«Comunicazioni sociali», 2023, n. 2, 191-202
Published online October 2023

ARNE FREYA ZILLICH - DANIELA SCHLÜTZ - EVA-MARIA ROEHSE -
ELENA LINK - WIEBKE MÖHRING*

PRACTICING INTERNET RESEARCH ETHICS: CHALLENGES AND SOLUTIONS FROM A GERMAN PERSPECTIVE

Abstract
Good research practice faces both ethical and methodological challenges that cannot always be met at the same time. This is also true for the study of digital trace data and digital media. However, questions of Internet research ethics (IRE) have so far played a subordinate role in German media and communication studies, especially regarding quantitative methods. Based on 24 semi-structured interviews with German media and communication scholars with experience in the field of (quantitative) Internet research, the current study explores perceived ethical challenges and proposed solutions. We identified two central themes, mirroring the theoretical debate about (1) privacy and informed consent and (2) minimizing harm. Overall, our results show that researchers have developed casuistic strategies to address the various ethical challenges in studying the Internet. These strategies are often based on practical research experiences rather than training or available information. This indicates the importance of having an ongoing discourse about IRE, integrating research ethics training into higher education, and providing practical recommendations.

Keywords
Internet research; digital trace data; digital media; semi-structured interviews; Internet research ethics (IRE).
ISSN: 03928667 (print) 18277969 (digital)
DOI: 10.26350/001200_000185
Creative commons license CC-BY-NC-ND 4.0

1. INTRODUCTION AND BACKGROUND

"There is little research that is not impacted in some way on or through the Internet"[1]. Buchanan and Zimmer's statement is, especially true for the field of media and communication research: digital data and digital methods shape the research practice of current

* Arne Freya Zillich, Film University Babelsberg KONRAD WOLF, Germany – a.zillich@filmuniversitaet.de; – Daniela Schlütz, Film University Babelsberg KONRAD WOLF, Germany – d.schluetz@filmuniversitaet.de; Eva-Maria Roehse, Technical University Dortmund, Germany – eva-maria.roehse@tu-dortmund.de; Elena Link, Johannes Gutenberg University Mainz – elena.link@uni-mainz.de; Wiebke Möhring, Technical University Dortmund, Germany – wiebke.moehring@tu-dortmund.de.

The authors thank all interviewees for their detailed and intriguing insights into their research practices. The study was part of the project "Forschungsethik in der Kommunikations- und Medienwissenschaft [Research Ethics in Communication and Media Science]" funded by the German Federal Ministry of Education and Research (BMBF) under grant 16PH20009A and 16PH20009B.

[1] E.A. Buchanan, M. Zimmer, "Internet Research Ethics", *Stanford Encyclopedia of Philosophy*, 2021. https://plato.stanford.edu/entries/ethics-internet-research.

© 2023 Vita e Pensiero / Pubblicazioni dell'Università Cattolica del Sacro Cuore

media and communication studies fundamentally[2]. Digital methods offer researchers new technical opportunities to collect, analyze, and process large amounts of digital data, such as user comments, search engine enquiries, or interactions on social media platforms[3]. These data provide valuable insights into social communication phenomena[4]. However, tapping into new research fields, such as big data or computational communication science, raises existing as well as new ethical questions[5], which IRE address. Internet research ethics are defined as the analysis of ethical issues and the application of IRE principles as they pertain to research conducted on and in the Internet[6]. IRE deals with research that uses the Internet as a field, a tool, or a space of research, and revolves around ethical issues such as minimizing harm, obtaining informed consent, and protecting subject privacy and confidentiality[7]. A key debate within IRE is the question concerning which digital spaces and data may be considered private, and which public[8]. Defining the Internet as a vast public space assumes not only random encounters of users and researchers who happen to be there at the same time but also releases researchers from having to decide whether to make their presence known and seek informed consent[9]. However, the argument of 'the data is already public' simplifies the complex privacy implications in IRE[10]. Instead, any discussion about issues of privacy may better be based not only on users' expectations of privacy but also their perception of imagined audiences[11]. A widely cited approach to determine what users will perceive as threats to their privacy is Nissenbaum's framework of contextual integrity[12], which is a useful heuristic to guide ethical decision-making in IRE. The framework "help[s] identify and explain why certain patterns of information flow are acceptable in one context, but viewed as problematic in another"[13]. However, researchers may find implementing this framework in research practice is a challenge. As Tiidenberg

[2] A. Hepp, W. Loosen, U. Hasebrink, "Jenseits des Computational Turn: Methodenentwicklung und Forschungssoftware in der Kommunikations- und Medienwissenschaft", *Medien & Kommunikationswissenschaft*, 69, 1 (2021): 3-24 (7). Accessed January 23, 2023. DOI: 10.5771/1615-634X-2021-1-3-1.

