Attitudes towards Client-Side Scanning for CSAM, Terrorism, Drug Trafficking, Drug Use and Tax Evasion in Germany

Lisa Geierhaas*, Fabian Otto†, Maximilian Häring*, Matthew Smith‡
*University of Bonn, {geierhaa,haering}@cs.uni-bonn.de
†OmniQuest, mailfabianotto@web.de
‡University of Bonn, Fraunhofer FKIE, smith@cs.uni-bonn.de

***Abstract*—In recent years, there have been a rising number of legislative efforts and proposed technical measures to weaken privacy-preserving technology, with the stated goal of countering serious crimes like child abuse. One of these proposed measures is Client-Side Scanning (CSS). CSS has been hotly debated both in the context of Apple stating their intention to deploy it in 2021 as well as EU legislation being proposed in 2022. Both sides of the argument state that they are working in the best interests of the people. To shed some light on this, we conducted a survey with a representative sample of German citizens. We investigated the general acceptance of CSS vs cloud-based scanning for different types of crimes and analyzed how trust in the German government and companies such as Google and Apple influenced our participants' views. We found that, by and large, the majority of participants were willing to accept CSS measures to combat serious crimes such as child abuse or terrorism, but support dropped significantly for other illegal activities. However, the majority of participants who supported CSS were also worried about potential abuse, with only 20% stating that they were not concerned. These results suggest that many of our participants would be willing to have their devices scanned and accept some risks in the hope of aiding law enforcement. In our analysis, we argue that there are good reasons to not see this as a carte blanche for the introduction of CSS but as a call to action for the S&P community. More research is needed into how a population's desire to prevent serious crime online can be achieved while mitigating the risks to privacy and society.**

I. Introduction

Apple's announcement of their plan to implement CSS for known child sexual abuse material (CSAM) in 2021 [1] was met with a predictable outcry from the security and privacy community.

Unlike other such scanning efforts like PhotoDNA [2], which is widely used by companies such as Microsoft, Google and Meta to scan for known material in their cloud offerings, Apple's approach would scan for known CSAM on end-users' devices. While there were a number of concerns raised by the S&P community, to us it seemed the mass surveillance of private devices as opposed to images shared via the Internet was the main point of contention. The outcry was strong enough that Apple withdrew their plan a couple of weeks later.[1] As is common in the long-running debate between privacy and surveillance in the name of public safety, accusations where leveled by both sides. For example, the proponents accuse the opponents of blocking essential progress in the fight against CSAM [3] and claim that the scanning mechanisms will be no more harmful than spam filters [4], [5]. The opponents accuse the proponents of unwarranted mass surveillance [6] - in this case, even of private devices - and point to inevitable feature creep which will go beyond CSAM and lead to misuse by (authoritarian) governments, pointing to examples such as WeChat [7]. A leaked internal memo sent by the National Center for Missing & Exploited Children (NCMEC) to the team at Apple caused particular outrage by calling the opponents *"screeching voices of the minority"*.[2] While Apple have dropped their CSS plans for now, the European Union (EU) Commission is pushing for it through new laws [8], so the debate rages on.

In this paper, we want to pick up some themes from this dispute and gather the German public's view using a representative survey. Consequently, we are focusing on the human aspects as opposed to technical ones. In particular, we are interested in the following:

- Is the move from scanning data shared via the cloud to scanning on people's private devices as salient to the general public as it seems to be to the S&P community?[3]
- Does the type of data (image or text) scanned make a difference?
- What are the levels of support and opposition to client-side or cloud-side scanning for different crimes: CSAM, terrorism, drug trafficking, drug use and tax evasion and are people concerned that the technology might be abused?
- Does it make a difference if law enforcement agencies or private companies are in charge of the scanning and analysis?
- How does trust in government, law enforcement agencies or companies affect the users' views on CSS?

To gather first insights into these questions, we conducted a survey with a representative sample of the German population, balanced for age, gender and federal state (n=1062). To our

[1] https://web.archive.org/web/20211210163051/https://www.apple.com/child-safety/

[2] For context the full memo can be found in Appendix A

[3] This is based on our perception of the community based on publications as well as public and private debates

DOI 10.1109/SP46215.2023.00178

surprise, we found no difference in attitudes concerning CSS or cloud scanning. However, there are large differences when the type of crime is taken into consideration, with CSAM scanning seeing fairly large support. Priming participants about potential benefits or feature creep and abuse of the technology did not have any effect that we could find. Trust in the institutions responsible for the analysis had a small but statistically significant effect. We will discuss these and further findings in detail in the rest of the paper, and we hope to offer useful data for a constructive debate.

II. Related Work

The debate between public safety and individual privacy in digital spaces has long been ongoing [9]–[16]. CSS has been a focal controversial technology in this debate during the past few years and, as such, has garnered a significant amount of academic attention [17]–[20]. Much of this research focuses on technical specifics [19], [20], and on putting the technology in context with societal or political considerations [17], [18], [21]–[23]. Surveys concerning public attitudes to this technology are often driven by interest groups such as child protection or law enforcement organizations and, thus, understandably present their position. In our work, we attempt a neutral assessment of public attitudes in Germany towards CSS - as best as we can. We discuss our positions and potential biases in Section IV-A.

This section presents an overview of currently prevalent opinions in this debate, split into a discussion on political and technical considerations and opinion research.

A. Political And Technical Debate

In 2021, Apple announced their plans to integrate a form of CSS into their operating system to help with the detection of CSAM [1]. In the years prior, Apple had been in repeated conflict with law enforcement, specifically U.S. American agencies, over their encryption policies, which made access to Apple's devices hard or even impossible, even with court orders (compare [24]).

This detection system was supposed to run on the end-users' devices, using an algorithm called Neural Hash to calculate cryptographic hashes of all images saved on the device. These hashes would then be compared to a database of hashes made of known CSAM material.

Even though public backlash led to Apple ultimately postponing the roll-out,[4] the highly publicized debate around this planned feature led to more awareness about CSS. Policymakers see it as a middle ground between preserving individual privacy and ensuring the safety of children, and the European Union recently decided to push for CSS in their legislation around so-called chat control [8].

The IT community reacted with calls for caution and numerous technical concerns in the face of the rising popularity of CSS. In 2022, Struppek et al. published their analysis of Neural Hash [20]. Specifically, they conclude that with Neural Hash, hashes can be manipulated to no longer match when they should and to match erroneously, both with relative ease. Also, the hashes still allow for inferences about the hashed data. They conclude that this form of hashing is not privacy-preserving and, therefore, not ready to be used in CSS.

In 2021, Abelson et al., a group of experts and researchers, published a report in which they collected an overview of several possible risks associated with the widespread implementation of CSS mechanisms [17]. They argue that CSS should be treated similarly to wiretapping from a jurisdictional perspective since it effectively grants government agencies access to private data on all devices, as opposed to just those of suspects or ex-offenders. As such, they argue that if universal wiretapping is illegal, so should be universal CSS.

Levy and Robinson, on the other hand, argue that CSS is an important potential technique to fight online CSAM [21]. They argue that it is possible to implement CSS in such a way that most technical concerns, such as false positives or mission creep, may be mostly mitigated.

