Privacy Concerns and Acceptance Factors of OSINT for Cybersecurity: A Representative Survey

Thea Riebe
Technical University of Darmstadt
Darmstadt, Hessen, Germany
riebe@peasec.tu-darmstadt.de

Tom Biselli
Technical University of Darmstadt
Darmstadt, Germany
biselli@peasec.tu-darmstadt.de

Marc-André Kaufhold
Technical University of Darmstadt
Darmstadt, Germany
kaufhold@peasec.tu-darmstadt.de

Christian Reuter
Technical University of Darmstadt
Darmstadt, Germany
reuter@peasec.tu-darmstadt.de

ABSTRACT

The use of Open Source Intelligence (OSINT) to monitor and detect cybersecurity threats is gaining popularity among Cybersecurity Emergency or Incident Response Teams (CERTs/CSIRTs). They increasingly use semi-automated OSINT approaches when monitoring cyber threats for public infrastructure services and incident response. Most of the systems use publicly available data, often focusing on social media due to timely data for situational assessment. As indirect and affected stakeholders, the acceptance of OSINT systems by users, as well as the conditions which influence the acceptance, are relevant for the development of OSINT systems for cybersecurity. Therefore, as part of the ethical and social technology assessment, we conducted a survey (N=1,093), in which we asked participants about their acceptance of OSINT systems, their perceived need for open source surveillance, as well as their privacy behavior and concerns. Further, we tested if the awareness of OSINT is an interactive factor that affects other factors. Our results indicate that cyber threat perception and the perceived need for OSINT are positively related to acceptance, while privacy concerns are negatively related. The awareness of OSINT, however, has only shown effects on people with higher privacy concerns. Here, particularly high OSINT awareness and limited privacy concerns were associated with higher OSINT acceptance. Lastly, we provide implications for further research and the use of OSINT systems for cybersecurity by authorities. As OSINT is a framework rather than a single technology, approaches can be selected and combined to adhere to data minimization and anonymization as well as to leverage improvements in privacy-preserving computation and machine learning innovations. Regarding the use of OSINT, the results suggest to favor approaches that provide transparency to users regarding the use of the systems and the data they gather.

KEYWORDS

cybersecurity, OSINT, online social networks, privacy, surveillance

1 INTRODUCTION

Open Source Intelligence (OSINT) is considered to be one of the most promising approaches to fight crime and corruption. It is a framework that consists of using publicly available data that is collected, processed, and correlated to provide timely information, e.g. for cyber situational awareness [40], or for investigative research and journalist teams, like Bellingcat. OSINT has also been used and has been specialized to detect cyber crime and cyber threats worldwide, using a semi-automated process [38]. When social media data is used exclusively, OSINT is referred to as Social Media Intelligence (SOCMINT).

The monitoring and crisis management by emergency services has been studied in the field of crisis informatics [44, 45]. The approach aims to use public data to gain situational awareness and provide effective incident prevention and response to improve public security in crisis situations. In the case of cybersecurity, (governmental) Computer Emergency Response Teams (CERTs), also known as Computer Security Incidents Response Teams (CSIRTs), have been adapting this approach from other emergency services and government agencies [48]. CERT members collect information on potential cyber threats from different public sources like Twitter, vulnerability databases and software vendor websites to gain situational awareness. This process is increasingly (semi)automated [26].

With growing numbers of cyber threats, OSINT has become an increasingly important approach, as more data is available, which can be used for early risk prevention. As OSINT approaches to detect cyber threats mostly use social media data [38, 49], many ethical and social questions arise at the complex intersection of privacy and security. In a systematic study of OSINT systems for cybersecurity, Riebe et al. [46] have found that in 73 OSINT systems, only 11 discussed ethical and social implications, such as privacy impact. Therefore, the principles of "privacy by design," such as data minimization, must first be applied when using such tools to collect and analyze any individual's data.

Such OSINT systems can be used to detect novel threats or to identify and profile individuals or groups. Profiling in the case of targeted advertising can have dramatically different consequences than profiling conducted by governments, law enforcement or emergency services. Studies have shown, that citizens express the want or need for government agencies to perform democratically legitimized forms of "surveillance-oriented security technologies"

Proceedings on Privacy Enhancing Technologies 2023(1), 477–493

https://doi.org/10.56553/popets-2023-0028

(SOSTs) to ensure protection from harm [12]. The kind of and the extent to which online surveillance is considered appropriate has been studied with regard to culture [13], trust in the government [28, 52], and fear of terrorism and crime [18, 55], as well as in relation to specific technologies, such as private communication, financial data, and camera use in public spaces [54, 58]. However, the acceptance of SOSTs is not static and can change due to technological or contextual events.

The effects of surveillance on people are described by Lyon's work on the surveillance society [32] and by Haggerty and Ericson [20] as the creation of data doubles that can be monitored. With more and more areas being represented digitally, the rhizomatic spread of surveillance into all spheres of daily life has shifted security and privacy norms [27]. On the one hand, people are aware of their privacy and advocate for privacy enhancement; on the other hand, they view surveillance as a necessary method to prevent harm, manage human-made or natural disasters or respond to incidents affecting public infrastructure. Due to this ambiguity in privacy concerns, the study of ethical, legal, and social implications (ELSI) has become increasingly relevant.

While this study focuses on the case of OSINT for cybersecurity for German governmental CERTs in particular, the ethical implications transcend this case and can therefore be used for other OSINT systems which are used for the public security. Focusing on these implications supports the development of technologies that anticipate the complex and non-binary nature of privacy and security regarding surveillance technologies [55]. Therefore, participatory approaches, such as ELSI-co design [30] and value-sensitive design (VSD) [17], offer research frameworks that include indirect and direct stakeholder perspectives in the design and development of technologies. Thus, this study aims at understanding the factors associated with the acceptance of OSINT systems for cybersecurity contexts to inform the design of these systems.

Therefore, this study's leading research question asks: **How do people evaluate the use of OSINT for cybersecurity by governmental organizations and which factors are associated with the acceptance of OSINT?**

Our results indicate that cyber threat perception and the perceived need for OSINT are positively correlated to acceptance, while privacy concerns show a negative correlation. The awareness of OSINT as an interactive factor only affects the association between privacy concerns and OSINT acceptance. Privacy behavior shows no correlation with OSINT acceptance, which shows that additional research needs to be done on contextual factors for privacy-decision making and data disclosure.

This paper is structured as follows: Section 2 discusses related studies on surveillance technologies, factors for their acceptance and privacy-preserving technologies and behavior. In section 3, the design and methodology of the representative survey of the German population as well as the construction of the concepts are explained. Afterwards, section 4 presents the results of the descriptive and statistical analysis. Section 5 discusses the results regarding the state of research and section 6 provides a comprehensive conclusion.

2 RELATED WORK

The related work section introduces the discourse on the acceptance of surveillance technologies (2.1) as well as on privacy concerns and behaviors regarding the use of OSINT for cybersecurity (2.2). The subsections introduce the concepts used in the study to answer the research question, on which basis, the hypotheses are developed (2.3).

2.1 Acceptance of Surveillance Technologies

Surveillance technologies, such as OSINT, have been researched in the context of public security to monitor terrorists and criminal groups. Their capability to monitor the public on a large scale, as well as existing cases thereof have also been researched [10, 39]. Surveillance can be defined as the systematic monitoring of individuals or groups for a given purpose [32, 52]. In the case of OSINT for cybersecurity, there are two kinds of systems: Cyber Threat Intelligence (CTI), which focuses on the detection and analysis of cyber threats, and risk mitigation systems, which work towards identifying actors and groups [46].

The legitimacy of SOSTs depends on the acceptance and the public approval [52]. Many scholars agree that the perception of threats is associated with the acceptance of the use of surveillance technologies [18, 52]. In this context, threat perception is conceptualized as the fear of crime or terror. For example, In their representative telephone survey with 2.176 participants, Trüdinger and Steckermeier [52] investigated how information is associated with trust in different institutions (legislative, executive, and judicature) and how this might affect the acceptance of surveillance technologies. They concluded that among other factors, threat perception and the perception of surveillance measures' effectiveness correlate with the acceptance of the same category. On the other side, the protection against these threats is also used to legitimize the application of surveillance technologies [5].

Furnham and Swami‘s [18] study shows that attitudes towards the government and public authorities are associated with the acceptance of surveillance measures. They developed a scale of 25 items to test attitudes towards surveillance based on personality traits and punitive attitudes, such as the threat perception and the attitude towards authorities. They have further shown that "demographic and personality factors were weakly related to attitudes to surveillance while general attitudes to authority were the strongest predictor" [18].

The prerequisites for the acceptance of surveillance measures can also be linked to various threats, which the COVID-19 pandemic has shown. For example, Ioannou and Tyssyaduah [23] studied the acceptance of surveillance and privacy protection behaviors during the global health crisis in the US. In accordance with Trüdinger and Steckermeier [52], they found that trust in the government and the need for proactive surveillance are positively associated with acceptance.

Threat perception and attitudes towards the government and authorities are also associated with the perceived need for surveillance measures. People with a higher threat perception are more likely to support government surveillance technologies [23, 52]. Trust in authorities is positively correlated to the increased acceptance of surveillance measures [18, 52]. Therefore, people who perceive the need for surveillance technologies and trust their government are more inclined to accept and support measures on this issue [12]. As norms and threat perceptions change, Wilton [57] argues, that there is a shift of the threat perception from a focus on predominantly commercial threats "to a recognition that government activities, in the sphere of intelligence and national security, also give rise to significant privacy risk". Thus, people might perceive threats different with regard to their culture or experiences [12, 13, 57]. As a result, people might not perceive OSINT as the best approach to deter crime or monitor cyber threats. Therefore, we developed a construct consisting of items that measure both aspects of threat perception, namely, acceptance and perceived need.

Like Furnham and Swami [18], many scholars agree that the perception of threats is associated with the acceptance of the use of surveillance technologies [18, 23, 52]. People with a higher threat perception are more likely to support government surveillance technologies. Therefore, people who perceive the need for surveillance technologies are also more inclined to accept them [12]. However, this does not imply that people are willing to share any kind of information in any situation. The context of a threat (like terrorism or a natural disaster) is associated with the information people are willing to provide to an organization [1]: "The more intimate the type of information, the lower the approval of the subjects. Telephone numbers, addresses and location information belong to the data that is not considered critically intimate and would be communicated by a large portion of subjects." Therefore, as the public data and the use cases of OSINT in emergency response and monitoring increases, studies are needed that focus on the context factors for acceptance of OSINT regarding specific data and information, as well as the machine learning-driven algorithms for analysis.