[3] M. Rogers, *Doing digital methods*, London: Sage, 2019, XVIII.

[4] S. Geise, A. Waldherr, "Computational Communication Science", in *Handbook of Computational Social Science*, edited by U. Engel *et al.*, London: Routledge, 2021: 66-82 (67). Accessed January 23, 2023. DOI: 10.4324/9781003024583-6.

[5] A.O. Larsson, "Studying Big Data – Ethical and Methodological Considerations", in *Internet Research Ethics*, edited by H. Fossheim and H. Ingried, Oslo: NOASP, 2016, 141-156 (141). Accessed January 23, 2023. DOI: 10.17585/noasp.3.1.

[6] Buchanan, Zimmer, *Internet Research Ethics*; a.s. franzke *et al*., *Internet Research: Ethical Guidelines 3.0*, 2020, 9. Accessed January 23, 2023. https://aoir.org/reports/ethics3.pdf.

[7] Buchanan, Zimmer, *Internet Research Ethics.*

[8] K. Tiidenberg, "Ethics in Digital Research", in *The SAGE Handbook of Qualitative Data Collection*, edited by U. Flick, London: Sage, 2018: Section "Public or Private". Accessed April 21, 2023. DOI: 10.4135/9781526416070; E. Locatelli, "Ethics of Social Media Research: State of the Debate and Future Challenges", in *Second International Handbook of Internet Research*, edited by J. Hunsinger, L. Klastrup, M. Allen, Dordrecht: Springer, 2018: Section "Public, Private, Privacy". Accessed April 21, 2023. DOI: 10.1007/978-94-024-1202-4_25-1.

[9] Tiidenberg, "Ethics in Digital Research", Section "Public or Private".

[10] M. Zimmer, "'But the Data Is Already Public': On the Ethics of Research in Facebook", *Ethics and Information Technology*, 12 (2010): 313-325 (323). Accessed April 21, 2023. DOI: 10.1007/s10676-010-9227-5.

[11] A. Markham, E. Buchanan, *Ethical Decision-Making and Internet Research 2.0: Recommendations from the AoIR Ethics Working Committee*, 2012. Accessed April 21, 2023. http://aoir.org/reports/ethics2.pdf.; K. Tiidenberg, "Research Ethics, Vulnerability, and Trust on the Internet", in *Second International Handbook of Internet Research*, edited by J. Hunsinger, L. Klastrup, M. Allen, Dordrecht: Springer, 2018: Section "Problems with Existing Research Ethics Guidance". Accessed April 21, 2023. DOI: 10.1007/978-94-024-1202-4_55-1.

[12] H. Nissenbaum, *Privacy in Context. Technology, Policy, and the Integrity of Social Life*, Stanford: Stanford University Press, 2010, 1-17.

[13] M. Zimmer, "Addressing Conceptual Gaps in Big Data Research Ethics: An Application of

points out, researchers cannot always ask users about their expectations or assume that those expectations are stable or informed[14].

Thus, good methodological practice includes constant ethical reflection and resolution of dilemmas that almost inevitably arise in the many stages of a research project[15]. By the same token, such ethical reflection may well result in a better research design[16]. The process of ethical deliberation is an ongoing case-based approach[17] guided by *prima facie* ethical principles like respect for autonomy, non-maleficence, beneficence, and justice[18]. There are cases, however, in which good research practice faces ethical and methodological challenges that cannot be met simultaneously.

While such an ongoing discourse on IRE has been institutionalized and integrated into research practice in international, especially U.S., research[19], corresponding questions have so far played a subordinate role in German media and communication studies and especially in the context of quantitative methods. However, in light of methodological developments (such as computational methods[20]) and recent discussions about controversial studies (such as the Facebook contagion experiment, the OKCupid study, or the Cambridge Analytica scandal)[21], IRE is becoming increasingly important in the German scientific community. In addition, the European approach to legislation has historically been designed to ensure the individual's "right to privacy"[22], making data protection an important issue for German citizens and researchers. The German context we consider in this study therefore offers additional insights into the cultural adaptation of existing IRE practices. Against this background, the present article asks: Which ethical challenges do German media and communication scholars perceive with regard to (quantitative) Internet research and how do they deal with them in the research process?