In a direct response to Levy and Robinson, Anderson raises several more technical concerns and considerations [18]. Among these arguments are real-life experiences which show that the automated detection of CSAM-related online content, e.g., grooming in text messages, is highly error-prone. Widespread implementation of these kinds of detection mechanisms may lead to an amount of false positive matches that would make any human review by law enforcement or government officials unmanageable.

B. Opinion Research

Public attitudes toward data-scanning mechanisms and the surrounding legislation is an integral aspect of the debate. The case of Apple retracting their plans for Neural Hash for the time being[4] shows that pressure from the general public can significantly impact the real-life implementation of new technology. On the other hand, if there is widespread support for new legislation despite the risks associated with it, mechanisms or systems might be written into law prematurely, opening the door to misuse, whether it be accidental or malicious.

In 2021, ECPAT, a non-governmental organization (NGO) fighting the sexual exploitation of children, surveyed representative samples of citizens from eight European countries [25]. Their results show that a majority of all participants (68%) would support legislation that fights online CSAM, even if it might negatively impact their own privacy. Among other aspects, the survey questioned participants' opinions regarding a sense of privacy online, and whether they deemed their privacy as being more important than CSAM detection. They were also asked to which extent they would support legislation mandating scanning mechanisms to fight CSAM on social media. We have compared our results to those from their specific German sample. German participants of ECPAT's survey generally felt there was little to no privacy online. They overwhelmingly supported legislation for social media content scanning (65%), even if that included their own private

[4] https://web.archive.org/web/20211210163051/https://www.apple.com/child-safety/

messages. When it comes to the comparison between privacy and CSAM detection, 36% of the German participants deemed detection more important. However, an equal amount (35%) expressed that they found detection and privacy to be equally important.

The questions in the survey by ECPAT asked participants to make a broad decision either in favor of the effective detection of child abuse online or in favor of their own, unspecified, privacy. This way of asking for privacy-related attitudes in the form of a two dimensional "would you rather" likely goes back to the privacy calculus [26], which posits that users make their privacy decisions based on a rational risk-benefit assessment. But neither the risks nor the benefits exist in a vacuum, and the attitudes and feelings of people on whether such a data scan is violating their privacy or not likely depend on a complex interaction of multiple factors.

That privacy and the norms connected with it depend on more than just one single factor is covered by Nissenbaum's theory of contextual integrity and the logical framework derived from it [27], [28]. This theory puts the focus on the information flow of personal data transfer and offers *"understanding [of] privacy expectations and the reasons that certain events cause moral indignation"* [28]. It derives peoples' attitudes, expectations and decisions around privacy from several contextual factors, such as which agents are involved (sender, receiver and data subject), the transferred data type, and the restrictions of or reason for the data flow. This idea has already been applied in different contexts, e.g., for understanding the privacy expectations and norms with IoT devices [29] or fitness wearables [30], or sharing data on a social network website [31]. Following the same idea, a contextually-aware permission system for android apps was build by Wijesekera et al. [32], allowing the reduction of privacy violations towards context unaware, ask-on-first-use rules. The design of our survey was inspired by the idea that different contexts evoke varying levels of acceptance, and that this discloses underlying norms and assumptions.

III. Methodology

To build a better understanding on how the German public views the debate on CSS, we designed an online survey experiment and distributed it to a representative sample of German citizens (N=1062). We assumed that it is specifically the details often omitted in such surveys, like which specific risks are connected to a technology or who will be in charge of scanning data, that may sway a person's opinion on whether they deem CSS or other potentially invasive technology acceptable or not. This assumption is in part built on the idea of Contextual Integrity (CI), which posits that several contextual variables in data flow influence what is seen as an unacceptable breech in privacy [27].

We conducted a randomized control trial (RCT) with the following main hypotheses and variables:

- ***H1*** (CSS) Participants are more accepting of cloud-side scanning than client-side scanning.[5]
- ***H2*** (Data) Participants show different levels of acceptance for different types of data (text vs image).
- ***H3*** (Organization) Participants show different levels of acceptance depending on the type of organization doing the analysis (law enforcement vs private company).
- ***H4*** (Crime) Participants show different levels of acceptance for different crimes (CSAM, terrorism, drug trafficking, drug use, tax evasion).
- ***H5*** (Priming) Priming participants with either positive or negative information about CSS will influence their support or opposition (no priming, positive priming: fighting child abuse, negative priming: feature creep).

For these hypotheses, we controlled the condition assignment and we conducted a ordinal logistic regression to analyze the effects.

In addition to these main hypotheses, we have nine exploratory hypotheses based on participant characteristics over which we had no control.

- ***EH1-EH3*** (Trust1): Trust in the government / law enforcement agencies / private companies is positively correlated with a likelihood to agree to CSS measures.
- ***EH4-EH6*** (Trust2): Trust in the government / law enforcement agencies / private companies is negatively correlated with a fear of CSS misuse.
- ***EH7*** (Technological knowledge): Participants who report having IT knowledge are more likely to oppose CSS-measures.
- ***EH8*** (Parenthood): Parents are more likely to agree to CSS-measures when they are intended to fight child abuse.
- ***EH9*** (East vs West Germany): Participants from East Germany are more likely to oppose surveillance measures.

EH9 was formulated based on the speculation that the history of political repression in East Germany until the reunion in 1990 may lead to more wariness in regard to surveillance measures.

In the following sections, we describe the survey design, the statistical tests we used for analysis of the collected data, and details of the survey distribution and participants.

A. Survey

The survey consists of three parts. First, participants were asked some general demographic questions in order to ensure that our sample is representative of the current German population. Then they were presented with a short introduction and six scenarios, in which we asked them to indicate how they felt about CSS given the context of the respective scenario. They were able to choose one of five Likert items, ranging from "I approve" to "I disapprove". After these scenarios, we asked the participants to answer several more questions on their

[5]While we formulated this hypothesis with a direction, we tested using two-tails.

general attitudes and circumstances. The full questionnaire can be found in Appendix B.

To test *H5*, we split participants into three groups. These groups differed in the way they were introduced to the CSS scenarios.

Group CG: The control group. Participants of this group were shown a short neutral base description of CSS: "There is software that allows a mass[6] automated search for files relevant to criminal activities on devices such as smartphones or laptops. If the software finds suspicious material, then the law enforcement authorities are informed."

Group NP: In addition to the above description, participants of this group received a negative primer: "Critics of this software point out that this represents mass surveillance[7] and that what is considered criminally relevant can be too easily expanded and abused. They also point out that errors in the software can cause innocent people to be suspected and then bear negative consequences, such as interrogations and blocked accounts." In this priming text, we attempted to highlight the most salient concerns of the opponents.

Group PP: Participants of this group received the following positive primer: "Proponents of this software point out that serious crimes such as child abuse or terrorism can be solved or even prevented with this technology. This is particularly relevant in the fight against organized crime and crimes where victims are sought on the Internet." In this priming text, we attempted to highlight the most salient arguments of the proponents.

To test *H1-3*, we designed different scenarios that were presented to the participants. Following CI, we varied the device (*H1*: either physical device or cloud), the type of data (*H2*: either text or picture) and the institution doing the analysis (*H3*: either a government institution, or a private company like Google or Apple), resulting in eight scenarios. To keep the time within a reasonable limit a random subset of six of these scenarios was shown to each participant.