2.2 Privacy Impact of OSINT

Privacy considerations are relevant to the development of OSINT systems, as they gather publicly available data, mostly from social media platforms. Privacy is a well-researched term, which has been explored by psychologists, sociologists as well as computer science, information systems and management research [7, 12, 13, 29, 50]. Privacy has been defined by Westin [56] as the right to control, edit, manage, and delete information about one-self and to decide when, how, and to what extent information is communicated to others.

Due to the rise of digital communication, privacy research has gained more interest and has been operationalized as privacy concerns, meaning "the anticipation of future possible loss of privacy" [13]. As research on privacy concerns has been conducted focusing on a variety of contexts, like online shopping, online social networks and IoT, our approach focuses on the context of online social media, in which people deliberately share personal and other information. Social theory and behavioral research have studied reasons why people take part in social media, such as benefits of participation, profiling and social connection [9]. Debatin et al. [9] found the reasons for self-disclosure to be "(a) the need for diversion and entertainment, (b) the need for social relationships, and (c) the need for identity construction".

Privacy concerns and their explanations have been studied mostly regarding commercial contexts [6, 50][1], and only a small number of studies have addressed governmental surveillance and monitoring [12, 23, 28, 52]. In their study on privacy concerns and their effects on the acceptance of surveillance in Australia, Kininmonth et al. [28] tested several factors associated with the acceptance of surveillance technologies. In particular, they examined the privacy concerns and practices, the concerns regarding secondary use of data, the perceived need for surveillance, the trust in the government, as well as the trust in data management and protection. They found that privacy concerns have a significant influence on the acceptance of surveillance technologies.

The relation between privacy concerns and the need for surveillance was studied by Dinev et al. [12]. They found that people who have privacy concerns would not perceive surveillance as necessary, and are less likely to disclose personal information. However, they also noted that "surveillance technology is being adopted and used faster than public awareness of it and is outpacing the public debate" while people are willing to give information to fight terrorism. Further, they added that this is also a result of "the nature of the search for a balance between security and privacy within the context of the continuous flow of information technology advancements and their implementation in private and public institutions." The pace of technological and political change on surveillance measures makes longitudinal studies necessary.

Such a longitudinal study has been conducted by Wester and Giesecke [55]. They investigated the attitudes towards privacy and surveillance and their change over time. In this context, they found that the risk perception of surveillance has decreased while the call for transparency has increased “dramatically” between 2009 and 2017, concluding that “this suggests that citizens not only make distinctions between different technologies, but also what actor is collecting and analyzing the data. Discussions about trust, transparency and accountability should then be held in relation to the different owners – and perhaps the relation between them.” This again strengthens the need for context-focused research, which also takes data-gathering institutions into account.

As privacy concerns are the anticipation of possible future privacy violations and/or the loss thereof, the risk perception of certain technologies helps to better understand user behavior. In their study, Gerber et al. [19] conducted an online survey with 942 participants on the risk perception of social networks, smart home and smart health devices. They found that participants perceived abstract risks to be more likely but moderately severe, while specific risks were perceived as moderately likely and more severe. Additionally, people did not seem to be aware of specific privacy risks in abstract scenarios, illustrated by standard disclaimers like "your data are collected and analyzed". As a result, the authors call for measures that raise people’s awareness about what is collected and analyzed and how information can be used or even abused.

Actual privacy behaviors in contrast to privacy concerns have been the subject matter of many studies, leading to the discourse on the privacy paradox [2, 29], which assumes a disconnect between desired privacy and potentially contradicting behavior. However,

[1] For a systematic overview on privacy behaviors and concerns see the meta study by Kokolakis [29].

other studies have questioned the privacy paradox and its resulting claims. This means, that observed behavior does not necessarily contradict privacy claims. For example, the privacy calculus suggest that all behaviors protecting one's privacy follow a rational choice in which giving up privacy can be rewarded [11]. Another branch of research has focused on privacy socialisation and the effects of groups and activism, especially as social media platforms have become relevant to political activists [33]. Therefore, people actively avoid surveillance as an expression of shifting privacy norms and question the legitimacy of government surveillance or individual measures [25].

In their study on the development of measures for privacy concerns and behavior during online shopping, Buchanan et al. [6] have found two separate factors that build the foundation to behavior aimed at protecting privacy. For the two factors, the general caution and common sense needs to be distinct from the "sophisticated use of hardware and software", which requires a more specialized knowledge and technical training for the actual protection.

Looking at the OSINT systems which were developed for cybersecurity purposes in a systematic study, Riebe et al. [46] identified 73 systems, from which 11 discussed ELSI implications. Especially when systems aim at focusing on particular actors, this could include the profiling of individuals [14]. Thus, privacy and legal implications must be assessed for systems which aim at profiling as well as detecting insider threats. However, systems which focus less on individuals and more on cyber threat intelligence (CTI) to detect and analyze threat early, can also impact privacy. First, the use of online social networks [49] and second the processing and analysis of data are relevant for privacy implications [42]. In this context, the trade-off between data protection requirements and the demand for forensic investigators are discussed by Nisioti et al. [35].

2.3 Research Gap and Hypotheses Development

The research on surveillance has focused on different application areas, such as camera surveillance in public spaces [54], and online surveillance [28], as well as causes of legitimization, such as fighting crime, terrorism but also public health monitoring [23, 24]. Thus, surveillance has been studied regarding its factors associated with acceptance as well as concerning different scenarios. The importance of the scenario, the surveillance actors and their use of information has been identified [1, 28, 52, 55]. Thus, for the use of OSINT in cybersecurity, these factors and scenarios need to be researched.

While OSINT systems can also be useful in the early detection and monitoring of cyber threats and incident communication, they, like other SOSTs, can create uncertainty [32]. Therefore, to understand the attitudes towards OSINT in the case of cybersecurity, we conducted a representative survey among the German population asking about the aforementioned constructs and how they relate to the acceptance of OSINT (see Table 1).

In the following, the research question is further operationalized in hypotheses. Furnham and Swami [18], Ioannou and Tyssyaduah [23], and Trüdinger and Steckermeier [52] have shown, that the level of threat perception is associated with the level of acceptance of surveillance measures. Threat perception is defined as the participant's fear of crime, terrorism, and of being harmed (see Table 1). As OSINT is a group of surveillance technologies, we derive the first hypothesis based on their research:

H1: People with a higher cyber threat perception are inclined to be more accepting of OSINT.

However, the perception of a threat might not necessarily mean that people would perceive the use of surveillance technology as the preferred approach to detect and to deter criminals and terrorist, or as the preferred measure of ones protection against these threats [12, 28]. Therefore, we asked participants for their perceived need of OSINT separately. This concept has been studied by Kinimonth et al. [28] and Dinev et al. [12]. Following their work, we defined the perceived need for OSINT as the perception that government surveillance is necessary to protect citizens (see Table 1). The concept is also used by Ioannou and Tyssyaduah [23] in the case of surveillance during the COVID-19 pandemic. Thus, the second hypothesis assumes a positive association between both concepts:

H2: People who think there is an overall need for OSINT are inclined to be more accepting of OSINT.

Privacy concerns, defined as the anticipation of a future possible loss of privacy (see Table 1, [13]), have also been negatively associated with the acceptance of surveillance in related studies, such as by Kinimonth et al. [28] and Dinev et al. [12]. They have shown, that people with higher privacy concerns are less likely to accept surveillance. Thus, regarding the use of OSINT for cybersecurity, we expect a negative correlation between privacy concerns and the acceptance of OSINT:

H3: People with greater privacy concerns are inclined to be less accepting of OSINT.

Privacy research has intensively studied how privacy concerns are related to privacy behavior [2, 11], and has found their association to be complex, with potentially divergent behavior [29]. Therefore, we separately investigate privacy behavior as the protective behavior to protect one's privacy (see Table 1). Because stronger privacy behavior can, to a certain degree, be viewed as a manifestation of higher privacy concerns, we assume that it is negatively correlated with OSINT acceptance:

H4: People with stronger privacy behavior are inclined to be less accepting of OSINT.

In their study on the acceptance of surveillance policy, Trüdinger and Steckermeier [52] research the effect of awareness of surveillance policies on the trust and acceptance of these policies. They use the concept of awareness, as an individual's knowledge on the existence and use of surveillance (see Table 1) is as an interactive factor. This is especially interesting for OSINT, as the gathering and analysis of public data online are not observable for the individual and the effects are rather abstract for people, which might affect their evaluation of the policies and measures [19]. However, in their study on the effects of the Snowden revelation on the public's opinion, Valentino et al. [53] show that awareness is not associated with the rejection of SOSTs. Thus, following Trüdinger and Steckermeier [52], we formulated a more exploratory hypothesis using awareness of OSINT as an interactive factor:

H5: The level of awareness of OSINT changes the associations between cyber threat perceptions, privacy concerns, privacy behavior, as well as perceived need for OSINT and OSINT acceptance.

Table 1: Constructs used in the survey

Construct	Definition	Source
OSINT Acceptance	Acceptance of a range of surveillance activities	[23, 28, 52]
Threat Perception	Fear of crime, terrorism, and of being harmed	[12, 28, 52]
Perceived Need for OSINT	Perception that government surveillance is necessary for the protection of citizens	[12, 23, 28]
Privacy Concerns	Anticipation of future possible privacy violation and/or the loss thereof	[12, 23, 28]
Privacy Behavior	Protective behaviors enacted to preserve online privacy	[6, 28, 29]
OSINT Awareness	Knowledge on the existence and use of OSINT	[19, 52, 53]

3 RESEARCH DESIGN

In this section, the research design is presented, including the design of the survey as a representative study (3.1). This section further introduces the questions posed in the questionnaire (3.2), as well as the data collected and the criteria for the representative survey (3.3). In section 3.4, the methodology for the data analysis is described and section 3.5 presents the ethical consideration of the survey design.

