# On the Tension Between Open Data and Data Protection in Research

Luisa Jansen[1,*], Nele Borgert[1] and Malte Elson[1]

[1]*University of Bern, Institute of Psychology, Fabrikstrasse 8, 3012 Bern, Switzerland*

## Abstract

The push for Open Science practices, including open data, clashes with the need for strict data protection of participant information. This creates a growing crisis in human-computer interaction (HCI) research. Despite its critical role, data protection remains an afterthought in metascience, leaving researchers without clear guidance and dedicated resources, endangering both participant privacy and scientific openness. We illustrate that Meta-HCI is uniquely positioned to address this challenge by investigating how researchers navigate these tensions and developing strategies that align openness with privacy. We propose starting points for solutions such as minimizing data collection and reusing datasets. This is a call to action—without urgent intervention, both the privacy of research data and Open Science are at risk.

## Keywords

data protection, open science, data sharing, metascience, human-computer interaction

## 1. Open Data and Data Protection

Scientific research is caught in a conflict: On the one hand, the Open Science Movement pushes for the sharing of research data, enabling assessment of research integrity and reuse of data, with growing success. On the other hand, legal and ethical frameworks require strict protection of participants' personal data, preventing potential harm to individuals. This tension affects human-computer interaction (HCI) research, which often involves personal, behavioral, and physiological data.

Data protection practices of researchers have been largely overlooked by empirical research. This lack of attention creates a gap in guidance, leaving researchers to navigate uncertainty with little institutional support. The result is a situation where both Open Science and privacy protections are at risk.

In this position paper, we explore the rising tension between these competing priorities. We argue that meta-scientists and HCI researchers have a unique expertise to lead the way in addressing the tension between scientific openness and privacy. We propose potential starting points for prioritizing these seemingly opposing principles.

---

*Meta-HCI '25: First Workshop on Meta-Research in HCI, April 26, 2025, Yokohama, Japan*

[*] Corresponding author.

luisa.jansen@unibe.ch (L. Jansen); nele.borgert@unibe.ch (N. Borgert); malte.elson@unibe.ch (M. Elson)
0000-0001-8126-1306 (L. Jansen); 0000-0002-0255-5822 (N. Borgert); 0000-0001-7806-9583 (M. Elson)

### 1.1. Data Protection in Research

The informational self-determination of participants is a central element of research ethics when conducting studies involving human subjects [1, 2]. Beyond this ethical obligation, researchers must often comply with data protection legislation [3, 4]. Adherence to these ethical and legal standards is essential for maintaining participants' trust in research [5, 6].

For researchers, compliance may require implementing a variety of measures: asking for informed consent of participants, abstaining from collecting identifying data, or adding noise at the data collection stage. After data collection, anonymizing or pseudonymizing data, blocking access to data, deleting no longer necessary data, providing secure data storage, ensuring participants' rights like access to or rectification of data, or using data masking techniques might be necessary.

Data protection, with all the measures it entails, can be viewed as a secondary task for researchers: It does not straightforwardly help with generating knowledge, finding truth, or—on a more practical note—publishing papers, but rather creates friction for a researcher's primary tasks.

### 1.2. Data Sharing in Research

The Open Science Movement calls for creating transparency at every stage of the research process [7], perhaps most prominently visible in the rise of open data [8]. The core goal is to enhance the reproducibility of research, with open data being especially important for estimations of outcome reproducibility and robustness, dominating much of meta-scientific research [9, 10]. This movement has rightfully gained much more ground in recent years, especially with findings showing a lack of reproducibility across several scientific areas [11, 12, 13]. We believe to be witnessing a paradigm shift in the social sciences, moving from closed to Open Science, evident in, e.g., the increase in open data [14, 15].

Much like data protection, Open Science is a secondary task for most researchers who do not work on metascience: Creating open data sets does (at least traditionally) not help researchers with their primary goal. On the contrary, some researchers might actively try to evade calls for data sharing to disguise their scientific misconduct. Even for researchers upholding high integrity standards, the creation of accessible data sets takes valuable time and might be perceived as annoying.