The scenarios were phrased like this: "A pre-installed software from the manufacturer automatically searches the [*images* / *text messages*] in/on your [*device (e.g., smartphone, laptop...)* / *cloud storage*]. In case of suspicion, an investigation is carried out by a [*law enforcement agency* / *private company (e.g. Apple, Google ...)*]. Would you approve or disapprove of this measure if the software was implemented to fight the following crimes? - *Child abuse* - *Terrorism* - *Drug trafficking* - *Consumption of drugs* - *Tax evasion*"

We asked the participants on a five-point scale from "I approve" to "I disapprove" to evaluate how much they would support the implementation of this technology. To test *H4* we presented five offenses of varying severity in a randomized order as options against which the technology might feasibly be used: child abuse, terrorism, drug trafficking, consumption of drugs and tax evasion. Child abuse and terrorism were chosen because they have historically been the focus of the political debate concerning CSS [18], [21]. We chose drug trafficking because although we expected it to be seen as less critical than CSAM and terrorism, it is still a serious crime which we expected the vast majority to oppose. We chose drug use and tax evasion because we suspected that these would be seen less critically by parts of the population.[8] We decided not to include actions like political protests or other non-terrorist anti-government behavior. These are not crimes in Germany and, thus, would firmly fall under feature creep, and mixing criminal and non-criminal (or "feature creep criminal") activities would have potentially confused this part of our experiment. This is, however, a necessary aspect to look into in future studies.

After the scenarios, we asked several follow-up questions. If participants had answered that they approved or somewhat approved of the proposed CSS measures in any of the presented scenarios, they were asked whether they generally feared the possible misuse of such a technology. They could answer this question with one of five items, ranging from "Yes" to "No". If participants had indicated disapproval in at least one of the scenarios, we asked them to state their reasons for this disapproval in a free text field. This question was not mandatory but 706 participants answered nevertheless.

Furthermore, we asked several other questions regarding the general attitude of our participants. We asked them to indicate on a five-item scale ranging from "I agree" to "I disagree" whether they agreed that they usually trusted the government to "do the right thing" and government agencies and private companies to "obey the law". We also asked several questions about the participants' background, such as income, family status and technological know-how. These questions were asked to make the comparison to the ECPAT survey possible [25].

B. Statistical Analysis

Since participants were asked to give their opinion on an ordinal scale as opposed to on an interval, we used an ordinal logistic regression for the attribute analysis for *H1-H4*. The regression model included four variables, three of them with two levels each (data type, device and institution have two possible values each) and one variable, crime, which has five levels. Additionally, we included demographic variables to control for them as possible confounding factors and added participant IDs as random effects to account for the repeated measures.

To test *H5*, we compared the average answers participants gave to the scenarios per group with Spearman's rank correlation.

We also used Spearman's rank correlation for several tests according to our exploratory hypotheses (*EH1-EH5*). We used the tests to look for possible influences on the participants' answers outside of the variables which we controlled in the scenarios. Accordingly, we assessed whether the answers

[6]In German we used the word "Flächendeckend" which we believe does not have the same negative emotional connotation as "mass" has in this context.

[7]Here we used the German term "Massenüberwachung" which has the same negative connotation as in English.

[8]There is an active movement to decriminalize drug use

TABLE I
NUMBER OF PARTICIPANTS PER SCENARIO

Scenario	Number
Image / Personal Device / Law enforcement agency	799
Image / Personal Device / Private company	802
Image / Cloud storage / Law enforcement agency	799
Image / Cloud storage / Private company	795
Text message / Personal Device / Law enforcement agency	794
Text message / Personal Device / Private company	795
Text message / Cloud storage / Law enforcement agency	798
Text message / Cloud storage / Private company	790

Each participant saw six out of eight scenarios.

participants gave on trust in the government, in government agencies or in companies, the side they took in the debate, their IT knowledge, parenthood or their state of residency had any correlation with the average answer they gave over all the scenarios they saw. To reduce the risk of false positives, we applied the Bonferroni correction on all the exploratory tests.

C. Power Calculation

To calculate the number of participants we would need to recruit for this survey, we adhered to the recommendations made by Green [33] since we did not have a reliable basis to guess population effect sizes, which would have been necessary for a more formal power analysis. In a regression, variables with more than two levels get recoded into $n-1$ dummy variables with two levels each. Three variables with two levels and one with five levels leave us with seven variables in total to consider. According to the rule-of-thumb introduced by Green, a regression with seven variables would need a total of at least 775 participants to be able to detect small effect sizes. Since we showed the participants only a subset of six scenarios out of eight, we calculated our sample size in such a way that each scenario would be seen by at least the necessary amount of participants according to Green. Table I contains an overview of the numbers of participants per scenario.

D. Distribution

All participants were recruited using the platform OmniQuest.[9] OmniQuest is a company that specializes in market and target audience research and they conduct representative market research via face2face, telephone and online panels. We used the online panel. They are a member of ADM,[10] a German organization for the scientific quality of market research. Participants were paid via OmniQuest, they could choose between an Amazon voucher or a money transfer.

E. Discarded First Study

We collected a representative sample for Germany (n=1014) with an earlier version of this survey, but during data analysis we discovered that we had not properly set up the randomization of the order in which the scenarios were presented to participants. This meant that we could not exclude ordering effects in *H1-H4* and, thus, we needed to repeat the survey. We could, however, analyze *H5* (Priming) and found no effect. Since we found that surprising, we wondered whether our priming had been too subtle and increased the intensity of the priming texts for this study. We also saw a fair number of "straightliners" who either agreed or disagreed with CSS regardless of the changing variables. In other surveys, it is not uncommon to remove straightliners since it is thought that they simply "click-through" without reading the questions. However, in our case, we believe it is completely feasible that participants are absolutely for or against CSS regardless of varying scenarios. To make sure, we added two questions to probe into the way people build and reflect on their attitudes to gain more insights into the straightliners.

F. Main Study

1087 participants were invited to take part in our survey, and 1076 of them completed it. We manually checked the data for plausibility, and removed several participants to improve data quality. 14 participants were removed prior to analysis because they were timing outliers. 11 of them completed the survey faster than 25% of the median time taken by all participants, and we decided to remove them. We removed three participants from the analysis because they took more than 24 hours to complete the survey. This is likely due to taking breaks. Since some questions build on previous ones, too much time between answering them might be detrimental.

As expected due to the results of our first study, we had a lot of "straightliners" who either wholly agreed (N=80) or wholly disagreed (N=114) with the proposed technology, regardless of the circumstances. We did not remove them since, given the context, we think it is plausible that this reflects their true opinion and is not an artifact of speeders. This is backed up by our observation that out of the 114 participants who indicated that they disapproved of CSS measures in any context, 89 gave sensible answers in the added free text question asking for their rationale. This to us suggests that they did indeed take the tasks seriously and answered in accordance with their real opinion.

Our resulting sample includes 1062 participants. They took an average time of 8.9 minutes with a standard deviation of 20.5. The sample is a fairly good representation of the German population based on gender and state of residency when compared to the numbers released by the Federal Statistical Office of Germany [34]. By comparison, old people (over 80) are slightly underrepresented with only 1% (instead of about 7%), and people between the ages of 20 and 79 are slightly overrepresented in all groups, by about 8 percentage points each. The age group of under 20 is underrepresented by necessity, because participants had to be 18 years or older. Table II shows an overview of the sample's demographics. Each of the scenarios had roughly the same demographics. For a detailed breakdown of the distributions by scenario, see Appendix C.