2. METHOD

To answer the research question, we will report findings from semi-structured interviews with scholars with various qualification levels from (quantitative) media and communication research in German-speaking countries. In total, 31 scholars were inter-

Contextual Integrity", *Social Media + Society*, 4, 2 (2018): 1-11 (6). Accessed January 26, 2023. DOI: 10.1177/2056305118768300.

[14] Tiidenberg, "Ethics in Digital Research", Section "Public or Private".

[15] franzke *et al., Internet Research: Ethical Guidelines 3.0*, 9.

[16] C.M. Ess, "Internet Research Ethics and Social Media", in *Handbook of Research Ethics and Scientific Integrity*, edited by R. Iphofen, Cham: Springer Nature Switzerland, 2019, 283-303 (296). Accessed January 23, 2023. DOI: 10.1007/978-3-319-76040-7_12-1.

[17] H.A. McKee, J.E. Porter, *The Ethics of Internet Research: A Rhetorical, Casebased Process*, New York, Peter Lang, 2009.

[18] T.L. Beauchamp, J.F. Childress, *Principles of Biomedical Ethics*, Oxford: Oxford University Press, 2013, 25.

[19] Buchanan, Zimmer, *Internet Research Ethics*; C.M. Ess, *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee*, 2002, Accessed April 21, 2023. https://aoir.org/reports/ethics.pdf.; Markham, Buchanan, *Ethical Decision-Making and Internet Research 2.0: Recommendations from the AoIR Ethics Working Committee.*

[20] Geise, Waldherr, "Computational Communication Science", 67.

[21] J. Isaak, M.J. Hanna, "User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection", *Computer*, 51, 8 (2018): 56-59; E.O. Kirkegaard, J.D. Bjerrekær, "The OKCupid Dataset: A Very Large Public Dataset of Dating Site Users", *Open Differential Psychology*, 2016. Accessed April 21, 2023. DOI: 10.26775/odp.2016.11.03.

[22] P. Elias, "A European Perspective on Research and Big Data Analysis", in *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, edited by J. Lane *et al.,* Cambridge: Cambridge University Press, 2014, 173-191.

viewed. We used a purposive sampling strategy. Our goal was to recruit a broad range of scholars with expertise across research fields (such as political communication, journalism, media use and effects), methods (such as surveys, content analysis, observation, psychophysiological measurements), and study designs (such as experimental studies, comparative studies, longitudinal studies). Based on these screening characteristics, we created a pool of participants and ranked them according to the relevance of their expertise in either the respective method or research field.

In this analysis, we focus on a subsample of 24 researchers with personal experience in the field of Internet research. We chose expert interviews to gather data as this method is "geared to the reconstruction of specific knowledge stocks"[23], particularly if the experts serve as "'crystallization points' for practical insider knowledge and are interviewed as surrogates for a wider circle of players"[24]. The participants had an average age of 47 years and a gender ratio of 19 (identifying as male) and 5 (as female). This ratio represents the current academic situation in Germany. On a qualification level, there were 15 professorships, 8 post-doctoral positions, and 1 was a PhD student and while 22 worked at a university, 2 had positions in non-university research organizations (see Table 1). The interviews lasted 60 minutes on average (range: 33 to 82 minutes). They were conducted via Zoom between April and July 2021. All participants were informed about the study goals and gave their informed consent prior to participation.

Table 1 - *Participants' sociodemographic characteristics*

Case number	Age group (years)	Gender	Qualification Level	Institutional Affiliation
#7	30-40	male	Post-Doc	Non-university research organization
#9	40-50	female	Professor	University
#25	50-60	male	Professor	University
#27	50-60	male	Professor	University
#28	40-50	female	Professor	University
#37	40-50	male	Professor	University
#38	40-50	male	Post-Doc	University
#42	40-50	male	Professor	University
#47	40-50	male	Post-Doc	University
#56	30-40	female	Post-Doc	University
#60	30-40	female	Post-Doc	University
#64	50-60	male	Professor	University
#65	30-40	male	PhD student	University
#68	30-40	male	Professor	University
#71	50-60	male	Professor	University
#75	40-50	male	Professor	University
#77	40-50	male	Professor	University
#79	40-50	male	Post-Doc	Non-university research organization
#81	40-50	male	Post-Doc	University
#83	50-60	male	Professor	University
#86	50-60	male	Professor	University
#107	40-50	female	Professor	University
#109	60+	male	Professor	University
#110	60+	male	Professor	University

[23] M. Pfadenhauer, "At Eye Level: The Expert Interview – A Talk between Expert and Quasi-Expert", in *Interviewing Experts. Research Methods Series*, edited by A. Bogner, B. Littig, W. Menz, London: Palgrave Macmillan, 2009, 81-97 (81). Accessed January 23, 2023. DOI: 10.1057/9780230244276_4.