### 1.3. The Tension Between Data Sharing and Data Protection

There is a clear tension between data protection, which essentially calls for the minimization of access to data, and data sharing, which essentially calls for the maximization of access to data. Figure 1 illustrates these opposing principles. In addition to this tension, both data protection and Open Science as secondary tasks may compete for resources in a fast-paced, output-oriented research landscape.

Little is known about how researchers weigh these priorities in practice. Some interview studies investigating Open Science practices showed researchers' data protection concerns keeping them from sharing their data [16, 17, 18]. Similarly, in the metascience literature, data protection seems to be a minor (and rather annoying) side note, treated as if it is not part of the

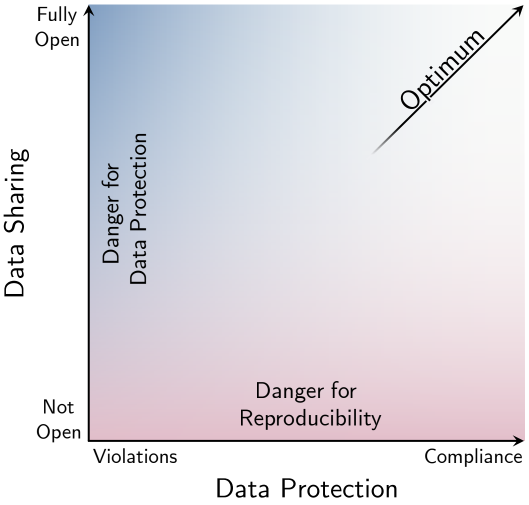

**Figure 1:** The tension created by data protection and data sharing, as well as potential areas of dangers.

scientific process. However, in recent years, data protection has gained some attention in terms of Open Science practices [19, 20].

Possibly due to the lack of research, there is virtually no concrete guidance for researchers on how to navigate this tension, aside from some general recommendations (see [21] for an exceptionally comprehensive, yet still very broad guide, or [20] for a "tutorial" presenting two case studies without offering further guidance). Moreover, the existing research infrastructure does not provide dedicated resources to address these challenges. Institutional review boards that could, in theory, fill this gap are not adequately equipped to do so [22]. In turn, first empirical assessments (that involve authors of this paper) indicate that researchers struggle to comply with data protection requirements [23, 24]: In the area of HCI, Martius et al. [24] conduct an exploratory study into data protection behaviors by Usable Security and Privacy (USP) researchers and institutional conditions as presented by the data protection officers of research institutions. Their results showcase many insecurities and obstacles regarding data protection, even for researchers who have theoretical knowledge about the topic. However, related works that discuss research ethics in HCI and touch on data protection by doing so are rather rare [25, 26, 27, 28, 15, 29].

We argue that the status quo endangers both Open Science as well as participants' privacy:

**The status quo endangers Open Science.** The current situation allows privacy concerns to become a scapegoat for data withholding: Researchers cite data protection as a reason for non-sharing of data, either due to legitimate uncertainty or strategic avoidance, as indicated by Hussey [30] even when promising to share the data upon request. If data remains unshared, empirical claims cannot be independently verified, making the assessment of reproducibility unfeasible and threatening the very foundation of good research.

**The status quo endangers participants' privacy.** Currently, researchers sharing data have to do so on their own—in spite of many uncertainties [24]. This might lead to the accidental disclosure of personal information in data sets. Even in the case of successful deletion of identifiers before sharing, traditional approaches such as anonymization may fail for large data sets in view of advancing re-identification techniques [21, 31]. Also, in the case of smaller study populations (e.g., marginalized groups), the combination of a few demographic data points might suffice to identify participants. These issues potentially expose participants to harm and, in turn, threaten the very trust that research involving human subjects depends on. This might, in the long term, endanger human subject research on a systemic level. More immediately, violations of data protection laws might also lead to legal repercussions for individual researchers or their institutions.