3.1 Survey Design

The survey was designed within the scope of a three-year research project, which aims to develop novel strategies and technologies for CERTs to analyze and communicate the security situation in cyberspace. To design and refine the questionnaire, the process included a review of published cybersecurity surveys and two workshops with four cybersecurity practitioners from German state CERTs (team leader, incident manager, information security officer and public safety answering point employee) and four interdisciplinary researchers (digital humanities, human-computer interaction, IT security and political sciences). The first workshop comprised these phases. First, we held a presentation (15 minutes) to introduce the overall topic, the procedure for conducting a representative survey, and the aim of this workshop to generate a questionnaire. Examples of closed and open-ended questions were also introduced. Second, we conducted a reflection phase (15 minutes) where, participants were instructed to note their ideas or questions on a digital board. Third, the workshop ended with a presentation phase (30 minutes) during which participants presented their ideas, which we subsequently arranged thematically on the digital board. Based on this input, we created a preliminary version of the questionnaire.

In the second workshop, we presented and discussed the preliminary questionnaire by reviewing all questions individually. Participants discussed and refined existing questions, generated new ones, and reflected upon their thematic grouping or relevance for the research project. Based on the workshops input, we created a second draft of the questionnaire and distributed it via email to the workshop participants for a final round of feedback and revision. The final version of the questionnaire is summarized within the next subsection.

3.2 Questionnaire

In its final version, the questionnaire comprised 20 closed questions. First, we obtained consent for participation (Q1) and then asked about demographic variables of age (Q2), gender (Q3), education (Q4), region (Q5), and monthly income (Q6).

Second, we wanted to gain insights into how citizens assess the current and future threat situation and possible protective measures in cyberspace. Thus, the participants were asked about their usage of internet devices (Q7), their general perception of cyber threats (Q8), how familiar they were with institutions that contribute to cybersecurity in Germany (Q9), how often they had been victims of specific cyberattacks in the past five years (Q10), whom they would ask for help in the event of a cyberattack (Q11), how they estimate the risk of becoming a victim of a cyberattack in the next five years (Q12) and how continuously they use security tools or measures on personal devices (Q13).

Third, we intended to gain insights into what disadvantages and advantages citizens see in the analysis of public data (OSINT) by authorities, government, and companies. Accordingly, participants were asked to evaluate statements regarding the prevalence, use, and impact of OSINT (Q14), as well as OSINT activities by security agencies (Q15). Finally, in a second part of the survey we posed questions concerning citizens' communication and information needs and behaviors (Q18-20), which were not analyzed in this study and had no halo-effect on the previous questions. The questionnaire and the items can be found in Table 5 in section A.

Most questions were designed as five-point verbal rating scales (VRS), with the exception of Q1 (binary consent), Q2 to Q6 (demographic variables), Q7 (four-point VRS), Q10 and Q12 (six-point VRS), and Q18 (multiple choice with up to three items). However, due to the broader scope of the research project, not all questions were incorporated within the analysis of this specific study. The use of a neutral midpoint option on a five-point rating scale is a debated issue. On the one hand, a neutral midpoint enables the accurate response for those with a truly neutral opinion, while the omission could lead to a potentially arbitrary, forced choice. On the other hand, the neutral option may be interpreted differently by individuals and potentially misused as a simple and quick response option (see [8, 34] for discussions on the use of a midpoint option). We included a midpoint mainly to provide an option for those with truly neutral opinions and thus reduce arbitrary choices.

3.3 Data Collection

We transmitted the questionnaire to GapFish who programmed and hosted the online survey. After final quality checks and mutual agreement, they invited participants from their panel to conduct the survey in September 2021. The sample of N=1,093 participants was selected to represent the German population in terms of age, gender, education, income, and state (represented by ISO 3166-2 codes).

- **Age**: 18-24 (8.9%), 25-34 (14.6%), 35-44 (15.0%), 45-54 (16.7%), 55-64 (18.2%), 65+ (26.5%)
- **Gender**: Female (50.2%), male (49.6%), diverse (0.1%), not stated (0.1%)
- **Education**: Lower secondary education (28.5%), middle or high school (55.3%), academic degree (16.3%)
- **Income**: <1,500€ (24.5%), 1,500€-2,600€ (30.8%), 2,600€-4,500€ (28.9%), >4,500€ (15.7%)
- **State**: DE-BB (2.6%), DE-BE (4.5%), DE-BW (13.4%), DE-BY (15.9%), DE-HB (0.8%), DE-HE (7.6%), DE-HH (2.3%), DE-MV (1.6%), DE-NI (9.7%), DE-NW (21.7%), DE-RP (4.9%), DE-SH (3.6%), DE-SL (1.2%), DE-SN (4.9%), DE-ST (2.7%), DE-TH (2.6%)

These criteria ensure that we can infer the German usage patterns with minimal biases, avoiding selection biases inherent in surveys, as a predominant bias includes favoring specific groups based on specific criteria, e.g., based on occupation and/or availability.

3.4 Statistical Analysis

The analysis was conducted using the software tools Microsoft Excel and RStudio Version 4.0.5. Answers with the rating of "no response" were excluded as missing values from the subsequent analysis. The sample was reduced by two participants because they did not answer quality check questions correctly, such as requests to mark a specific answering box. Initially, a descriptive analysis with response distributions for separate items related to the acceptance of OSINT was conducted. For the statistical analysis, items were combined with regard to their corresponding superordinate construct. Since summed values of the Likert-Scores were used in this course, the corresponding scales were treated as interval-scaled for the subsequent statistical analysis. The reliability of the corresponding scales was established based on the internal consistency with Cronbach's Alpha and also Omega (as a measure of congeneric reliability [43]).

To analyze the hypothesized associations between cyber threat perception, OSINT need, privacy concerns, privacy behavior, and OSINT acceptance, a multiple linear regression was applied. In this course, the former were used as predictors while the latter (OSINT acceptance) represented the dependent variable in an ordinary least squares (OLS) regression model. Several assumptions for running multiple linear regression (linearity of associations, multicollinearity, normality of residuals, homogeneity of residuals) were checked and did not reveal any severe issues. The regression was conducted twice. Whereas the first model represented independent effects, the second, more exploratory model, represented interaction effects with the factor OSINT awareness.

3.5 Ethics

The study was conducted in accordance with the requirements of the local ethics committee at our university. These requirements include, but are not limited to, avoiding unnecessary stress, excluding risk and harm, and anonymizing participants. In the study, the demographic variables of age (Q2), gender (G3), education (Q4), region (Q5), and monthly income (Q6) were collected. Particularly sensitive data (e.g. ethnicity, religion, health data) was not collected. Participants were not misled, but were transparently informed about the study's procedure and goals, and subsequently gave their informed consent to participate. GapFish (Berlin), as the selected panel provider, is ISO-certified and ensures panel quality, data quality, security, and survey quality through various (segmentation) measurements for each survey within their panel of 500,000 active participants.

4 RESULTS

In this section, the descriptive results focusing on the awareness and acceptance of OSINT for cybersecurity are be presented in the first subsection (4.1). In the second subsection 4.2, the factors associated with the acceptance of OSINT are analyzed.

4.1 Descriptive Results

To understand the general and case-specific **OSINT acceptance**, we asked participants about their agreement regarding seven abstract and specific use cases for OSINT in cybersecurity. Scenarios asked for the acceptance of the use cases and the use of artificial intelligence algorithms for the analysis of data. Here, the analysis of publicly available data by the police to pursue criminal activities has the highest acceptance. However, the acceptance is lower when participants were asked if they would agree that artificial intelligence should be used to analyze publicly available data. Notably, they expressed greater dissent to having information shared without their knowledge, i.e. without their consent. However, the neutral positions are among the largest groups among the participants (see Figure 1). This could mean, that many participants had not yet formed an opinion due to the lack of public discourse on this topic. Overall, the combination of the six items shows a high internal consistency for the construct with an alpha of 0.85 and omega of 0.89. The values for all constructs can be found in the appendix (see Table 3).

To understand participants' cyber **threat perception**, we asked them to assess the likelihood of them, or society, or the information infrastructure becoming a victim to a range twelve cyber threats within the next five years (see Figure 8, in the Appendix). Regarding infrastructure threats, malicious software was considered to be the most likely threat in the next five years, while the threat by distributed denial of service-attacks and advanced persistent threat were perceived to be less likely, both with high numbers of neutral participants. Regarding the individual threat perception, participants perceived the following risks to be most likely: Spam messages, spyware phishing and unauthorized access to personal social media channels. Identity theft and social engineering to obtain personal information were perceived as the least likely risks . All 17 items show high internal consistency (alpha = 0.97, omega = 0.98).

We used three items to ask participants to evaluate the **need for OSINT** for cybersecurity. There was a higher perceived need and level of acceptance to use OSINT to prevent terrorism and crimes, than for authorities to have more OSINT powers in general see Figure 2). Participants responded similarly, when asked if the use of OSINT would support preventing crime. However, it is also interesting that all items show high rates of neutral positions. The

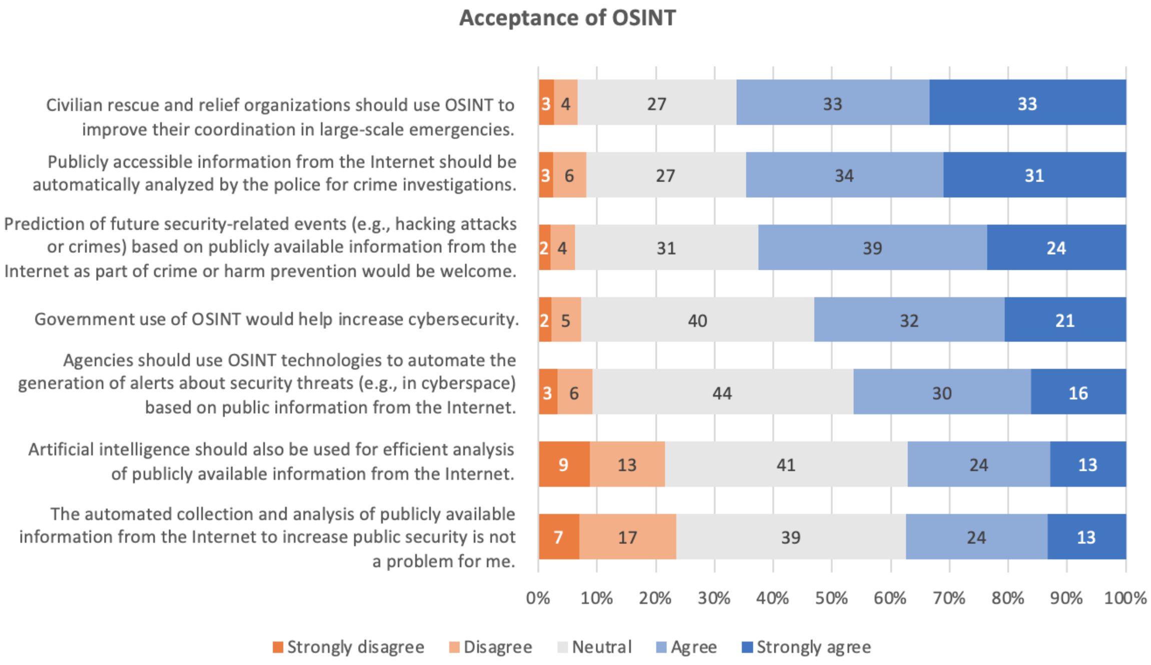

Figure 1: Relative results in all items regarding the acceptance of OSINT.

consistency of the items as part of the construct is high with an alpha of 0.84 and omega of 0.85.