[24] A. Bogner, B. Littig, W. Menz, "Introduction: Expert Interviews – An Introduction to a New Methodological Debate", in *Interviewing Experts. Research Methods Series*, edited by A. Bogner, B. Littig, W. Menz, London: Palgrave Macmillan, 2009: 1-13 (2). Accessed January 23, 2023. DOI: 10.1057/9780230244276_1.

The interview guideline addressed seven main topics such as personal experiences, and assessments of research ethics and perceived challenges in empirical (quantitative) research projects, their relevance to academic publishing, third-party funding, and teaching, as well as the interviewees' frame of reference with regard to ethical questions (i.e. their sources of information and discourse formats). The interviews were audio-recorded, transcribed, coded and analyzed using MAXQDA software[25]. For this publication, all names were replaced by case numbers (marked with #) to protect participants' privacy. We performed a thematic analysis of the text to identify and describe themes raised by participants[26]. The coding was carried out deductively-inductively[27]: We started from a coding scheme that was derived from the conceptual framework and theoretical background of the study. The coding scheme comprised nine categories: 1) challenges in the research process, 2) standards for dealing with these challenges, 3) predicaments, 4) relevance of research ethics in academic publishing, 5) teaching, 6) supervision of students, and 7) third-party funding, 8) participants' frame of reference, as well as 9) development perspective for the German scientific community. Applying this scheme, a line-by-line coding of each transcript was conducted to identify common themes among participants. In the course of the analysis, subcategories were added to the main categories and some were merged. Afterwards, we selectively retrieved and organized the coded transcript segments to compare them and identify recurring patterns[28]. For this article, we focused on the challenges that the interviewed researchers perceived when studying digital trace data and digital media as well as the solutions they pursued.

3. RESULTS

The interviewees very vividly described the challenges of studying digital trace data and digital media and how they dealt with them. In the analysis, we identified two central themes, mirroring the theoretical debate about 1) privacy and informed consent and 2) minimizing harm. In the following sub-sections, we will describe each of these challenges and detail the practices used to address them. We will illustrate our interpretation with quotes in the original wording (translated from German and referenced by case number).

3.1. *Privacy and informed consent*

The results show that the public-private debate was also a central aspect of IRE in Germany. For each research project, the interviewees deliberated which spaces should be

[25] U. Kuckartz, S. Rädiker, *Analyzing Qualitative Data with MAXQDA. Text, Audio, and Video*, Cham: Springer, 2019. Accessed April 19, 2023. DOI: 10.1007/978-3-030-15671-8.

[26] V. Braun, V. Clarke, "Using Thematic Analysis in Psychology", *Qualitative Research in Psychology*, 3, 2 (2008): 77-101; U. Flick, *An Introduction to Qualitative Research*, London: Sage, 2018, 474-475; M. Vaismoradi, S. Snelgrove, "Theme in Qualitative Content Analysis and Thematic Analysis", *Forum Qualitative Social Research*, 20, 3 (2019). Accessed April 19, 2023. DOI: 10.17169/fqs-20.3.3376.

[27] J. Gläser, G. Laudel, "Life With and Without Coding. Two Methods for Early-Stage Data Analysis in Qualitative Research Aiming at Causal Explanations", *Forum Qualitative Social Research*, 14, 2 (2013). Accessed December 08, 2022. DOI: 10.17169/fqs-14.2.1886.