Recent developments further amplify this tension: With the rise of the Open Science Movement, more researchers than ever publish their data and are incentivized to do so [14]. At the same time, advances in re-identification techniques and the multitude of already existing data sets that might be linked to a new one make it even harder for individual researchers to determine the extent of data modification required to achieve complete and lasting anonymization. Since there is no indication that these trends will slow down, we believe the urgency of this problem will continue to grow.

## 2. HCI's Potential Impact on Data Protection

The scientific community working with human participants needs to investigate how to effectively prioritize data protection and data sharing to mitigate these dangers. We regard HCI as uniquely equipped in matters of data protection: For HCI researchers, privacy is important not only because of general ethical mandates or legal obligations that apply to all researchers but also because it is the community that actually investigates data protection behavior of humans. For human-centered security and privacy researchers, a sub-community of HCI, data protection is not only a secondary task but a primary one [32, 33]. Analyzing data protection behaviors, creating support for the protection of data, and researching conditions surrounding data protection are central elements of USP research, where general users, policymakers, as well as developers are investigated.

We believe that resolving the current problems arising from the tension between Open Science and privacy requires more information, enabling improved guidance and stronger advocacy for institutional support and resources. We strongly encourage HCI researchers to leverage their expertise in privacy and the data handling of other groups (e.g., developers) by collaborating with meta-scientists who can contribute knowledge about general research conditions and processes. By joining forces, they may be able to answer those pressing questions about data protection behavior in general and, more specifically, within the context of Open Science:

- How often do data protection and Open Science conflict in reality?
- How often are data protection concerns used as a pretext for unwillingness to publish data?
- How successful are data protection efforts of research data?
- How common are data protection violations when sharing data?
- How can one find a fair trade-off between data protection and data sharing?
- What resources are necessary to enable fair prioritization of data protection and data sharing?

- What resources are already provided to enable fair trade-offs between data protection and data sharing?
- How can one involve different stakeholders (especially participants) in prioritizing data protection and data sharing?
- How can one develop effective guidance on these issues?

We encourage Meta-HCI research teams not to limit their investigations to their own community but rather extend them to all scientists working with human participants.

## 3. Starting Points to Resolving the Tension

We want to highlight some potential starting points to ease the tension between data protection and data sharing. We propose to mitigate the presented dangers by improving the status quo towards fewer risks (i.e., decreasing the risks associated with data sharing) and more utility (i.e., increasing the usefulness of collected data).

### 3.1. Decrease Risks Associated With Data Sharing

**Involve third-party actors with data protection knowledge.** Professionals such as data stewards with privacy expertise can provide an external perspective, helping researchers navigate complex data protection regulations, implement best practices, and assess privacy risks before sharing data [34]. By making their involvement obligatory, institutions can ensure that data protection is not left entirely to individual researchers, who may lack the necessary expertise or time to properly evaluate privacy concerns.

**Promote data obfuscation techniques.** Data obfuscation techniques (e.g., masking, differential privacy, data aggregation, or synthetic data generation) can be employed to minimize privacy risks while still allowing for data sharing [35, 36, 21]. However, the implementation of these techniques may require specialized knowledge and additional effort, potentially discouraging researchers. Additionally, obfuscation may reduce the perceived value of data for secondary research.

**Integrate data protection in technology for data sharing.** Existing data-sharing infrastructures should include built-in privacy protection mechanisms, e.g., automated tools to remove identifiers. By integrating these features directly into centralized data repositories, research platforms, or institutional data management systems, privacy protection becomes a part of the sharing process rather than an afterthought. The data-sharing platform ICSPR, for example, checks for identifying data in published data sets [21, 37] and the medical analytics platform OpenSAFELY enables the analysis of health records while protecting confidentiality [38].

**Commit to the risks.** Seemingly, the anonymization of research data is indispensable. Brown et al. [25] propose to breach this necessity when participants might benefit from the publication of identifying data. Recognizing the participants as contributors and, therefore, making their involvement explicit can be a solution in the case of qualitative or participatory studies [25]. This strategy does not technically reduce the risks associated with data sharing but rather states there are none in special cases. Similarly, risks may be more tolerable when participants are fully informed and agree freely, as ensured by consent comprehension quizzes [21, 39].