Furthermore, we asked participants about their **privacy concerns** regarding the use of OSINT for cybersecurity. For this purpose, we used four items covering the effects of surveillance on individuals and society, as well as its effects on privacy. Overall, 48% strongly disagreed or disagreed with the feeling of being constantly watched and monitored. Interestingly, many participants (49%) showed neutral positions towards the use of OSINT by government agencies and the majority did not view OSINT as a violation of their privacy online (see Figure 3). All items show a high internal consistency (alpha = 0.83, omega = 0.84).

Concerning **privacy behavior**, participants answered seven items on how often they used certain measures to protect their privacy. With 38-52%, the percentage of people who had never used any of the protective measures was high for all requested items. However, a few differences in privacy behavior can be observed. Among the different behaviors, measures such as covering the camera lens of laptops and smartphones as well as the use of encrypted messengers were much more common than the other measures listed. These are followed by different forms of encryption, e.g. for emails and files, as well as the use of VPN connections which help to encrypt online traffic. The least used methods are the use of anonymization services, e.g. proxy services, and meta search engines that protect user data (see Figure 4) . Again, the items show a high internal consistency (alpha = 0.84, omega = 0.87).

To research the **OSINT awareness**, participants were asked to what extent they were aware of OSINT activities using public and social media data. The three items asked participants for their awareness of publicity, the gathering of such public data as well as for actors conducting the OSINT activities. Seventy-two percent of participants were aware that their shared information can be analyzed by other actors online, but the use of OSINT was lesser-known to participants (51%, see Figure 5). The internal consistency of these items is also rather high with an alpha of 0.77 and omega of 0.80.

4.2 Factors Associated with OSINT Acceptance

4.2.1 Main Effects. To assess the main hypotheses, a multiple linear regression was applied to predict OSINT acceptance based on (H1) cyber threat perception, (H2) need for OSINT, (H3) privacy concerns, and (H4) privacy behavior. The overall regression equation was found to be significant (F(4,1086) = 456, p < .001) with an R-squared of .63. Thus, the regression model contained significant predictors and the overall model explains around 63% of the variance observed in the dependent variable OSINT acceptance (see Table 4 for an overview of the regression results).

Of the hypothesized factors, privacy concerns (β= -.06, p < .001), threat perception (β= .02, p < .001) and OSINT need (β= .63, p < .001) significantly predicted OSINT acceptance, whereas privacy behavior did not (β= .02, p < .145). Hence, the parameter estimates indicated a positive relationship for all predictors except for privacy concerns, which is in line with the hypotheses. Among the latter, an increase in privacy concerns was associated with a decrease in OSINT acceptance.

When comparing the relative size of the parameter estimates, the strongest increase in OSINT acceptance was observed based on perceived OSINT need. For an increase in self-reported OSINT

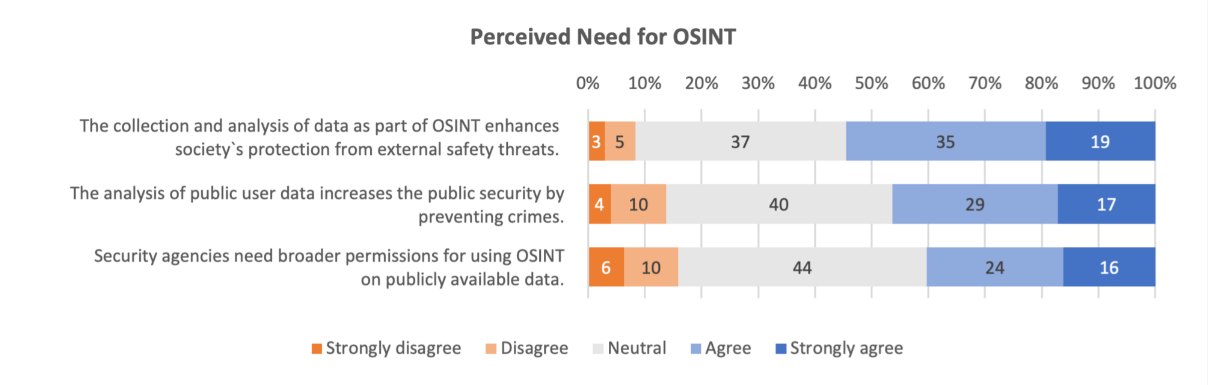

Figure 2: Relative results in all items regarding the perceived need of OSINT.

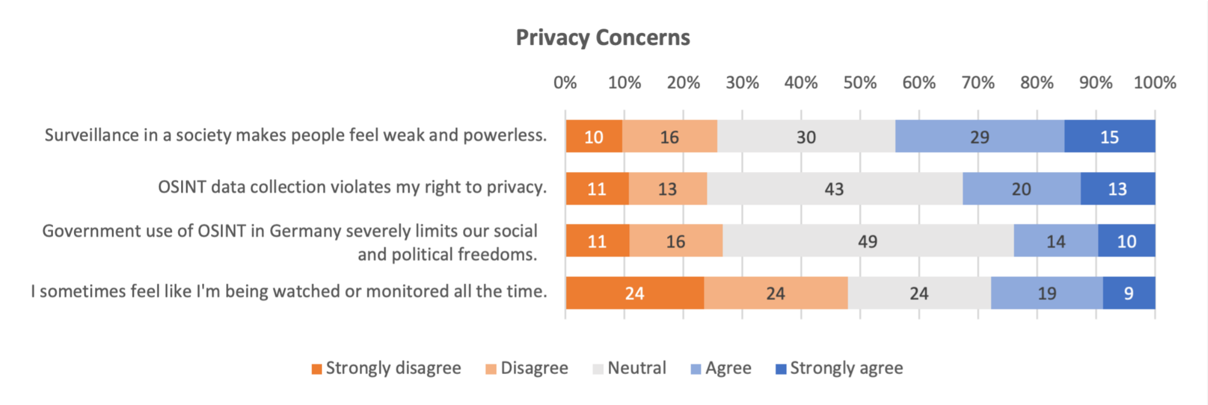

Figure 3: Relative results in all items regarding the privacy concerns.

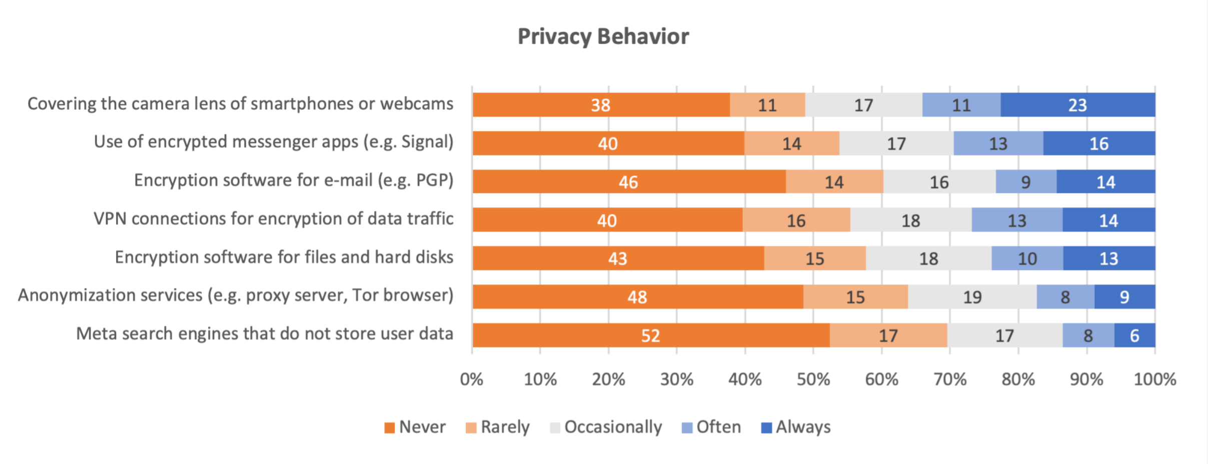

Figure 4: Relative results in all items regarding the privacy behavior.

need, OSINT acceptance increased by .63 points, which is at least 10 times higher than the change in OSINT acceptance based on the other predictors (see Figure 6 for a graphical representation of the relative size of the parameter estimates). This is theoretically plausible, since OSINT need is more closely linked to potential OSINT acceptance than to the other predictors.