[28] M.B. Miles, A.M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, Thousand Oakes: Sage, 1994, 55-58.

considered private and which could be assessed as public forums. This decision was critical in determining whether informed consent was considered a necessity or an add-on. Some participants regarded social media as a new form of public sphere comparable to a traditional marketplace: "The cathedral square of [a city] is also an enclosed space, so if you stand there and say 'I'm against equal rights of men and women', you have to expect people to notice" (#64). Accordingly, these researchers treated openly accessible data, e.g., from Twitter or Facebook as public statements or behavior. Therefore, they considered it ethical to use and publish these data within the framework of applicable laws, i.e. within the General Data Protection Regulation (GDPR) and each platform's terms of service (ToS). Other participants took a more deliberative stance. For example, they argued that "what is being communicated on social media is public in the sense of 'accessible', but that does not mean public in the sense of 'meant for the public'" (#79). Hence, researchers should ask themselves whether Internet users really do assume their utterances can be read and analyzed by other people who do not participate in the chat on the platform, forum, or community. Researchers who concluded that Internet users do not expect to be studied were consciously accepting that their research was violating users' privacy, "because it always violates the contextual integrity of the statements I am examining" (#38). This varying understanding of public vs. private spheres would particularly become apparent in an interdisciplinary research group: Some participants pointed out that in computer science, for example, it is more common to think of web texts or web links as publicly available data, whereas in social science they are more commonly understood as data from human subjects. As a result, many participants had to rethink how anonymity can be granted to protect Internet users from being identified, while still being able to publish meaningful results. One participant summed up this challenge when they asked:

> How do you manage to make it anonymous, to make it legally and ethically secure, so that the people involved can work with it, and so that other scientists, including external scientists, can be involved without drawing conclusions about individuals? (#38).

This challenge is exacerbated by the structural characteristics and sensitivity of digital data. Participants mentioned that full-text data per se pose a risk for identifying individuals. In the case of text data or video data, intellectual property and copyright issues also arise. Furthermore, participants addressed their concern that data can be linked or de-anonymized. They anticipated that technological advances in computational communication science will make it even easier to combine data, such as digital network data and location data and draw further conclusions about individual users that they would likely not have consented to:

> If you master these tools, it is indeed very easy to collect a lot of data about a lot of users and to aggregate information from different sources about one person. And by linking information, you undermine contextual integrity because you are bundling things that, from the people's point of view, belong in two different worlds (#38).

The interviewees therefore did not consider it sufficient that users had agreed to the ToS of each platform. Users' consent to the ToS of a platform should not be equated with consent to participate in a research project. Instead, gaining explicit informed consent was seen as essential – particularly on social media platforms, when studying egocentric networks, or when using mobile tracking data. The claim to be "as transparent as possible" (#68) includes, among other things, informing users which digital data traces

have been tracked and which have not, and which third parties, such as market research companies, have been involved in the process.

However, several researchers who applied big data approaches and automated analyses of digital trace data stated that it is not feasible to gain consent from each individual user; the same is true for consent from uninvolved third parties, such as friends featured in users' photos or linked in comments. Therefore, these participants often regarded public accessibility of data as equivalent to (implicit) informed consent. Others distinguished more clearly between private and (semi-)public spaces, or between private individuals and public figures. For communication between private individuals, e.g., in messenger chat groups, the researchers considered it necessary to obtain informed consent. By contrast, consent was assumed for public figures in the digital sphere, such as politicians, journalists, members of political parties, or media representatives. When individual consent could not be gained, one participant suggested obtaining proxy consent by contacting the platform moderators and asking them to inform the community about the proposed study. In doing so, users would be given the possibility to "opt out and say 'I don't want to be included'" (#9). Another participant suggested to obtain granular consent, meaning that users actively give consent for each type of data collection. For example, as participants in an online access panel, users would agree to the terms and conditions presented by the market research company and would be asked for their informed consent for each online survey to which they were invited. Finally, they were asked to share the data of those social media platforms with the research team that they had indicated to use in the online survey.

In addition to obtaining or assuming informed consent, the interviewees employed the following strategies to ensure anonymity and avoid (future) identification of users. Several participants stressed the importance of knowing data protection laws and seeking advice from their institution's data protection officer. The crucial question in this context is: "What is personal data?" (#38). Instead of relying on general assumptions or personal opinions, researchers should consult the GDPR (as well as the specific regulations in the individual German federal states), particularly regarding copyright, personal rights, and right to informational self-determination. This is particularly important in the context of content analyses of digital media. Nowadays, "when we fetch [online] media articles, for example, you're not allowed to distribute them. So, we are allowed to store them, but we are not allowed to distribute them" (#77).

Several participants followed the principle of data economy and refrained from collecting certain information:

> I think that's almost the most important step when you're thinking about the design of the study, it is to think about 'What data do I need?' So, that I really don't collect too much, that I don't collect everything now, and then it is stored somewhere (#107).

Some participants refrained from collecting personal data altogether, especially when studying children or other vulnerable groups. Other researchers preferred to use data donations that "allow people to make quasi self-determined decisions about the data they provide" (#75) or use so-called 'whitelists', to analyze just data traces from certain designated domains or about particular activities to reduce potential privacy violations.