[9]https://www.omniquest.de/

[10]https://www.adm-ev.de/standards-richtlinien/

TABLE II
DEMOGRAPHICS

Gender	Male: 48.7%	Female: 51.1%	Diverse: 0.2%
Age	under 20: 0.8%	20 to 39: 30.7%	40 to 59: 35.3%
	60 to 79: 32.2%	80 to 99: 1.0%	
State of residency	Baden-Württemberg: 12.4%	Hesse: 7.7%	Saxony: 5.2%
	Bavaria: 15.7%	Mecklenburg-Vorpommern: 2.4%	Saxony-Anhalt: 2.8%
	Berlin: 4.3%	Lower Saxony: 9.3%	Schleswig-Holstein: 3.5%
	Brandenburg: 2.9%	North Rhine-Westphalia: 21.0%	Thuringia: 2.7%
	Bremen: 0.7%	Rhineland-Palatinate: 4.9%	
	Hamburg: 2.4%	Saarland: 1.3%	

Overview of the demographics of the participants of our study (n=1062).

G. Ethics

The Institutional Review Board (IRB) of our institution reviewed this project and raised no concerns. We adhere to the General Data Protection Regulation (GDPR) of the EU in handling all data we collect. Prior to the survey, a consent form informed participants about the content of the survey and their rights. Participation was voluntary and they could withdraw at any point during the study. To debrief the participants, we showed them all three priming introductions upon survey completion.

IV. Limitations

There is a number of limitations to keep in mind when engaging with the results of our work. Firstly, the participants we recruited were all from Germany and it is to be expected that cultural differences mean these results cannot be generalized to other countries. We expect this to be especially true when the social structure or form of government differs more, e.g., in countries with a repressive regime.

Also, the distribution of ages in our demographic is not an exact match to their real distribution among the German population, with old people being underrepresented. This is likely a result of the study being conducted online.

Apart from that, it is important to keep in mind the inherent limitations of studies conducted with representative online samples, especially when researching issues that might disproportionately affect marginalized communities (e.g. BiPOC, LGBTQ+, ...). Measures such as CSS, if abused, have a danger of being weaponized for discrimination. A representative sample is by definition ill-suited to give a voice to minorities and more targeted research needs to be done to cover these communities.

Our chosen method of recruiting, an online panel by OmniQuest, suffers from the typical biases of such methods, mainly less reach towards those who have not enough technological know-how to navigate the Internet.

Another possible limitation to keep in mind when it comes to the format of our survey is that we decided against using attention check questions. We did so mainly because we have had mixed experiences with attention checks when distributing surveys via companies that run online panels in the past. We have seen many speeders who have given conflicting answers but who correctly answered all attention checks, but have also seen failed attention checks in otherwise consistent data with lengthy and sensible free-text answers. So, we found ourselves regularly questioning our own attention checks and instead relying on timing, internal consistency and free text answers to prune participants who seem not to be paying attention.

Also, even though we attempted to keep participant fatigue to a minimum by only presenting six instead of eight scenarios, fatigue can still be an issue.

Lastly, while we endeavored to be as neutral as we could be, our personal viewpoints and biases certainly influenced the survey design. So, our results should only be taken as one viewpoint in a very complex debate and we encourage other researchers to tackle this issue from other points of view, especially adding additional factors that potentially explain the decision process. To aid in understanding our biases the following section contains positionality statements of the authors.

A. Positionality Statements

The authors are listed in no particular order.

1) Author: This author is a white German computer scientist with a background in Usable Security and Privacy. They are employed by a German university, i.e., a German state employee. They are firmly against mass government surveillance because of the dangers it poses to any person belonging to a demographic that is at risk of discrimination (e.g., women, PoC/BIPoC or LGBTQIA). However, they acknowledge the fact that without any form of content moderation online, digital spaces often facilitate abuse for other (or even the same) high-risk groups, specifically children. They are hoping for a middle ground in which the ensured safety of one at-risk group will not compromise the safety of the other.

2) Author: This author is a German/British white male computer scientist with a background in Usable Security and Privacy. This author is employed by a German university, i.e., a German state employee. This author's view on the privacy vs surveillance debate is that it currently is not particularly constructive or productive. Tech companies are implementing privacy protections based on their ideology with less regard for law enforcement and politicians than these would like. And, conversely, governments are enacting laws and surveillance programs with little regard for the tech community. Despite this, in this author's experience, the actors involved on both sides genuinely believe that their side is right and are working with good intentions. This author's view is that there are

legitimate reasons for surveillance and it should not rest in the hands of private (tech) companies to decide what capabilities law enforcement has. That decision should lie with the people via the democratic process of law-making (with all its flaws). Conversely, law enforcement and governments need to take heed of the warnings presented by the tech community, both with respect to immediate misuse dangers, as well as potential long-term dangers. Both sides should invest more effort in understanding the other side. They also should include the views of the populations affected by these decisions. This includes special interest groups such as victims of abuse, victims of surveillance, and the general population in many different countries.

3) Author: This author is a German white male computer scientist with a background in Usable Security and Privacy. He is employed by a German university, i.e., a German state employee. The author acknowledges that (some kind of) surveillance is part of what the government has to do and, consequently, be able to do so, to fulfill its duty towards the citizens and society. However, he thinks government-level surveillance of citizens has to be restricted and controlled. If in doubt, the maxim "privacy first" should be used. Countermeasures against mission creep and abuse should be based on technology and not only on trust in good behavior. This author's pre-study assumption was that most people who understand and think about the consequences wouldn't want the currently available technology of mass surveillance (e.g., CSS) to be deployed.

4) Author: This author is a white German employee in market research with a degree in Computer Science. The author thinks that the implementation of state surveillance would have serious implications for freedom of expression in the digital space and that it also poses an enormous security risk for all relevant end devices. There is no question to this author that risk groups, especially children, must be protected, but this intention cannot be stated endlessly without considering the consequences towards everyone. In this author's opinion, restricting everyone's freedoms should only be considered when all other means of protection have been exhausted.

V. Results

In this section, we present the results of the analysis of the data we collected. We start with the general overview, proceed with the testing of the main hypotheses and then present further more exploratory insights.

Overall, the average rate of acceptance for the presented CSS scenarios was 49.7%, the average disapproval rate 35.3%. So, overall, the participants were generally in favor of the proposed measures. Figures 1 and 2 show an overview of the answers participants gave, divided by crime and the institution responsible for the analysis.

