future research directions. European Journal of Information Systems, 00(00), 1–18. https://doi.org/10.1080/0960085X.2023.2297974

Figure 3: Consumer awareness of smart products and brands.



Note. n = 94. Each participant could provide up to six responses

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Plattform
Privatheit

PROJEKTPARTNER

Fraunhofer

Natur
Technik
Kultur
Gesellschaft

UNIKASSEL
VERSITÄT