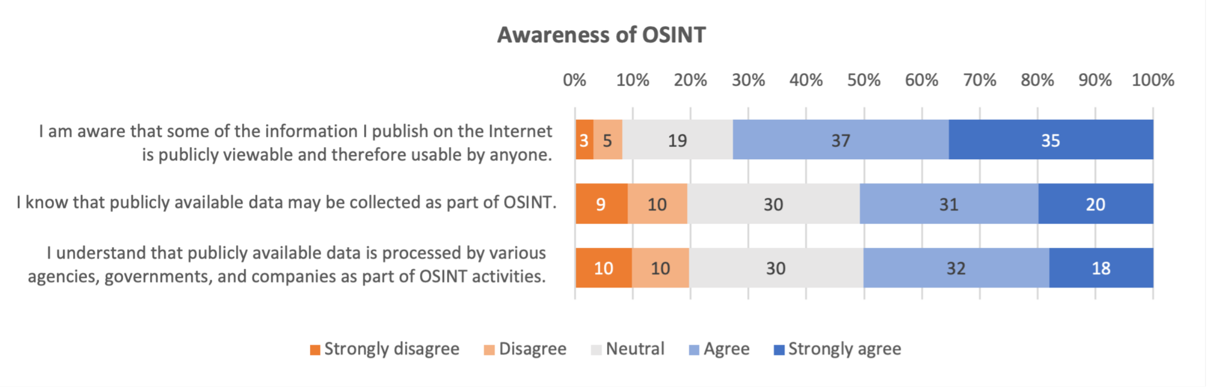

Figure 5: Relative results in all items regarding the awareness of OSINT.

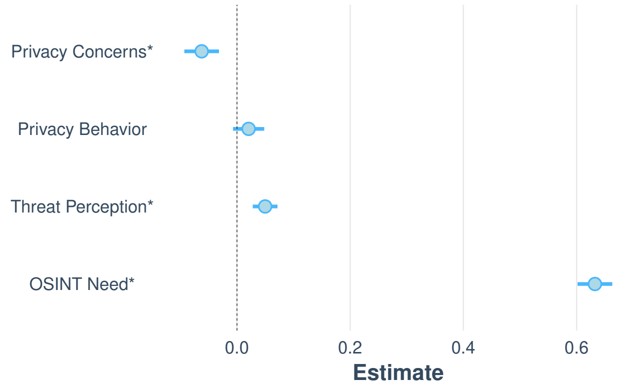

Figure 6: Coefficient estimates from multiple linear regression (Significant predictors for OSINT acceptance are marked with a *). The horizontal lines indicate the 95 % Confidence Interval of the point estimates.

4.2.2 *Effect of OSINT Awareness.*

A second, more exploratory regression model was created to examine the extent (H5) to which the awareness of OSINT technologies might affect the associations between the previously analyzed factors. In this course, the same model was established, except for one addition: OSINT awareness was added as an interaction term for (H1) cyber threat perception, (H2) need for OSINT, (H3) privacy concerns, and (H4) privacy behavior. The objective was to evaluate whether significant interactions exist that might provide additional information on the dynamics of OSINT acceptance.

The resulting overall regression equation was found to be significant (F(9,1081) = 224.4, p < .001) with an R-squared of .65. The complete regression results can be found in Table 2. Through the novel interaction term OSINT awareness, slight changes in the original model became apparent. For example, the previously significant predictor threat perception did not represent a significant predictor anymore (β= .01, p = .89). Moreover, the interaction model did not represent a superior explanatory model compared to the initial model. The 2% increase in explained variance by the interaction model can be considered rather negligible. Furthermore, the interaction model was actually more exploratory from a theoretical point of view. Here, the focus was on the interactions with OSINT awareness, in particular. All but one interaction was found not to significantly predict OSINT acceptance. The interaction of privacy concerns and OSINT awareness was the one factor that significantly predicted OSINT acceptance (β= -.06, p < .001). The pattern of interaction can be seen in Figure 7. While OSINT acceptance differed only slightly at best for high levels of privacy concerns, OSINT acceptance was dependent on OSINT awareness for lower levels of privacy concern. Higher OSINT awareness was associated with higher OSINT acceptance, whereas lower OSINT awareness was associated with lower OSINT acceptance. Thus, minor privacy concerns and low OSINT awareness were associated with low OSINT acceptance, while minor privacy concerns and high OSINT awareness accompanied higher OSINT acceptance.

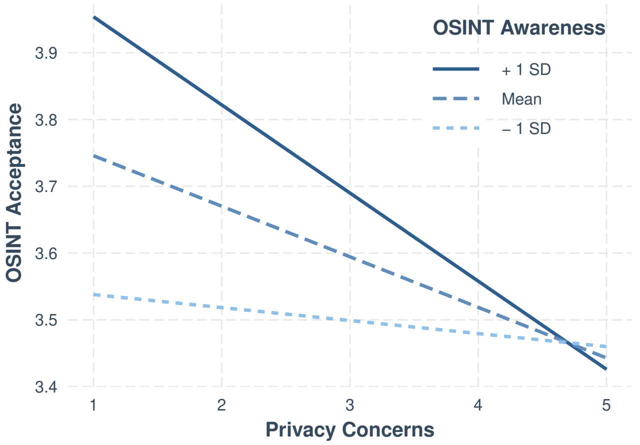

Figure 7: Interaction of Privacy Concerns and OSINT Awareness

5 DISCUSSION

In this section, empirical results are discussed in relation to related work in subsection 5.1. Afterwards, implications for the design and

Table 2: Regression Model with Interaction

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5775	0.2265	2.55	0.0109
Privacy Concerns	0.1406	0.0530	2.65	0.0081
Privacy Behavior	-0.0246	0.0531	-0.46	0.6431
Threat Perception	0.0061	0.0436	0.14	0.8881
OSINT Need	0.6525	0.0499	13.06	0.0000
OSINT Awareness	0.2799	0.0603	4.64	0.0000
Privacy Concerns: OSINT Awareness	-0.0602	0.0133	-4.52	0.0000
Privacy Behavior: OSINT Awareness	0.0103	0.0139	0.74	0.4600
Threat Perception: OSINT Awareness	0.0091	0.0114	0.79	0.4275
OSINT Need: OSINT Awareness	-0.0138	0.0129	-1.07	0.2851

Residual standard error: 0.4324 on 1081 degrees of freedom
Multiple R-squared: 0.6513, Adjusted R-squared: 0.6484
F-statistic: 224.4 on 9 and 1081 DF, p-value: < 2.2e-16

organizational factors for the use of OSINT in the context of cybersecurity are presented in subsection 5.2. Finally, the implications and venture points for future work are discussed in subsection 5.3.

5.1 Factors Associated with the Acceptance of OSINT for Cybersecurity

This study aims to answer the following research question: which factors are associated with the acceptance of OSINT in the context of cybersecurity? Surveillance of public data is increasingly being applied in many areas, from public health policy [4], to the protection against crime and terrorism [18, 28, 52].

The factors positively associated with the acceptance of OSINT are cyber threat perception and the perceived need for OSINT. As the threat perception has been studied, this finding is in accordance with the literature [18, 28, 52]. However, most studies have focused on the unspecific fears of crime and terrorism associated with higher trust in authorities and the state [18]. We examined the specific fear of cyber threats against the individual as well as infrastructure. Similar to other areas of application in cybersecurity, one might assume that the effective use of surveillance to prevent crime can help to gain trust in authorities, should people be aware of this measure [52].

However, people who had already trusted authorities were more inclined to accept surveillance measures [18]. This might be also the reason why some people believe in the need for surveillance as a measure and are thus more inclined to accept surveillance technologies. In our study, the perception that OSINT is needed has been the strongest effect regarding the acceptance of surveillance, which is not surprising.

The correlation between acceptance and privacy concerns is negative, meaning that people with higher concerns tend to be less accepting of surveillance (see Table 4). This hypothesis is supported by related work [12, 25] and can be explained by the fact that OSINT may pose risks to people's privacy. In contrast to other SOSTs, gathering public data on social media platforms is invisible and poses abstract risks. Gerber et al. [19] have shown that people estimate more abstract risks as less severe than specific ones.

Interestingly, our hypothesis on the effect of privacy behavior was not supported by the analysis. This could be explained by the so-called privacy paradox [2]. The privacy paradox has been contested by research on the privacy calculus, which assumes a rational choice in which people weigh the benefits against their privacy concerns [11]. As our model has shown, privacy behavior is not associated with the acceptance of OSINT, in contrast to the effect of privacy concerns. Whether participants suffer from the paradox or follow rational choice might also depend on their awareness of actual surveillance. Kokolakis [29] has discussed the literature on divergent privacy concerns and behavior, and identified five different explanations for divergent concerns and behavior in the literature: (a) privacy calculus theory, (b) social theory, (c) cognitive biases and heuristics in decision-making, (d) decision-making under bounded rationality and information asymmetry conditions, and (e) quantum theory homomorphism. Most of the approaches offer explanations for decision-making which take context factors into account. However, among others, particular trade-offs, social settings, as well as heuristics in decision-making and information asymmetries within the process, have yet to be researched in detail in our field of application.

According to social theory, which explains self-disclosure, Taddicken [51] has shown that privacy concerns hardly impact self-disclosure behavior. Nevertheless, there are factors which seem to moderate the relation. As the majority of users disclose personal information, the author found that there might be different degrees of self-disclosure "with clearly defined communities where users feel safe from privacy invasion". Therefore, as users might have divergent privacy concerns regarding more or less public communities, the author concluded that it would be helpful for users to know their audience.

In the final and exploratory hypothesis, we tested awareness as an interactive factor. Our results show that awareness only changes the dynamic of association between privacy concerns and OSINT acceptance. Here, awareness showed the effect that lower awareness and few privacy concerns were associated with lower rates of acceptance, while higher awareness and fewer privacy concerns were associated with higher acceptance. Thus, the results highlight the importance of transparency and information about OSINT for its acceptance by participants. This has been suggested by Wester and Giesecke [55], who in a longitudinal study showed that the risk perception of privacy loss decreased between 2009 and 2017, while calls for transparency had increased "dramatically". They further suggest "that citizens not only make distinctions between different technologies, but also what actor is collecting and analyzing the data. Discussions about trust, transparency, and accountability should then be held in relation to the different owners – and perhaps the relation between them." Thus, Wester and Giesecke [55] are indicating the role of contextual factors, such as the actors and technologies being used. Research on privacy behavior and its motivations has shown how the role of context influences sharing decisions [29] as well as how context can be conceptualized for the technological design as "privacy in context" [36, 37]. Nissenbaum has identified two norms for contextual integrity to be followed: the norm of appropriateness and the norm of flow or distribution [36].

Hence, a privacy violation would be when informational norms are breached. For this, the parameters are information type, the actors and transmission principles. On the other side, Nissenbaum's concept enables to the secondary use of data as long as the social context of the self-disclosure is respected [37].