79.8% of participants indicated that they approved or somewhat approved of the CSS measures in at least one of the scenarios. They were asked if they worried about misuse of the technology (**Q5.1** as corresponding to Appendix B). 54.6%

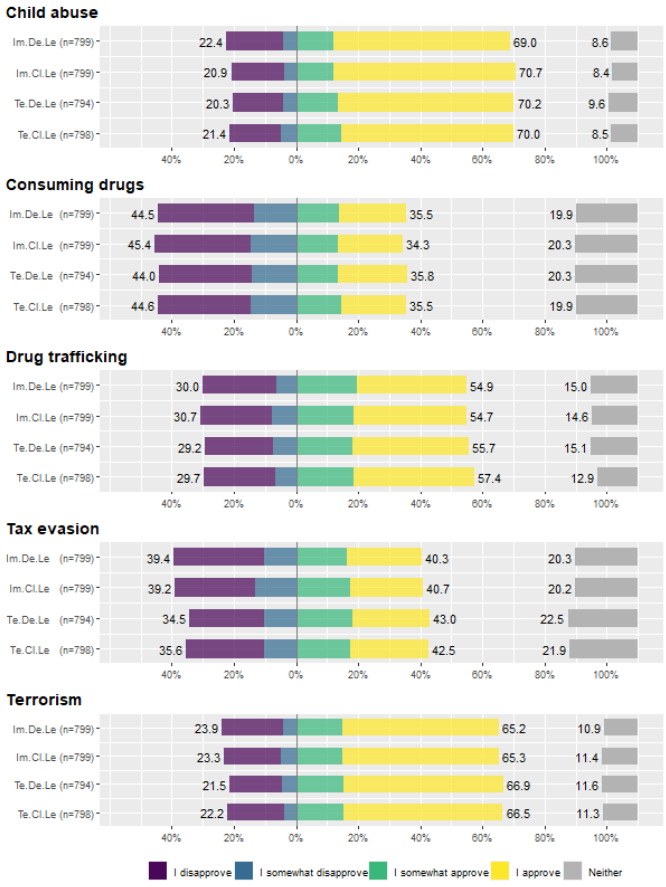

Fig. 1. Answers to scenarios where an investigation is carried out by a law enforcement agency. With Im = Image, Te = Text message, De = Device, Cl = Cloud and Le = Law enforcement agency (see Appendix B for more details)

TABLE III
Answers to Q5.1 (N=848)

Answer	Percent
Yes	24.76%
Somewhat yes	29.83%
Undecided	27.12%
Somewhat no	14.74%
No	3.54%

Q5.1: "You have just indicated that you support one or more of the measures just presented. Are you concerned about possible misuse of the technology?"

of them answered with "Yes" or "Somewhat yes". Table III shows an overview of the numbers for this question.

We will present our statistical analyses in the following subsections.

A. Priming and Scenario Hypotheses

1) H5:Priming participants with positive or negative information about CSS had no effect: To test whether priming had an effect on participants' attitudes, we compared the three groups CG (N=346), PG (N=351) and NG (N=365). Overall, 51% of the responses given by participants to the scenarios in the control group CG approved of CSS measures. In the positive priming group PG, 50% of the answers approved of CSS. In the negative priming group NG, it was 48%.

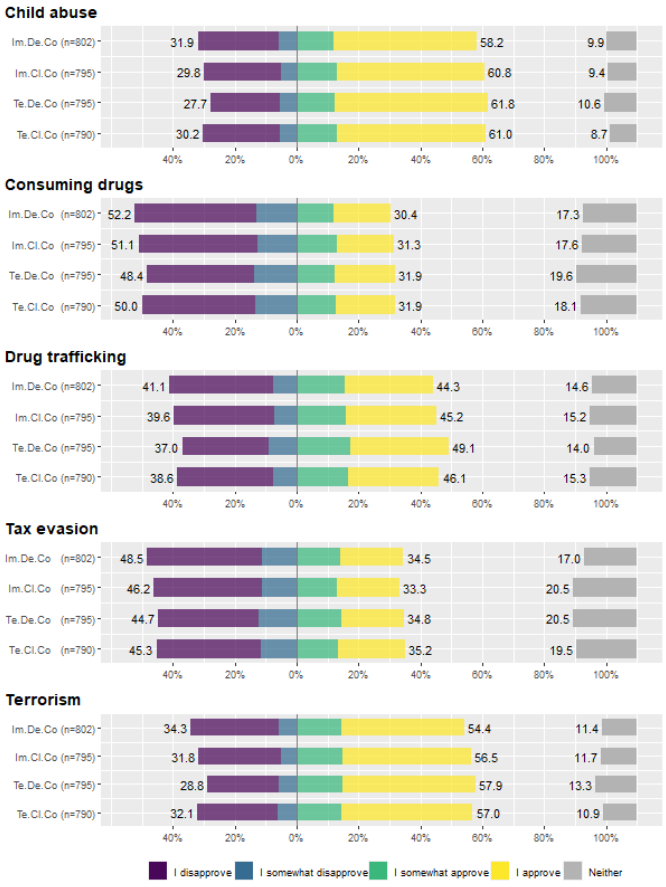

Fig. 2. Answers to scenarios where an investigation is carried out by a law enforcement agency. With Im = Image, Te = Text message, De = Device, Cl = Cloud and Co = Private Company (see Appendix B for more details)

For Spearman's rank correlation, we calculated each participant's average answer to all CSS scenarios with which they were presented. The test showed that the three groups had no statistically significant difference in attitudes ($\rho = -0.037$, $p = 0.23$). So, bringing up positive or negative aspects did not influence our participants' views in any relevant way.

2) H1-H4: Data Location, Data Type, Organization and Crime all have effects on CSS acceptance: To test whether participants' attitudes would change based on contextual factors, we presented the participants with several scenarios on CSS. We varied four variables in these scenarios, scanned data type, device, involved institution and prosecuted crime. To analyze the relationship between these variables and the response that participants gave on a five-point Likert scale, we used an ordinal logistic regression in a cumulative link model with random effects to account for the repeated measures. We added the demographic factors of gender, age, state of residency, income and education as variables to the regression model to be able to control for them. None of these variables proved to be confounding factors, which shows that the randomization of the scenarios balanced these factors well. We included random intercepts to account for the within-subjects design, however, since the focus is not on predicting the outcome for individuals, we omitted reporting on detailed numbers. An overview of the regression results including the demographic variables can be found in Table IV, and the results for the demographics are described in V-B.

Participants were more willing to approve of a scan if the crime under investigation was more severe (*H4*). Participants were much more likely to say they approved or somewhat approved of the CSS measures if they were to be employed to detect cases of child abuse, as compared to drug consumption, which was the least likely to garner acceptance (OR=11.57, 97.5%CI [10.64, 12.59]). Child abuse, terrorism and drug trafficking received on average 65.2%, 61.2% and 50.7% approval, while for tax evasion and drug consumption, this average dropped to 38.0% and 33.3%.

The organization responsible for investigating the data (H3) also made a statistically significant difference in acceptance, with participants having higher odds of accepting measures in the case of an involvement of a law enforcement agency as compared to a private company (OR=1.87, 97.5%CI [1.78, 1.97]).

The type of data that was scanned (*H2*), while still statistically significant, had only a small impact on participants' opinions. Participants were more likely to agree to the scanning measures if the data type in the scenario was text messages as opposed to images (OR=1.14, 97.5%CI [1.09, 1.20]).

Whether data was scanned on a device, like a smartphone, or in the cloud (*H1*), also had a statistically significant influence on the answers participants gave in the scenarios, however, the OR is so small that it is of little relevance in practice (OR=1.06, 97.5%CI [1.01, 1.12]). We find this noteworthy insofar as that this is one of the major differences between CSS and existing cloud-based approaches such as PhotoDNA which have been in use and accepted for many years. The fact that CSS would scan personal devices was also one of the major points of contention for the opponents, but at least with the way we framed our questions, it was no different than a cloud scan. We do want to point out that our survey only presented a high-level view and, as such, lacked nuances, such as the fact that cloud sync can be turned off, which means it is easier to avoid cloud-based scanning compared to CSS. Nonetheless, we think it is important to know that the general population might not weigh the distinction in the same way as the tech-community.