Some of the participants took special precautions for data security. These included encrypting hard drives, not using removable storage devices (e.g., USB sticks) or external communication systems, and giving preference to internal university services or providers from the EU that comply with the GDPR. As an additional organizational safe-

guard, participants stated that every member of their research teams had signed a data protection declaration, a privacy statement, and a confidentiality agreement. Since big data analytics is primarily about pattern recognition, in most cases interviewees were not interested in individual users' online behavior. "Due to the large amount of data, the need and desire to track individual cases is not that great" (#75). Hence, they analyzed digital data only at the aggregate level and provided access to it to other researchers only in aggregate form.

However, scientific publications often use examples or verbatim quotes to illustrate results. To ensure privacy, most participants pseudonymized identifiable information, paraphrased users' online communication, or blacked out images and text segments and especially in the case of public figures, such as politicians, interviewees were careful not to mention the name or refer to the online profile. They also informed the person in question about the study. Another criterion was whether the commenter was being studied in their role as a private individual or as an author. If the researchers' interest was in the author, it may be appropriate to quote the utterance verbatim and by name. "This tweet that I cite as an example of particularly successful aphorisms on Twitter, I also want to reference that it is from XY" (#79).

3.2. *Minimizing harm*

The interviewees made it clear that avoiding doing any harm with their research – to either the participants or members of the research team – was of extreme importance. For instance, many participants found it challenging to deal with (unintentional) exposure to problematic or sensitive content, especially when studying social media. Among the solutions discussed were participants refraining entirely from studying specific research topics, such as child pornography or suicide memes, and researchers prohibiting their students from doing so, since "this should be done by someone who really has a lot of experience" (#9). The interviewees also tried to avoid problematic stimuli in their studies and publications, but if that was impossible, they ensured that study participants were debriefed adequately. Ethical challenges also arose with regard to dual use, i.e. research and technology with the potential of the pairing to yield valuable scientific knowledge and be used for nefarious purposes[29]. For example, automated chatbots designed to combat hate speech in user comments can be trained to replace hurtful expressions with harmless code words. This makes hate speech unrecognizable and encourages its spread. Consequently, the bots contribute to the dissemination of problematic content instead of preventing it. "It's like weapons development. You kind of develop a technology, and it can just be used in a different way. That is already a problem" (#77), for which there are not any established solutions in place. One participant suggested that such tools should be released under a creative commons license so that their code is transparent and may be used by anyone. Even if this does not prevent possible abuse, at least the principle of justice is guaranteed.

Another aspect less often mentioned in theoretical discussions concerned possible harm to the researchers themselves, especially when research assistants and students were involved. Many participants reported that analyzing Internet content could be

[29] S.V. Pustovit, E.D. Williams, "Philosophical Aspects of Dual Use Technologies", *Science and Engineering Ethics*, 16, 1 (2010): 17-31. Accessed January 26, 2023. DOI: 10.1007/s11948-008-9086-1.

stressful when coders were repeatedly exposed to incriminating content, such as violence against women, pornography, or right-wing extremism. “Some of this is such disgusting crap. How long can I actually let a twenty year-old work on it a day?” (#64). As supervisors, the participants had responsibility for their employees.

> With things like this, the question is always how resilient are the research assistants, the student assistants, who work on it? In this field, for example, we have to deal with the fact that we see pornography, that we see swastikas, and theoretically something like beheading videos can also occur. This means that we have an ethical responsibility to the people involved in the project to protect them accordingly (#38).

To minimize harm, respondents took great care in recruiting their research teams. When hiring student assistants, participants carefully selected applicants, attempting to consider their presumed mental stability. “Do I need to look at whether they’re stable or whether they’re interested in it out of some misguided interest? Is the person stable? Of course, this is always a gut decision” (#64). One strategy was to use trigger warnings in order to allow sensitive or vulnerable students to prepare emotionally for disturbing content or to decide to forgo interacting with it altogether by dropping out. Other researchers prepared their teams by proactively addressing potential strains that they might encounter and discussed coping strategies. They suggested appropriate measures to cope with stress, challenging situations, and anxiety. For example, students were informed about the importance of taking breaks from coding and being mindful with themselves. “As part of the project, we also said ‘It’s winter now, it’s dark outside, I’m reading the twentieth article where a woman was killed. I must stop” (#56). To reduce workload and associated stress, some participants scheduled a longer time for coding in their teaching and research projects. For example, rather than insisting on a tight time frame within one semester, students were allowed to move the completion of coding to the semester’s break. If needed, assistants were offered professional help, such as psychological counseling or supervision. In addition, several participants emphasized that they themselves acted as low-threshold contacts for the student assistants involved. This included offering appointments for talks at short notice and finding needs-based solutions for individual challenges.