### 3.2. Increase the Utility of Collected Data

We see a promising way forward that might support both data protection and data sharing interests: increasing the usefulness of collected data. The Open Science Movement's aim for an increase in scientific quality and, therefore, the reproducibility of science can also indirectly assist data protection interests:

**Collect less data.** One way to align data protection with the goals of Open Science is to focus on collecting less data while maximizing its utility. A more selective approach to data collection encourages researchers to be more diligent in designing studies and prioritizing quality over quantity. This shift necessitates alternative incentives that reward rigorous, high-impact research rather than sheer output. Large-scale socioeconomic studies, for example, demonstrate how carefully designed research can address multiple questions while minimizing redundant data collection (e.g., the Socio-Economic Panel in Germany [40]).

**Share and reuse data.** This might seem contradictory, but although data sharing can increase privacy risks, it also enables the reuse of data. That is why we advocate for more responsible and FAIR (i.e., findable, accessible, interoperable, reusable) data sharing [41]. The reuse and reanalysis of already existing data lead to less data collection and, therefore, align with data minimization interests. But, as of today, the vast amount of open data remains underutilized. Propagating the reuse of data aligns with both Open Science and data protection interests. For HCI, a cultural change toward data reuse might be especially challenging: Here, the predominance of novelty as a quality criterion conflicts with the usage of another researcher's data [42, 43].

## 4. Conclusion

Empirical research in itself conflicts with data protection. This tension is amplified by the increase in scientific openness (i.e., primarily, data sharing). HCI, as a research community, investigates human behavior in the context of technology and, consequently, data and its protection. Therefore, the HCI community is in a prime position to investigate researchers' data behavior within researcher-centered privacy studies. In spite of first efforts to that end, a multitude of research questions remain open. This leaves researchers abandoned when making data-sharing decisions with both data protection and sharing exposed to dangers, undermining good human-centered research in HCI and elsewhere. In light of the increasing urgency, we strongly encourage researchers to investigate this topic. We propose starting points for research and interventions to decrease the privacy risks associated with data sharing and to increase the utility of collected data, advancing both data protection efforts and the quality of science.

## Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT-4o and DeepL in order to perform grammar, translation, and spelling checks. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

## References

[1] ACM, ACM Publications Policy on Research Involving Human Participants and Subjects, 2021. URL: https://www.acm.org/publications/policies/research-involving-human-participants-and-subjects.

[2] A. Rouvroy, Y. Poullet, The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy, in: S. Gutwirth, Y. Poullet, P. De Hert, C. De Terwangne, S. Nouwt (Eds.), Reinventing Data Protection?, Springer Netherlands, Dordrecht, 2009, pp. 45–76. URL: http://link.springer.com/10.1007/978-1-4020-9498-9_2. doi:`10.1007/978-1-4020-9498-9_2`.

[3] E. S. Dove, The EU General Data Protection Regulation: Implications for International Scientific Research in the Digital Era, Journal of Law, Medicine & Ethics 46 (2018) 1013–1030. doi:`10.1177/1073110518822003`.

[4] T. Greene, G. Shmueli, S. Ray, J. Fell, Adjusting to the GDPR: The Impact on Data Scientists and Behavioral Researchers, Big Data 7 (2019) 140–162. URL: https://www.liebertpub.com/doi/abs/10.1089/big.2018.0176. doi:`10.1089/big.2018.0176`, publisher: Mary Ann Liebert, Inc., publishers.

[5] M. Guillemin, L. Gillam, E. Barnard, P. Stewart, H. Walker, D. Rosenthal, "Doing Trust": How Researchers Conceptualize and Enact Trust in Their Research Practice, Journal of Empirical Research on Human Research Ethics 11 (2016) 370–381. URL: http://journals.sagepub.com/doi/10.1177/1556264616668975. doi:`10.1177/1556264616668975`.