In summary, of our five main hypotheses, we found statistically significant results for four of them.

B. Exploratory analyses

In addition to testing our main hypotheses, we added several demographic variables (age, gender, state of residency, income and education) to our regression model in order to control for them as possible confounding factors. We also conducted further analyses to test our exploratory hypotheses. Table V contains an overview of the exploratory statistical tests and their results. To reduce the risk of false positives, we applied the Bonferroni correction on all the exploratory tests.

TABLE IV
OVERVIEW OVER THE RESULTS OF THE ORDINAL REGRESSION FOR *MH2*

IV (value)	Baseline	OR	2.5 %	97.5 %	*p*-value
Data location (phys. device)	Cloud storage	1.0617	1.0101	1.1159	0.02*
Data type (text message)	Image	1.1415	1.0860	1.1998	0.00*
Institution (l.e. agency)	Private company	1.8690	1.7771	1.9657	0.00*
Crime (drug trafficking)	Consuming drugs	3.1435	2.9140	3.3911	0.00*
Crime (child abuse)	Consuming drugs	11.5690	10.6351	12.5850	0.00*
Crime (tax evasion)	Consuming drugs	1.4195	1.3173	1.5296	0.00*
Crime (terrorism)	Consuming drugs	7.8615	7.2519	8.5223	0.00*
Demographic variables					
Gender (female)	Male	1.6476	1.0500	2.5851	0.03*
Gender (diverse)	Male	0.5494	0.0036	84.5310	0.82
Age (40 to 59)	under 20	2.6224	1.5170	4.5334	0.00*
Age (over 60)	under 20	4.1947	2.3910	7.3593	0.00*
Residency (West Germany)	East Germany	2.4359	1.3221	4.4880	0.00*
Income (€1300 to 2599)	below €1300	1.4387	0.6926	2.9885	0.33
Income (€2600 to 5000)	below €1300	1.2947	0.6284	2.6673	0.48
Income (above €5000)	below €1300	3.2563	1.1706	9.0583	0.02*
Education (tertiary)	Basic education (no degree)	7.1584	1.7814	28.7650	0.01*
Education (secondary)	Basic education (no degree)	9.7537	2.4425	38.9492	0.00*
Education (other)	Basic education (no degree)	3.6275	0.4878	26.9737	0.21

IV: Independent Variable (In brackets: the IV's value); Baseline: the IV's value used as a baseline in the regression; Dependent variable: Acceptance measured as items from 1 to 5; OR: Odds ratio; 2.5% and 97.5%: lower and upper bound of the confidence interval; tests marked with *: statistically significant.

1) EH1-EH6 Trust in government / law enforcement agencies / private companies had an effect on CSS acceptance / fear of misuse: Several of the exploratory analyses we conducted centered around the speculation that trust, either in the government or the private companies implementing such technical measures, may be an integral factor in participants' level of acceptance. Similar findings were made in the field of contact tracing [35], [36], and in other surveillance research [37]. Furthermore, in [38], trust was identified as one of the four "pillars" necessary for consumers to deem the privacy conduct of companies appropriate. So, we calculated rank correlation tests to compare several trust variables against average responses, as well as against the level of fear of misuse reported by participants. Table V offers an overview of these tests and the results as well.

We asked participants to indicate how much they trust that the government is generally doing the right thing, and how much they trust that government agencies are generally doing the right thing. We compared both variables against the average response participants gave in those scenarios where law enforcement agencies would be responsible for handling any scanned data, and found that in both cases, the level of trust correlated positively with the level of approval of CSS measures (government: $r(df) = 0.155$, $p < 0.01$; law enforcement agencies: $r(df) = 0.185$, $p < 0.01$). Figure 4 shows the distribution of answers over the level of trust in the government.

Similarly, when companies were said to be responsible for handling scanned data, participants who indicated they generally trusted private companies to do the right thing were more likely to approve of measures ($r(df) = 0.262$, $p < 0.01$).

Additionally, participants who reported a higher level of trust towards government, law enforcement agencies, or private companies were significantly less likely to report fearing the misuse of CSS technologies (government: $r(df) = 0.128$, $p < 0.01$; law enforcement agencies: $r(df) = 0.143$, $p < 0.01$; private companies: $r(df) = 0.243$, $p < 0.01$).

2) EH7 Technological knowledge had a negligible effect on CSS opposition: It is a common, although not necessarily true [39], stereotype that people with an IT background and more technical knowledge tend to err on the side of more privacy. We asked participants if they had any specific computer skills. While participants who reported to have relevant technical know-how were statistically significantly less likely to support the scanning measures presented in the scenarios, the effect size is very small ($r(df) = -0.088$, $p = 0.04$) and consequently does not play an important role in the overall picture.

3) EH8 Parenthood had no effect on CSS acceptance: Whether participants are parents had no significant influence on the answers participants gave in the scenarios relating to CSAM ($r(df) = -0.040$, $p = 1.00$). Even participants who indicated having no connection at all with any children under the age of 18 were just as likely to approve of CSS measures when they were intended to fight child abuse.

4) EH9 East versus West Germany had no effect on CSS acceptance: Until 1990, Germany was divided, and citizens of East Germany (the DDR / GDR) suffered political repression which was aided by high levels of government surveillance. We speculated as to whether this history might still affect the participants from East Germany, leading them to be warier of measures such as CSS and less trusting of the government. However, being a resident of one of the five states that used to be part of the GDR had no significant influence on the answers participants gave, after the Bonferroni correction ($r(df) = -0.064$, $p = 0.33$, $orig_p = 0.04$). It even had a significant positive correlation with trust in the government, albeit with a small effect size ($r(df) = 0.095$, $p = 0.02$).

5) Regression analysis with demographic variables: To be able to conduct the regression with the added variables age, gender, state of residency, income and education, we grouped German federal states of residency into East and West, binned age and income into three and four roughly equally sized groups respectively, and used gender as is, with the three levels male, female and diverse. Education has four levels, condensing the German education system into categories corresponding to basic, secondary and tertiary education, and "other".

The variables did not affect the main scenario effects. Most income brackets did not have a statistically significant effect on acceptance, with the exception of participants with an income of more than €5000 (OR=3.26, 97.5%CI [1.17, 9.06]). Female participants were more likely to approve of the CSS measures (OR=1.65, 97.5%CI [1.05, 2.59]). Participants who reported having a higher education (secondary or tertiary) were statistically significantly more likely to agree to the measures (OR=9.76, 97.5%CI [2.44, 38.95] and OR=7.16, 97.5%CI [1.78, 28.77] respectively). Being a resident of West Germany led to a statistically significant OR of 2.44 (97.5%CI [1.32, 4.49]) and participants older than 20 were more likely to agree with CSS measures, even more so when they were older than 60 (OR=4.20, 97.5%CI [2.39, 7.36]).

However, to highlight the risk of false positives, we want to mention that our original Spearman's test for the difference between East and West Germany was only statistically significant before we corrected for multiple testing but not anymore after correction. Also, several of these effects have large CIs, further indicating uncertainty. So all results in this exploratory section need to be taken with due caution.