4. DISCUSSION

Overall, our results show that German media and communication scholars face two key challenges in conducting Internet research: first, respecting respondents’ privacy by obtaining informed consent, and second, protecting all involved parties in a research project from harm. These findings are largely comparable to the existing debate on IRE described above[30]. In Germany, there are extensive public and private debates, framed by the legal particularities of the GDPR. Deciding whether obtaining informed consent in quantitative research projects is necessary or not, the interviewees argued on two levels: 1) technical (im)possibilities and 2) users’ alleged attitudes, which concerning privacy are highly context-dependent[31]. For example, previous studies have shown that

[30] Buchanan, Zimmer, *Internet Research Ethics.*
[31] Nissenbaum, *Privacy in Context*, 1-17.

users often have an "imagined audience"[32] that, while not necessarily corresponding to reality, does not include researchers who use their data. In addition, users' privacy attitudes depend on factors such as how the study is conducted or disseminated, who is conducting it, and what the study is about[33]. Therefore, McKee and Porter propose a heuristic approach that takes the specific contexts of a particular study into account and helps researchers to decide whether they should seek informed consent. It represents several continua that a researcher might consider, including public vs. private spheres[34]. However, when obtaining informed consent is simply not possible, researchers can take other precautions, such as a consistent anonymization, to protect users' privacy.

Another aspect that frequently came up in the interviews was the importance of the "do no harm" principle. This was applied both to participants in studies and members of research teams. The interviewees cared strongly about the well-being of their teams, especially student assistants, and took appropriate precautions. While safe-guarding users' harm is a basic principle in research ethics, protecting the researchers themselves is less frequently discussed[35]. However, this is a vital part of research ethics, particularly regarding content analyses of social media[36].

The findings of our study are in line with Favaretto *et al.* who conclude that scholars generally consider traditional principles of ethics to be an appropriate guide for conducting ethical research, but at the same time have recognized the challenges associated with their practical application[37]. The above insights illustrate how these challenges are overcome, mainly on a case-by-case basis. The interviewed scholars have developed casuistic strategies to address the various ethical challenges in studying the Internet. However, the strategies described were often based on practical research experiences rather than professional training or available information. The interviewees felt that it was primarily the individual researcher's responsibility to ensure ethical conduct, with only little support from the institution or the scientific community as a whole[38]. Considering the complexities of research conducted on and in the Internet – both in terms of method and ethics – this is unsatisfactory.

4.1. *Limitations*

While our study contributes to a better understanding of the ethical challenges and solutions that German media and communication scholars perceive in Internet research, it has some limitations. The main limitation relates to the homogeneity of the sample. Since most participants held a professorship, the results cannot be generalized to academics of all qualification levels. Early career scholars may perceive different challeng-

[32] E. Litt, "Knock, Knock. Who's there? The Imagined Audience", *Journal of Broadcasting & Electronic Media*, 56, 3 (2012): 330-345. Accessed May 2, 2023. DOI: 10.1080/08838151.2012.705195.

[33] C. Fiesler, N. Proferes, "Participant Perceptions of Twitter Research Ethics", *Social Media + Society*, 4, 1 (2018). Accessed May 2, 2023. DOI: 10.1177/2056305118763366.

[34] McKee, Porter, *The Ethics of Internet Research*, 86-90.

[35] But see: N. Podschuweit, "How Ethical Challenges of Covert Observations Can Be Met in Practice", *Research Ethics*, 17, 3 (2021). Accessed May 2, 2023. DOI: 10.1177/17470161211008218.

[36] S.R. Stern, "Encountering Distressing Information in Online Research: A Consideration of Legal and Ethical Responsibilities", *New Media & Society*, 5, 2 (2003): 249-266. Accessed May 2, 2023. DOI: 10.1177/1461444803005002006

[37] M. Favaretto *et al.*, "First Do No Harm: An Exploration of Researchers' Ethics of Conduct in Big Data Behavioral Studies", *PLoS One*, 15, 11 (2020). Accessed May 2, 2023. DOI: 10.1371/journal.pone.0241865.