[6] N. Eyal, Using informed consent to save trust, Journal of Medical Ethics 40 (2014) 437–444. URL: https://jme.bmj.com/lookup/doi/10.1136/medethics-2012-100490. doi:`10.1136/medethics-2012-100490`.

[7] M. R. Munafò, B. A. Nosek, D. V. M. Bishop, K. S. Button, C. D. Chambers, N. Percie Du Sert, U. Simonsohn, E.-J. Wagenmakers, J. J. Ware, J. P. A. Ioannidis, A manifesto for reproducible science, Nature Human Behaviour 1 (2017) 0021. URL: https://www.nature.com/articles/s41562-016-0021. doi:`10.1038/s41562-016-0021`.

[8] V. Gewin, Data sharing: An open mind on open data, Nature 529 (2016) 117–119. URL: https://www.nature.com/articles/nj7584-117a. doi:`10.1038/nj7584-117a`.

[9] B. A. Nosek, G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. L. Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, T. Yarkoni, Promoting an open research culture, Science

348 (2015) 1422–1425. URL: https://www.science.org/doi/10.1126/science.aab2374. doi:10.1126/science.aab2374.

[10] B. A. Nosek, T. E. Hardwicke, H. Moshontz, A. Allard, K. S. Corker, A. Dreber, F. Fidler, J. Hilgard, M. Kline Struhl, M. B. Nuijten, J. M. Rohrer, F. Romero, A. M. Scheel, L. D. Scherer, F. D. Schönbrodt, S. Vazire, Replicability, Robustness, and Reproducibility in Psychological Science, Annual Review of Psychology 73 (2022) 719–748. URL: https://www.annualreviews.org/doi/10.1146/annurev-psych-020821-114157. doi:10.1146/annurev-psych-020821-114157.

[11] Open Science Collaboration, Estimating the reproducibility of psychological science, Science 349 (2015). URL: https://www.science.org/doi/10.1126/science.aac4716. doi:10.1126/science.aac4716.

[12] C. G. Begley, L. M. Ellis, Raise standards for preclinical cancer research, Nature 483 (2012) 531–533. URL: https://www.nature.com/articles/483531a. doi:10.1038/483531a.

[13] C. F. Camerer, A. Dreber, F. Holzmeister, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, G. Nave, B. A. Nosek, T. Pfeiffer, A. Altmejd, N. Buttrick, T. Chan, Y. Chen, E. Forsell, A. Gampa, E. Heikensten, L. Hummer, T. Imai, S. Isaksson, D. Manfredi, J. Rose, E.-J. Wagenmakers, H. Wu, Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015, Nature Human Behaviour 2 (2018) 637–644. URL: https://www.nature.com/articles/s41562-018-0399-z. doi:10.1038/s41562-018-0399-z.

[14] M. Hahnel, G. Smith, A. Campbell, The state of Open Data 2024: Special report - bridging policy and practice in data sharing, Technical Report, Digital Science, 2024. URL: https://digitalscience.figshare.com/articles/report/_b_The_State_of_Open_Data_2024_Special_Report_b_Bridging_policy_and_practice_in_data_sharing/27337476/2. doi:10.6084/M9.FIGSHARE.27337476.V2.

[15] K. Salehzadeh Niksirat, L. Goswami, P. S. B. Rao, J. Tyler, A. Silacci, S. Aliyu, A. Aebli, C. Wacharamanotham, M. Cherubini, Changes in Research Ethics, Openness, and Transparency in Empirical Studies between CHI 2017 and CHI 2022, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, ACM, Hamburg Germany, 2023, pp. 1–23. URL: https://dl.acm.org/doi/10.1145/3544548.3580848. doi:10.1145/3544548.3580848.

[16] S. Bangani, M. Moyo, Data Sharing Practices among Researchers at South African Universities, Data Science Journal 18 (2019) 28. URL: https://datascience.codata.org/article/10.5334/dsj-2019-028/. doi:10.5334/dsj-2019-028.