An overview of the results regarding the demographic variables can be found in Table IV.

6) Comparison of CSS acceptance between our survey and the ECPAT survey: To see how the general attitude of our participants compared to the data gathered in the survey published by ECPAT [25], we asked participants about the level of privacy they perceived on the Internet. 61% of the participants we surveyed felt that there is little to no privacy online, 33% indicated that they think there is some or even a lot of privacy. 6% said they were unsure. How people felt about the amount of privacy online had no statistically significant correlation with the attitude they showed towards CSS measures ($r(df) = -0.007$, $p = 0.82$).

The numbers are similar to those reported by ECPAT, where 68% of participants felt that little or no privacy was currently to be found online. ECPAT interpreted this to mean that further privacy-invading measures would not make much of a difference to people who already feel that privacy is rare, stating that the concern for protecting privacy "while valid, may not be shared by the public" [25, p. 6]. Despite our survey producing similar results, our interpretation differs, as we will discuss in Section VI.

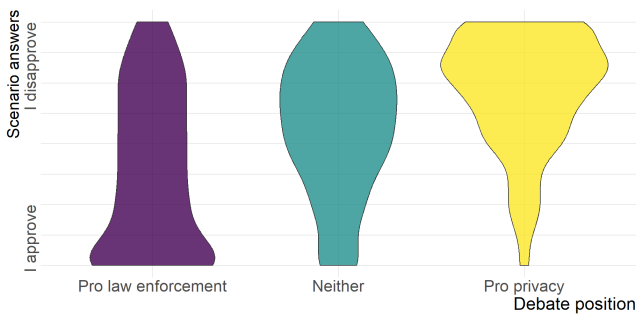

Fig. 3. Distribution of participant answers to **Q-Scenarios**, divided by the side of the privacy debate they see themselves on, as indicated in their answers to **Q-21** (see Appendix B).

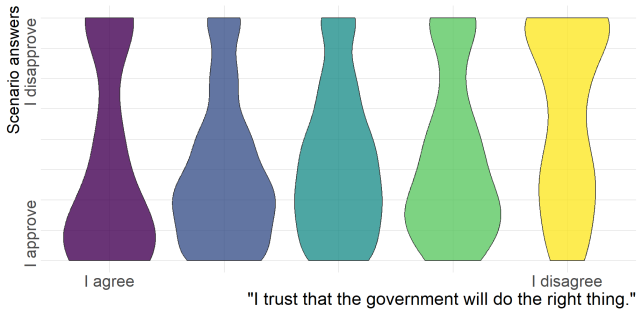

Fig. 4. Distribution of participant answers to **Q-Scenarios**, divided by the level of trust in the government they reported in their answers to **Q-15** (see Appendix B).

VI. Discussion

In this section, we will discuss the most noteworthy results of our survey, and provide some additional thoughts and context.

The data we collected suggests that the participants in our sample, which consisted of 1062 people representative of the German population, are generally supportive of technological solutions to scan for criminally relevant data both in the cloud as well as on devices themselves, even if we highlighted downsides such as feature creep and abuse.

In our sample, 194 participants indicated either full approval (80) or full opposition (114) to the described measures, regardless of our scenarios. So, we see slightly more hard-liners on the opposing side but the majority of participants did vary their support based on the scenarios. We couldn't find a statistically significant effect from the priming texts presented to participants.

While the acceptance for measures was generally high, the crime for which such a detection mechanism may be employed did make a large difference. This highlights how important it is for proponents to take heed of feature creep concerns. While support for CSAM scanning is high, scanning to detect cases of drug consumption or tax evasion was significantly less likely to be met with acceptance. And we did not include the possibility of CSS being used for purposes which

TABLE V
OVERVIEW OVER THE EXPLORATORY RESULTS

IV	DV	test	ρ	p-value	cor - p-value
Side of debate	CSS acceptance	Spearman	0.448	0.0000	0.00*
Trust in government	CSS acceptance	Spearman	0.155	0.0000	0.00*
Trust in law enforcement agencies	CSS acceptance	Spearman	0.185	0.0000	0.00*
Trust in companies	CSS acceptance	Spearman	0.262	0.0000	0.00*
Trust in government	Fear of misuse	Spearman	0.128	0.0002	0.00*
Trust in law enforcement agencies	Fear of misuse	Spearman	0.143	0.0000	0.00*
Trust in companies	Fear of misuse	Spearman	0.243	0.0000	0.00*
Technical knowledge	CSS acceptance	Spearman	-0.088	0.0041	0.04*
Parenthood	CSS acceptance	Spearman	-0.040	0.1909	1.00
East / West	CSS acceptance	Spearman	-0.064	0.0367	0.33
East / West	Trust in government	Spearman	0.095	0.0020	0.02*

Results of the exploratory statistical analyses. DV: Dependent variable; IV: Independent variable;
ρ: Spearman's rho; cor - p-value: p-value, Bonferroni corrected;
tests marked with *: statistically significant.

currently do not fall under German criminal law, such as anti-government protest, which in all likelihood would have even lower acceptance rates. Another indication that feature creep needs to be taken more seriously than just asking for trust is that amongst supporters of scanning mechanisms, 54.6% stated that they were worried about potential abuse of the system.

A. General Acceptance

As mentioned, the described measures for CSS were largely met with approval by our participants. For the detection of CSAM, 65% overall said they were either likely to agree or agreed fully, which is in line with the findings presented by ECPAT in their survey last year, in which equally, 65% stated their support [25].

Unlike the ECPAT survey, we presented participants with several different scenarios and tested variables to gain more context on participants' views. We are able to tell that this acceptance does come with several caveats. First of all, in our sample, the crime being prosecuted made a large difference, and while the majority of our participants stated they were in favor of CSS technology for fighting CSAM, fighting other (legitimate) crimes such as tax evasion would lead them to reject the technology. We think robust mechanisms to prevent the scope being extended are consequently one of the most important areas for research.

Despite our effort to prime participants to think about the risks in the negative priming group, the text about potential misuse was not enough to sway our participants. The same goes for our attempt at increasing support with positive priming. There are several possible explanations for this. Potentially our priming texts were not detailed, salient or extreme enough. It may also in part be due to the fact that, generally, our participants were very likely to trust that the German government was on the whole law abiding and, thus, would not abuse the technology, therefore, rendering the negative priming useless. We base the assumption of the strong effect of the trust in the government on two findings: 1) a high trust in the government was a predictor for likely acceptance, and 2) so was having a government agency responsible for handling the data, as opposed to a private company.

This stands somewhat in contrast to the fact that over half of participants who supported CSS stated that they were worried about abuse and a further 30% stating they were undecided about abuse of such a system. Only about 20% were unconcerned. This leads us to our final potential explanation: participants weighed the risks and benefits as they perceived them and made the conscious choice of accepting the privacy and government abuse risks for the benefit of protecting others. Note that we did not study the assumed size of the benefits or risks and merely use the indirect measure of support to judge the perceived risk / benefit trade-off. Further studies and triangulation will be needed to examine the underlying factors. Getting good quantitative data on the harms and benefits is difficult. In particular since the usefulness of CSS has been called into doubt [18], we think it would be very beneficial to the discussion if law-enforcement agencies could share more data on the scale of the benefit they expect and how this can be measured and verified.