5.2 Implications for Design and Organization

As OSINT in cybersecurity draws insights from the fields of crisis informatics [4, 44] as well as surveillance studies [39], the research from similar cases can be used to derive implications for the design of OSINT systems and their evaluation regarding the factors which are associated with the acceptance of such systems. The discourse on privacy and relevant context factors [36, 37] has shown that the following aspects need further consideration: the kind of data which is collected (1), the actors or organization gathering the data (2) and the transmission principles and platforms which allow for the data gathering (3).

The introduction of the General Data Protection Regulation (GDPR) in the European Union has resulted in a shift in data gathering and analysis by increasing users' power over data processing, retention periods, and use [31]. The GDRP has greatly influenced the coverage of privacy topics in data protection, such as safeguarding user data "with the options to access and rectify their information" [31]. In contrast to other organizations, authorities can only collect data for a legitimate and legally approved reason and have to comply with retention periods [47].

OSINT uses a variety of machine learning and deep learning approaches for threat detection and analysis [38]. Thus, machine learning research with greater attention to privacy is a promising area, which helps to include privacy preserving requirements. In particular, machine learning as a service raises privacy concerns, while privacy-preserving computation techniques still demonstrate a lack of "standard tools and programming interfaces, or lack of integration with [deep learning] frameworks commonly used by the data science community" [7].

Regarding the organizations using OSINT beyond the context of cybersecurity, investigative journalistic organisations like Bellingcat and OCCRP follow different interests than, e.g, organisations from crisis management and law enforcement. Emergency services are not allowed to collect personal data without a reasonable suspicion. However, in crisis management, social media data can become a useful tool for situational assessment. In this case, however, the aim is not to collect personal data, but to complete the situational assessment. Research in crisis informatics has shown [1] that the tendency to share more personal data with emergency services changes in crisis situations. Thus, OSINT systems can be used for event detection and analysis, while profiling and analyzing personal data have higher legal barriers. However, this discourse will continue, as questions of accountability and transparency have to be discussed [15].

As many OSINT systems rely on social media platforms, they face the challenge that social media platforms not only provide information about individuals, but sometimes support the sharing of information of third parties. This touches upon issues of interdependent privacy [22, 41], in which a person shares information about another individual.

Therefore, particularly in the context of sensitive cybersecurity information, approaches that help to assess and manage risks from privacy conflicts in collaborative data sharing need to be taken into account for further research [21]. As Riebe et al. [46] have shown, most OSINT systems for cyber security focus on detecting new cyber threats, and might offer additional analysis for incident managers. Concepts of contextual privacy could support this approach, for example, as part of limited data gathering approaches.

Thus, when designing OSINT systems which work on the basis of CTI, the following **implications** should be considered:

- OSINT systems should consider the types of data which is gathered and how it can follow data minimization approaches. In changing threat situations, the gathering strategy could be adopted, and thus could react to changed threat perceptions which are associated with higher OSINT acceptance.
- The system should stay within the social context, which could be achieved by using professional sources (vendors, vulnerability and cybersecurity experts, ...), and could react to larger threats by expanded beyond the context when the need changes.
- Platforms should update their safeguards against disproportionate data collection and support data minimization.
- Individuals' awareness could be raised by using participatory design and maintenance methods.

The results of our study indicate that the awareness of OSINT in combination with lower privacy concerns is a relevant factor associated with the acceptance of OSINT in cybersecurity. This supports findings by Trüdinger and Steckermeier [52]. Therefore, authorities and organizations planning to implement such OSINT systems need to develop strategies to inform affected indirect stakeholders, as well as to include them in the development and implementation process [17, 30]. Research on risk assessment has shown that people assess the severity and likelihood based on specific scenarios [19].

Discourses on social media analysis and emergency management provide venture points for ethical impact assessment. Scholars have argued not to follow a simple logic of "privacy v.s. security", but to consider a wider field of arguments from digital ethics [4, 16]. Participatory approaches, such as ELSI co-design and Value Sensitive Design (VDS), can make use of the identification of factors which are associated with the acceptance of OSINT systems in the context of cybersecurity. Further, such approaches include civil society in the design and implementation of security-oriented technologies [30]. Participatory approaches could also aim to increase the knowledge of non-experts regarding OSINT systems. This would help to raise awareness on these systems.

5.3 Limitations and Future Work

Limitations to our study are presented in the following: As the sample is representative for the German population only, the results are not directly transferable to other countries. Studies have shown that factors like privacy concerns are associated with cultural socialization [13]. Therefore, the effect of the factors we have identified, especially regarding cyber threat perception and privacy concerns, may differ in other cultural contexts. Individualistic cultures, like the United States, might be an interesting and relevant case for further studies on the acceptance of OSINT for cybersecurity. A direct comparison of different cultures would be especially promising in this context. Further, studies have shown, that people have varying understandings of what they perceive as personal, private or even intimate information [51], which was not part of the questionnaire and should be investigated further. Similar, contextual factors [37] and interdependent concepts of privacy [22] need to be studied in greater detail.

Another limitation arrives due to the random sample of the representative study. Many items scored high for "neutral" positions among the participants. While this is to be expected for normally distributed response patterns, this pattern was also evident for asymmetric items. This might have been due to participants' lack of technology-specific expertise, as the development of OSINT systems has not been part of a broader public debate yet, and the use of different technological approaches might not be known to many participants. This also makes specific questions regarding single technological approaches difficult and not feasible in the study design. Thus, further studies need to take expert perspectives into account, as well. We also did not control for affiliation and familiarity with cyber security topics, which could have provided an additional opportunity for analyzing differences between individuals with more and less expertise. However, this was not one of our study's primary goals - an occasional presence of a higher level of expertise, as would be expected in a random sample, was sufficient for our purpose.

With regard to the items, the issue of neutral responses could also have been avoided by not offering a neutral option at all. The usefulness and pitfalls of using such a category have already been discussed in relevant, related literature (see [8, 34] for an overview). One the one hand, we could have potentially gained more insights by not providing such a neutral option. On the other hand – particularly in light of the non-expert sample and potentially actual neutral opinions – this would have introduced more noise into the data by forcing arbitrary choices. Nonetheless, the issue of providing a neutral midpoint should be considered in future studies. Regarding the specific items assessing OSINT needs, it should be noted that some of them were not ideally worded and represented compound questions, which may have increased noise in the data. Lastly, the study uses the data based on self-reported behavior. Thus, the actual privacy behavior might differ from the data in this study.

6 CONCLUSION

The use of OSINT for cybersecurity is a growing research topic, and the number of systems for cyber threat detection and analysis using public online data is increasing [38, 49]. In areas of emergency and crisis management, the surveillance of public data has increased, not only since the COVID-19 pandemic [4, 23]. Research on factors associated with the acceptance of such systems has studied the use of surveillance technology to fight crime and terror, as well as to support people during human-made and natural disasters. Particularly in cybersecurity, new information on cyber threats appears early on online social networks [3]. This is increasingly being used by security operators and government authorities to detect cyber threats early on.

As research on other areas of application has shown, the acceptance of surveillance and of specific measures depends on the context of application, the measures themselves, the information about the measures, and the implementing institutions. The acceptance of OSINT for cybersecurity as a particular case of security-oriented surveillance is an important piece in the puzzle. Thus, a representative study with 1,093 participants in Germany was conducted to understand how people evaluate OSINT in the context of cybersecurity and which factors are associated with the acceptance thereof. The results indicate that:

- Cyber threat perception and the perceived need for OSINT are positively correlated to acceptance, while privacy concerns show a negative correlation.
- The awareness of OSINT, however, has only affected the association between privacy concerns and OSINT acceptance.
- Specifically, high OSINT awareness and minor privacy concerns were associated with higher OSINT acceptance, whereas low OSINT awareness and minor privacy concerns were associated with lower OSINT acceptance.

Implications for further research and the use of OSINT systems for cybersecurity by authorities include conducting research on the implementation of data minimization as a design principle, as well as the association of contextual factors for acceptance, such as the data types and scenarios for situational adaption of gathering strategies. Such approaches should additionally make use of improvements in privacy-preserving computation and machine learning innovations. In terms of OSINT use, we support approaches that provide transparency to people regarding the use of the systems and the data they gather, analyse and retention periods.

ACKNOWLEDGMENTS

This research was supported by the German Federal Ministry for Education and Research (BMBF) in the project CYWARN (13N15407) and by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

REFERENCES

[1] Larissa Aldehoff, Meri Dankenbring, and Christian Reuter. 2019. Renouncing Privacy in Crisis Management? People's View on Social Media Monitoring and Surveillance. In *Proceedings of the Information Systems for Crisis Response and Management (ISCRAM)*. ISCRAM Association, València, Spain, 1184–1197.

[2] Susan B Barnes. 2006. A privacy paradox: Social networking in the United States. *First Monday* 11, 9 (2006).

[3] Vahid Behzadan, Carlos Aguirre, Avishek Bose, and William Hsu. 2018. Corpus and deep learning classifier for collection of cyber threat indicators in twitter stream. In *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 5002–5007.

[4] Kees Boersma, Monika Büscher, and Chiara Fonio. 2022. Crisis management, surveillance, and digital ethics in the COVID-19 era. *Journal of Contingencies and Crisis Management* 13, 1 (2022), 2–9.

[5] Ian Brown and Douwe Korff. 2009. Terrorism and the proportionality of internet surveillance. *European Journal of Criminology* 6, 2 (2009), 119–134.

[6] T. Buchanan, C. Paine, A. Joinson, and U. Reips. 2007. Development of measures of online privacy concern and protection for use on the Internet. *Journal Of The American Society For Information Science And Technology* 58 (2007), 2.

[7] José Cabrero-Holgueras and Sergio Pastrana. 2021. SoK: Privacy-preserving computation techniques for deep learning. *Proceedings on Privacy Enhancing Technologies* 2021, 4 (2021), 139–162.