[17] O. S. Buhomoli, P. S. Muneja, Research Data Handling by Researchers in the Selected Universities in Tanzania, University of Dar es Salaam Library Journal 16 (2022) 53–69. URL: https://www.ajol.info/index.php/udslj/article/view/220959. doi:10.4314/udslj.v16i2.5.

[18] T. Weller, A. Monroe-Gulick, Understanding methodological and disciplinary differences in the data practices of academic researchers, Library Hi Tech 32 (2014) 467–482. URL: https://www.emerald.com/insight/content/doi/10.1108/LHT-02-2014-0021/full/html. doi:10.1108/LHT-02-2014-0021.

[19] D. Hallinan, F. Boehm, A. Külpmann, M. Elson, Information Provision for Informed Consent Procedures in Psychological Research Under the General Data Protection Regulation: A Practical Guide, Advances in Methods and Practices in Psychological Science 6 (2023). URL: https://journals.sagepub.com/doi/10.1177/25152459231151944.

doi:10.1177/25152459231151944.

[20] D. Van Ravenzwaaij, M. De Jong, R. Hoekstra, S. Scheibe, M. M. Span, I. Wessel, V. E. Heininga, De-identification when making datasets FAIR: Two worked examples from the behavioral and social sciences, 2023. URL: https://osf.io/acpm3. doi:10.31234/osf.io/acpm3.

[21] M. N. Meyer, Practical Tips for Ethical Data Sharing, Advances in Methods and Practices in Psychological Science 1 (2018) 131–144. URL: http://journals.sagepub.com/doi/10.1177/2515245917747656. doi:10.1177/2515245917747656.

[22] K. Ložnjak, A. Malešević, M. Čargo, A. Mladinić, Z. Koporc, L. Puljak, Comparison of Heads of Research Ethics Committees with Data Protection Officers on Personal Data Protection in Research: A Mixed-Methods Study with Structured Interviews, Journal of Academic Ethics 22 (2024) 577–598. URL: https://link.springer.com/10.1007/s10805-024-09509-8. doi:10.1007/s10805-024-09509-8.

[23] D. Hallinan, F. Boehm, A. Külpmann, M. Elson, (Un)informed Consent in Psychological Research: An empirical study on consent in psychological research and the GDPR, Journal of Open Access to Law 11 (2023) 1–28. URL: https://ojs.law.cornell.edu/index.php/joal/article/view/140.

[24] F. Martius, L. Jansen, L. Struck, A. Arumugam, L. Geierhaas, A.-M. Ortloff, M. Smith, C. Tiefenau, Out of Sight, Out of Mind? Exploring Data Protection Practices for Personal Data in Usable Security & Privacy Studies, ACM, Yokohama, Japan, 2025. doi:10.1145/3706598.3713654, in press.

[25] B. Brown, A. Weilenmann, D. McMillan, A. Lampinen, Five Provocations for Ethical HCI Research, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM, San Jose California USA, 2016, pp. 852–863. URL: https://dl.acm.org/doi/10.1145/2858036.2858313. doi:10.1145/2858036.2858313.

[26] C. Frauenberger, A. S. Bruckman, C. Munteanu, M. Densmore, J. Waycott, Research Ethics in HCI: A Town Hall Meeting, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM, Denver Colorado USA, 2017, pp. 1295–1299. URL: https://dl.acm.org/doi/10.1145/3027063.3051135. doi:10.1145/3027063.3051135.

[27] H. Hochheiser, R. S. Valdez, Human-Computer Interaction, Ethics, and Biomedical Informatics, Yearbook of Medical Informatics 29 (2020) 93–98. URL: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0040-1701990. doi:10.1055/s-0040-1701990.

[28] C. Munteanu, H. Molyneaux, W. Moncur, M. Romero, S. O'Donnell, J. Vines, Situational Ethics: Re-thinking Approaches to Formal Ethics Requirements for Human-Computer Interaction, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM, Seoul Republic of Korea, 2015, pp. 105–114. URL: https://dl.acm.org/doi/10.1145/2702123.2702481. doi:10.1145/2702123.2702481.