[38] See also *ibid*.

es, because they are less experienced. On the other hand, recent changes in education may have made them more familiar with IRE. Furthermore, our sample was rather small and consisted only of media and communication scholars in German-speaking countries. Further research in other cultural or academic contexts is needed to better understand the influence of the legal, institutional, and normative frameworks underlying ethical decision making in studying the Internet. Another limiting aspect is the focus on ethical challenges in quantitative Internet research, as qualitative research might have other challenges to overcome[39]. Therefore, further studies with a more heterogeneous sample of scholars from different career levels, cultures and both methodological and epistemological backgrounds are needed. Finally, a methodological limitation is that academic scholars tend to make generalizing statements in expert interviews and make assessments that are important for their public image[40]. We attempted to address the issue of performativity[41] through a cautious interpretation.

4.2. *Practical implications*

Despite these limitations, our study offers an important insight into the dynamic and challenging field of IRE from a German perspective. The study shows that a constant balancing act is required between the technically feasible and the ethically acceptable. To achieve this balance, a principle based casuistic approach, guided by an "ethics of care", seems to be helpful[42]. For example, franzke *et al.* propose the Data Ethics Decision Aid to guide case-deliberation and advance the development of responsible data practices[43]. Moreover, an ongoing discourse around these questions is crucial – both within the broader scientific community and the single research team. Furthermore, it is vital to integrate research ethics training into higher education to immunize researchers against questionable research practices[44]. The great wealth of theoretical concepts and research approaches developed in the study of the Internet can guide researchers on key ethical issues[45]. Our findings underscore the need for practical recommendations like the ones mentioned above. It would be useful to develop even more low-threshold materials – such as best-practice examples, checklists, tutorials, or short videos – that encourage

[39] M. Guillemin, L. Gillam, "Ethics, Reflexivity, and 'Ethically Important Moments' in Research", *Qualitative Inquiry*, 10, 2 (2004): 261-280 (262). Accessed December 08, 2022. DOI: 10.1177/1077800403262360; M.K.E. Lahman, *Ethics in Social Science Research: Becoming Culturally Responsive*, Thousand Oakes: Sage, 2017.

[40] J. Gläser, G. Laudel, *Experteninterviews und Qualitative Inhaltsanalyse als Instrumente Rekonstruierender Untersuchungen*, Wiesbaden: Springer, 2009, 181.

[41] R. Diaz-Bone, "The Performativity of Qualitative Social Research", *Forum Qualitative Social Research*, 12, 3 (2011). Accessed December 08, 2022. DOI: 10.17169/fqs-12.3.1750.

[42] Tiidenberg, "Research Ethics, Vulnerability, and Trust on the Internet", Section "How Not to Be an Asshole?".

[43] a.s. franzke, I. Muis, M.T. Schäfer, "Data Ethics Decision Aid (DEDA): A Dialogical Framework for Ethical Inquiry of AI and Data Projects in the Netherlands", *Ethics and Information Technology*, 23, 3 (2021): 551-567. Accessed May 2, 2023. DOI: 10.1007/s10676-020-09577-5.

[44] N.S.L. Yeo-The, B.L. Tang, "Research Ethics Courses as a Vaccination against a Toxic Research Environment or Culture", *Research Ethics*, 17, 1 (2021): 55-65. Accessed May 2, 2023. DOI: 10.1177/1747016120926686.

[45] See for example: M.W. Kalichman, M.L. Devereaux, D.K. Plemmons, "A Course for Teaching and Learning about the Responsible Conduct of Research", *Journal of Empirical Research on Human Research Ethics*, 17, 3 (2022): 284-291. Accessed May 2, 2023. DOI: 10.1177/15562646211047163.

self-reflection and guide ethical decision making in Internet research[46]. We do not advocate rigid catalogs of solutions as standardized procedural research ethics are inadequate[47], but recommend guidelines that include relational aspects and, which are applied on a case-by-case basis to achieve continuous ethical reflection in the research process[48]. Furthermore, we consider it important to discuss these reflection processes transparently (e.g., at conferences or in publications) in order to increase awareness of ethical challenges and their methodological consequences, as well as to initiate exchange and mutual support within the scientific community.

[46] The project website https://www.forschungsethik-kmw.de/en is a first step towards such a collection.
[47] Tiidenberg, "Research Ethics, Vulnerability, and Trust on the Internet", Section "Abstract".
[48] franzke *et al*., *Internet Research: Ethical Guidelines 3.0*, 4.