[8] S. Chyung, K. Roberts, I. Swanson, and A. Hankinson. 2017. Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement* 56, 10 (2017), 15–23. https://doi.org/10.1002/pfi.21727

[9] Bernhard Debatin, Jennette P Lovejoy, Ann-Kathrin Horn, and Brittany N Hughes. 2009. Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of computer-mediated communication* 15, 1 (2009), 83–108. https://doi.org/10.1111/j.1083-6101.2009.01494.x

[10] Sara Degli Esposti and Elvira Santiago-Gomez. 2015. Acceptable surveillance-orientated security technologies: Insights from the SurPRISE Project. *Surveillance & Society* 13, 3/4 (2015), 437–454.

[11] Tobias Dienlin and Miriam J Metzger. 2016. An extended privacy calculus model for SNSs: Analyzing self-disclosure and self-withdrawal in a representative US sample. *Journal of Computer-Mediated Communication* 21, 5 (2016), 368–383.

[12] Tamara Dinev, Paul Hart, and Michael R Mullen. 2008. Internet privacy concerns and beliefs about government surveillance–An empirical investigation. *The Journal of Strategic Information Systems* 17, 3 (2008), 214–233.

[13] Tamara Dinev, Bellotto Masssimo, Paul Hart, Colautti Christian, Russo Vincenzo, and Serra Ilaria. 2005. Internet Users, Privacy Concerns and Attitudes towards Government Surveillance-An Exploratory Study of Cross-Cultural Differences between Italy and the United States. *BLED 2005 Proceedings* (2005), 30.

[14] Matthew Edwards, Robert Larson, Benjamin Green, Awais Rashid, and Alistair Baron. 2017. Panning for gold: Automatically analysing online social engineering attack surfaces. *computers & security* 69 (2017), 18–34.

[15] Quirine Eijkman and Daan Weggemans. 2012. Open source intelligence and privacy dilemmas: Is it time to reassess state accountability. *Sec. & Hum. Rts.* 23, 4 (2012), 285–296.

[16] Luciano Floridi, Corinne Cath, and Mariarosaria Taddeo. 2019. Digital Ethics: Its Nature and Scope. In *The 2018 Yearbook of the Digital Ethics Lab*. Springer, Cham, 9–17.

[17] Batya Friedman, Peter H Kahn, Alan Borning, and Alina Huldtgren. 2013. Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory*. Doorn, Neelke and Schuurbiers, Daan and van de Poel, Ibo and Gorman, Michael E., Dordrecht, 55–95.

[18] Adrian Furnham and Viren Swami. 2019. Attitudes toward surveillance: Personality, belief and value correlates. *Psychology* 10, 5 (2019), 609–623.

[19] Nina Gerber, Benjamin Reinheimer, and Melanie Volkamer. 2019. Investigating people's privacy risk perception. *Proceedings on Privacy Enhancing Technologies* 2019, 3 (2019), 267–288.

[20] Kevin D Haggerty and Richard V Ericson. 2017. The surveillant assemblage. In *Surveillance, Crime and Social Control*, Clive Norris and Clive Wilson (Eds.). Routledge, London, 61–78.

[21] H. Hu, G. Ahn, and J. Jorgensen. 2011. Detecting and Resolving Privacy Conflicts for Collaborative Data Sharing in Online Social Networks. In *Proceedings Of The 27th Annual Computer Security Applications Conference*. New York, NY, USA, 103–112. https://doi.org/10.1145/2076732.2076747

[22] Mathias Humbert, Benjamin Trubert, and Kévin Huguenin. 2019. A survey on interdependent privacy. *ACM Computing Surveys (CSUR)* 52, 6 (2019), 1–40. https://doi.org/10.1145/3360498

[23] Athina Ioannou and Iis Tussyadiah. 2021. Privacy and surveillance attitudes during health crises: Acceptance of surveillance and privacy protection behaviours. *Technology in Society* 67 (2021), 101774. https://doi.org/10.1016/j.techsoc.2021.101774

[24] Georgy Ishmaev, Matthew Dennis, and M Jeroen van den Hoven. 2021. Ethics in the COVID-19 pandemic: myths, false dilemmas, and moral overload. *Ethics and Information Technology* 23, 1 (2021), 19–34.

[25] Elizabeth E Joh. 2013. Privacy protests: Surveillance evasion and fourth amendment suspicion. *Ariz. L. Rev.* 55 (2013), 997.

[26] Sharifah Roziah Binti Mohd Kassim, Shujun Li, and Budi Arief. 2022. How national CSIRTs leverage public data, OSINT and free tools in operational practices: An empirical study. *Cyber Security: A Peer-Reviewed Journal* 5, 3 (2022), 251–276.

[27] Stefan Kaufmann. 2016. Security through technology? Logic, ambivalence and paradoxes of technologised security. *European Journal for Security Research* 1, 1 (2016), 77–95.

[28] Joel Kininmonth, Nik Thompson, Tanya McGill, and Anna Bunn. 2018. Privacy Concerns and Acceptance of Government Surveillance in Australia. In *29th Australasian Conference on Information Systems (ACIS2018)*. Sydney.

[29] S. Kokolakis. 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security* 64 (2017), 122–134.

[30] Michael Liegl, Alexander Boden, Monika Büscher, Rachel Oliphant, and Xaroula Kerasidou. 2016. Designing for ethical innovation: A case study on ELSI co-design in emergency. *International Journal of Human-Computer Studies* 95 (2016), 80–95.

[31] Thomas Linden, Rishabh Khandelwal, Hamza Harkous, and Kassem Fawaz. 2020. The privacy policy landscape after the GDPR. *Proceedings on Privacy Enhancing Technologies* 2020, 1 (2020), 47–64.

[32] David Lyon. 2001. *Surveillance society: Monitoring everyday life.* McGraw-Hill Education (UK).

[33] Linda Monsees. 2020. Cryptoparties: empowerment in internet security? *Internet Policy Review* 9, 4 (2020), 1–19.

[34] J. Nadler, R. Weston, and E. Voyles. 2015. Stuck in the middle: the use and interpretation of mid-points in items on questionnaires. *The Journal Of General Psychology* 142 (2015), 2. https://doi.org/10.1080/00221309.2014.994590

[35] Antonia Nisioti, George Loukas, Aron Laszka, and Emmanouil Panaousis. 2021. Data-Driven Decision Support for Optimizing Cyber Forensic Investigations. *IEEE Transactions on Information Forensics and Security* 16 (2021), 2397–2412. https://doi.org/10.1109/TIFS.2021.3054966

[36] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[37] Helen Nissenbaum. 2018. Respecting context to protect privacy: Why meaning matters. *Science And Engineering Ethics* 24 (2018), 831–852.

[38] Javier Pastor-Galindo, Pantaleone Nespoli, Félix Gómez Mármol, and Gregorio Martínez Pérez. 2020. The not yet exploited goldmine of OSINT: Opportunities, open challenges and future trends. *IEEE Access* 8 (2020), 10282–10304.

[39] Roman Pauli, Hares Sarwary, Peter Imbusch, and Tim Lukas. 2016. "Accepting the Rules of the Game": Institutional Rhetorics in Legitimizing Surveillance. *European Journal for Security Research* 1, 2 (2016), 115–133.

[40] Sina Pournouri and Babak Akhgar. 2015. Improving cyber situational awareness through data mining and predictive analytic techniques. In *International Conference on Global Security, Safety, and Sustainability*. Springer, Cham, 21–34.

[41] Yu Pu and Jens Grossklags. 2016. Towards a model on the factors influencing social app users' valuation of interdependent privacy. *Proceedings On Privacy Enhancing Technologies* 2016, 2 (2016), 61–81.

[42] Priyanka Ranade, Sudip Mittal, Anupam Joshi, and Karuna Joshi. 2018. Using Deep Neural Networks to Translate Multi-lingual Threat Intelligence. *International Conference On Intelligence And Security Informatics* 2018 (2018), 11. https://ieeexplore.ieee.org/document/8587374/

[43] Tenko Raykov. 2001. Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *Brit. J. Math. Statist. Psych.* 2 (2001), 315–323.

[44] Christian Reuter, Amanda Lee Hughes, and Marc-André Kaufhold. 2018. Social media in crisis management: An evaluation and analysis of crisis informatics research. *International Journal of Human–Computer Interaction* 34, 4 (2018), 280–294.

[45] Christian Reuter and Marc-André Kaufhold. 2018. Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. *Journal of Contingencies and Crisis Management* 26, 1 (2018), 41–57.

[46] Thea Riebe, Julian Bäumler, Marc-André Kaufhold, and Christian Reuter. 2022. Values and Value Conflicts in the Context of OSINT Technologies for Cybersecurity Incident Response: A Value Sensitive Design Perspective. *submitted to CSCW* (2022).

[47] Thea Riebe, Jasmin Haunschild, Felix Divo, Matthias Lang, Gerbert Roitburd, Jonas Franken, and Christian Reuter. 2020. Die Veränderung der Vorratsdatenspeicherung in Europa. *Datenschutz und Datensicherheit - DuD* 44, 5 (2020), 316–321. https://doi.org/10.1007/s11623-020-1275-3

[48] Thea Riebe, Marc-André Kaufhold, and Christian Reuter. 2021. The Impact of Organizational Structure and Technology Use on Collaborative Practices in Computer Emergency Response Teams: An Empirical Study. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–30.

[49] Thea Riebe, Tristan Wirth, Markus Bayer, Philipp Kühn, Marc-André Kaufhold, Volker Knauthe, Stefan Guthe, and Christian Reuter. 2021. CySecAlert: An Alert Generation System for Cyber Security Events Using Open Source Intelligence Data. In *International Conference on Information and Communications Security*. Springer, Cham, 429–446.

[50] H Jeff Smith, Sandra J Milberg, and Sandra J Burke. 1996. Information privacy: Measuring individuals' concerns about organizational practices. *MIS quarterly* 20, 2 (1996), 167–196.

[51] Monika Taddicken. 2014. The 'privacy paradox' in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal Of Computer-mediated Communication* 19 (2014), 2.

[52] Eva-Maria Trüdinger and Leonie C Steckermeier. 2017. Trusting and controlling? Political trust, information and acceptance of surveillance policies: The case of Germany. *Government Information Quarterly* 34, 3 (2017), 421–433.

[53] N. Valentino, F. Neuner, J. Kamin, and M. Bailey. 2020. Testing Snowden's Hypothesis Does Mere Awareness Drive Opposition to Government Surveillance? *Public Opinion Quarterly* 84 (2020), 4. https://doi.org/10.1093/poq/nfaa050

[54] Helen Wells and David Wills. 2009. Individualism and identity: Resistance to speed cameras in the UK. *Surveillance & Society* 6, 3 (2009), 259–274.