[29] J. Waycott, C. Munteanu, H. Davis, A. Thieme, S. Branham, W. Moncur, R. McNaney, J. Vines, Ethical Encounters in HCI: Implications for Research in Sensitive Settings, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM, Denver Colorado USA, 2017, pp. 518–525. URL: https://dl.acm.org/doi/10.1145/3027063.3027089. doi:10.1145/3027063.3027089.

[30] I. Hussey, Data is not available upon request, 2023. URL: https://osf.io/jbu9r. doi:10.31234/

osf.io/jbu9r.

[31] A. Narayanan, V. Shmatikov, Robust De-anonymization of Large Sparse Datasets, in: 2008 IEEE Symposium on Security and Privacy (sp 2008), IEEE, Oakland, CA, USA, 2008, pp. 111–125. URL: http://ieeexplore.ieee.org/document/4531148/. doi:10.1109/SP.2008.33.

[32] S. Dodier-Lazaro, R. Abu-Salma, I. Becker, M. A. Sasse, From paternalistic to user-centred security: Putting users first with value-sensitive design (2017). URL: https://discovery.ucl.ac.uk/id/eprint/1557895/.

[33] V. Distler, M. Fassl, H. Habib, K. Krombholz, G. Lenzini, C. Lallemand, V. Koenig, L. F. Cranor, Empirical Research Methods in Usable Privacy and Security, in: N. Gerber, A. Stöver, K. Marky (Eds.), Human Factors in Privacy Research, Springer International Publishing, Cham, 2023, pp. 29–53. URL: https://link.springer.com/10.1007/978-3-031-28643-8_3. doi:10.1007/978-3-031-28643-8_3.

[34] C. L. Borgman, Open Data, Grey Data, and Stewardship: Universities at the Privacy Frontier, Berkeley Technology Law Journal 2 (2018) 365–412. URL: https://lawcat.berkeley.edu/record/1128546. doi:10.15779/Z38B56D489.

[35] G. Sarada, N. Abitha, G. Manikandan, N. Sairam, A few new approaches for data masking, in: 2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015], IEEE, Nagercoil, India, 2015, pp. 1–4. URL: http://ieeexplore.ieee.org/document/7159301/. doi:10.1109/ICCPCT.2015.7159301.

[36] D. Bakken, R. Rarameswaran, D. Blough, A. Franz, T. Palmer, Data obfuscation: Anonymity and desensitization of usable data sets, IEEE Security and Privacy Magazine 2 (2004) 34–41. URL: http://ieeexplore.ieee.org/document/1366117/. doi:10.1109/MSP.2004.97.

[37] Institute for Social Research at the University of Michigan, Preserving Respondent Confidentiality, 2025. URL: https://www.icpsr.umich.edu/web/pages/deposit/confidentiality.html.

[38] Bennett Institute for Applied Data Science, OpenSAFELY, 2025. URL: https://www.opensafely.org/.

[39] J. E. Lunshof, R. Chadwick, D. B. Vorhaus, G. M. Church, From genetic privacy to open consent, Nature Reviews Genetics 9 (2008) 406–411. URL: https://www.nature.com/articles/nrg2360. doi:10.1038/nrg2360.

[40] DIW Berlin, Research Infrastructure 'Socio Economic Panel (SOEP)', 2025. URL: https://www.diw.de/en/diw_01.c.615551.en/research_infrastructure__socio-economic_panel__soep.html.

[41] GO FAIR, FAIR Principles, 2025. URL: https://www.go-fair.org/fair-principles/.

[42] ACM, Contributions to CHI – CHI 2025, 2025. URL: https://chi2025.acm.org/contributions-to-chi/.

[43] J. Oppenlaender, S. Malacria, X. Fang, N. van Berkel, F. Chevalier, K. Yatani, S. Hosio, Meta-HCI: First Workshop on Meta-Research in HCI, CHI EA '25, Yokohama, Japan (2025). URL: https://meta-hci.github.io/. doi:10.1145/3706599.3706723.