[55] Misse Wester and Johan Giesecke. 2019. Accepting surveillance–An increased sense of security after terror strikes? *Safety Science* 120 (2019), 383–387.

[56] A. Westin. 1976. *Privacy and Freedom.* Atheneum.

[57] Robin Wilton. 2017. After Snowden–the evolving landscape of privacy and technology. *Journal of Information, Communication and Ethics in Society* 15, 3 (2017), 328–335.

[58] Shikun Zhang, Yuanyuan Feng, Lujo Bauer, Lorrie Faith Cranor, Anupam Das, and Norman Sadeh. 2021. "Did you know this camera tracks your mood?": Understanding privacy expectations and preferences in the age of video analytics. *Proceedings on Privacy Enhancing Technologies* 2021, 2 (2021), 282–304.

A APPENDIX

Table 3 shows the internal reliability of the constructs, while Table 4 presents the regression model. The questionnaire and its items are displayed in detail in Table 5. Figure 8 shows the relative results of all items regarding the participants' threat perception.

Table 3: Reliability of the constructs

Construct	Reliability (Alpha, Omega)
OSINT Acceptance	.85, .89
OSINT Awareness	.76, .85
Threat Perception	.97, .98
Perceived Need for OSINT	.84, .85
Privacy Concerns	.83, .84
Privacy Behavior	.84, .87

Table 4: Regression Model

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	1.3783	0.0759	18.16	0.0000
Privacy Concerns	-0.0624	0.0156	-4.00	0.0001
Privacy Behavior	0.0205	0.0141	1.46	0.1445
Threat Perception	0.0497	0.0111	4.46	0.0000
OSINT Need	0.6323	0.0156	40.46	0.0000

Residual standard error: 0.4464 on 1086 degrees of freedom
Multiple R-squared: 0.6268, Adjusted R-squared: 0.6254
F-statistic: 456 on 4 and 1086 DF, p-value: < 2.2e-16

Table 5: Questionnaire

Construct	Item
Acceptance of OSINT: How would you evaluate the following statements about OSINT activities by security agencies?	1 Strongly disagree; 2 Disagree; 3 Neutral; 4 Agree; 5 Strongly agree
	• The automated collection and analysis of publicly available information from the Internet to increase public safety is not a problem for me. • Artificial intelligence should also be used for efficient analysis of publicly available information from the Internet. • Agencies should use OSINT technologies to automate the generation of alerts about security threats (e.g., in cyberspace) based on public information from the Internet. • Government use of OSINT technologies to automate the collection and analysis of cyber threat and vulnerability information would help increase cybersecurity. • Prediction of future security-related events (e.g., hacking attacks or crimes) based on publicly available information from the Internet as part of crime or harm prevention would be welcome. • Publicly accessible information from the Internet should be automatically analyzed by the police for the purpose of investigating and prosecuting criminal offenses (e.g., to preserve evidence or identify criminal networks). • Civilian rescue and relief organizations (e.g., fire departments, THW, Red Cross) should use automated OSINT technologies to improve coordination of responders and volunteers in large-scale emergencies (e.g., floods).
Awareness of OSINT: How much do you agree with the following statements about the prevalence, use and impact of OSINT?	1 Strongly disagree; 2 Disagree; 3 Neutral; 4 Agree; 5 Strongly Agree
	• I know that publicly available data may be collected as part of OSINT. • I understand that publicly available data is processed by various agencies, governments, and companies as part of OSINT activities. • I am aware that some of the information I publish on the Internet is publicly viewable and therefore usable by anyone.
Percived Need for OSINT: How would you evaluate the following statements about OSINT activities by security agencies?	1 Strongly disagree; 2 Disagree; 3 Neutral; 4 Agree; 5 Strongly agree
	• The collection and analysis of data as part of OSINT helps protect society from threats such as crime, cyberattacks, or terrorism. • The analysis of public user data and posts increases public safety, as security authorities can intervene before a crime is committed. • Security agencies need broader powers to use OSINT technologies for automated collection and analysis of publicly available data on the Internet.
Threat Perception: How high do you estimate the risk of becoming a victim of one of the following types of cyberattacks in the next five years?	1 I cannot judge; 2 Very low; 3 Rather low; 4 Average; 5 Rather high; 6 Very high
a) Threats to infrastructure	• Malicious software such as viruses or worms • No access to online services due to a cyber attack (DDoS attack) • Theft of computing power, for example through cryptomining • Ongoing complex, targeted and effective attack against IT infrastructures (Advanced Persistent Threats)

Construct	Item
b) Threats against oneself / personal data	• Unsolicited, mass delivery of messages (spam)
	• Exclusion, insult, spreading of rumors or sexual harassment on the Internet (cyberbullying) • Threatening or stalking on the Internet (cyberstalking) • A person steals your personal information and pretends to be you (identity theft) • Malicious payment request to regain control over their data or device (ransomware or extortion software). • Loss of money or goods due to online shopping fraud • Malware that coerces me into buying security software (scareware) • Software that spies on me in the background (spyware) • Spying on or stealing confidential data (phishing) • Unwanted publication of private data on the Internet (doxing) • Disclosure of confidential information through manipulation (social engineering) • Unauthorized third-party access to an online or social media account
Privacy Concerns: How much do you agree with the following statements about the prevalence, use, and impact of OSINT?	1 Strongly disagree; 2 Disagree; 3 Neutral; 4 Agree; 5 Strongly agree
	• Government use of OSINT in Germany severely limits our social and political freedoms. • I sometimes feel like I'm being watched or monitored all the time. • OSINT data collection violates my right to privacy. • Surveillance in a society makes people feel weak and powerless.
Privacy Behavior: How continuously do you use the following security programs or security measures on your personal devices (computer, smartphone, etc.) to be protected against cyber threats?	1 I do not know; 2 Never; 3 Once; 4 Rarely; 5 Occasionally; 6 Often
	• Encrypted messenger apps (e.g. Signal) • Encryption software for files and hard disks • Encryption software for e-mail (e.g. PGP) • Anonymization services (e.g. proxy server, Tor browser) • VPN connections for encryption of data traffic • Covering the camera lens of smartphones or webcams • Metasearch engines that do not store user data

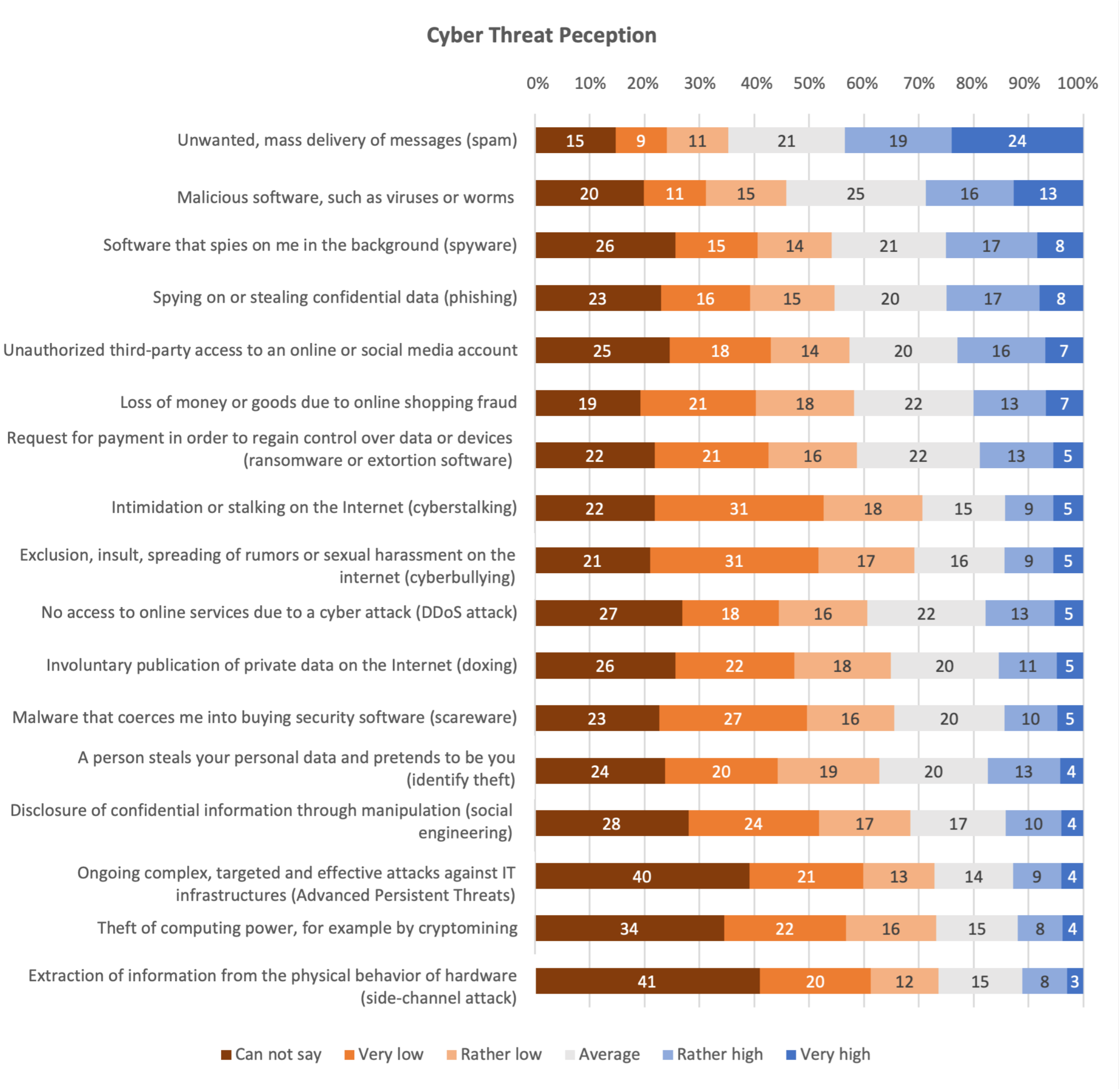

Figure 8: Relative results in all items regrading participants' threat perception.