B. Governments vs Companies

Our German participants expressed more support for scanning if it was in the hands of the government as opposed to private companies such as Google or Apple. This is noteworthy in two contexts. Firstly, from a German perspective, US companies currently play an outsized role in this process. Eight out of the ten most used social media platforms in Germany belong to US-American companies, most notably Google and Meta [40]. Microsoft, Google and Apple dominate the market share in operating systems with a combined 97.8% [41], which naturally also extends to a degree to cloud storage solutions. According to the views of our participants, giving these companies too much responsibility may not be the ideal solution. Conversely, we want to highlight the enormous technical, legal but most of all moral challenges of potentially integrating all interested governments into big tech platforms. Our survey only looked at the German public's perspective, and studying the moral dilemmas of different values and norms of different countries was outside of the scope of our study. But our results suggest that it is worth future research and the current US-company and US-norms-heavy situation is not

necessarily the optimal approach for citizens of other countries or their governments.

C. Other Countries and Societal Factors

While we got similar overall results as the ECPAT study which covered eight major EU countries, we want to caution that our results are unlikely to carry over into other nations. Our results show that trust in the German government plays a part in participants' willingness to agree to CSS measures. Consequently, our findings are unlikely to carry over just based on this one variable and there are likely more variables that differ between countries - or even within. Further studies will be needed to shed light on these factors.

D. Comparison to ECPAT's conclusion

Despite getting very similar numbers for the acceptance of CSAM scanning (both 65%) and low levels of perceived privacy on the Internet (61% vs 68%), we come to a different conclusion than that of the ECPAT survey authors.

The authors state: *"Our conclusions: These findings show that in 8 major EU countries, online privacy is seen by a majority of people to have disappeared. Protecting privacy is often used as a counter argument against specific actions to tackle the problem of online child sexual abuse. However, this data suggests that this concern, while valid, may not be shared by the public."* [25]

We do not believe that because perceived privacy is already low, adding further privacy invasive technology is a non-issue. Just because the general privacy situation is bad does not mean that more surveillance cannot make it worse. We also do not believe that the fact that the majority of our participants were in favor of CSAM scanning is a carte blanche for the introduction of this technology.

However, we do think our results show that a substantial part of our sample expressed support for CSS to help prevent or prosecute serious crimes, such as CSAM. These participants were willing to trade some privacy to aid law enforcement concerning these crimes, despite their worries about potential abuse.

Based on the discussions we have followed, S&P activists by-and-large view those trade-offs differently. We do not want to make any claims about the underlying truth or accuracy of knowledge and assumptions on which these trade-offs are based. And we want to explicitly state that the majority opinion is not necessarily the right or moral option. There are most definitively technical, legal, moral and societal issues (especially long term) that our survey did not and could not capture and some issues might be beyond any kind of survey. However, our participants had an underlying willingness to accept this trade-off when the assumed stakes are high enough. We believe this is true for both our survey instrument in Germany and the ECPAT survey instrument in the EU. And in the case of our survey, half of the participants who supported CSS explicitly acknowledge their worry about the risks of abuse, but still chose the support option. Our participants' view that they are supportive of a technology aimed at aiding law enforcement to combat serious crimes, despite the risks, is one we think is important for the S&P community to take seriously.

A possible response is to think that the general population is not capable of properly understanding the implications and that we as gate-keepers need to protect them from themselves. However, we think this age-old debate would benefit from more inclusion of representative views of the public. We believe we need to better understand which risks and benefits the public correctly assessses and accept and differentiate these from risks and benefits that are not seen or misunderstood, since these can lead to unintended consequences. In order to achieve this, it would be particularly useful if proponents could offer better evidence of the effectiveness of existing measures and more transparent estimations of the effectiveness of planned measures. Proponents should also be mindful and take the concerns of potential misuse seriously. Both sides would benefit from working with the public in a scientific manner to gather representative views of this complex topic. We also believe there should be a more concerted effort to research technical solutions that prevent governmental over-reach and abuse while at the same time enabling democratic processes to define the capabilities of law enforcement, instead of this being in the hands of big tech companies which are not accountable to the public in a democratic sense.

VII. Conclusion

We conducted an online survey with a representative sample for Germany (n=1062). We conducted a randomized control trial in which we examined the effects of several variables, including client-side vs cloud-side scanning, types of crime, and positive and negative priming of participants. We also conducted an exploratory analysis of participants' attitudes, such as trust in governments. Our results show that our sample of German citizens had wishes and concerns that are not fully met by either side of the debate between surveillance and privacy. More research is needed, both in understanding people's wishes as well as in technical measures to fulfill them where this is feasible.

Acknowledgment

We thank the Werner Siemens Stiftung for their generous support on this project. The authors would also like to thank Leonhard Karsch from Mentorium for reviewing the statistical analyses.

References

[1] "Csam detection - technical summary," Tech. Rep., 2021. [Online]. Available: https://www.apple.com/child-safety/pdf/CSAM_Detection_Technical_Summary.pdf

[2] (2022) Microsoft. [Online]. Available: https://www.microsoft.com/en-us/photodna

[3] (2021) A welcome step: a statement of support for apple's plans to detect child sexual abuse images. WeProtect. [Online]. Available: https://www.weprotect.org/library/a-welcome-step-a-statement-of-support-for-apples-plans-to-detect-child\\-sexual-abuse-images/

[4] P. Breyer. (2021). [Online]. Available: https://www.patrick-breyer.de/en/chat-control-internal-documents-show-how-divided-the-eu-member-\\states-are/

[5] David Lega [@DavidLega]. @vonderleyen @EU_commission @WeProtect @aplusk @childmanifesto @europarl_en @thorn 2/2 these tools are no more privacy-invasive than tools to detect spam & malware that we all use and i don't recall any of my colleagues complaining about their privacy being violated by those tools. #endCSA. [Online]. Available: https://twitter.com/DavidLega/status/1328275355241746433
[6] T. Claburn. (2021). [Online]. Available: https://www.theregister.com/2021/10/15/clientside_side_scanning/
[7] J. Knockel, C. Parsons, L. Ruan, R. Xiong, J. Crandall, and R. Deibert, "We chat, they watch: How international users unwittingly build up WeChat's chinese censorship apparatus," section: Free Expression Online. [Online]. Available: https://citizenlab.ca/2020/05/we-chat-they-watch/
[8] (2022) Fighting child sexual abuse: Commission proposes new rules to protect children. Commission of the European Union. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2976
[9] H. Abelson, R. Anderson, S. M. Bellovin, J. Benaloh, M. Blaze, W. W. Diffie, J. Gilmore, M. Green, S. Landau, P. G. Neumann, R. L. Rivest, J. I. Schiller, B. Schneier, M. A. Specter, and D. J. Weitzner, "Keys under doormats," *Communications of the ACM*, vol. 58, no. 10, pp. 24–26, Sep. 2015. [Online]. Available: https://dl.acm.org/doi/10.1145/2814825
[10] P. Patel, W. P. Barr, K. K. McAleenan, and P. Dutton. (2019) Open Letter to Facebook. [Online]. Available: https://www.justice.gov/opa/press-release/file/1207081/download
[11] J. Mayer, "Content moderation for end-to-end encrypted messaging," *Princeton University*, 2019.
[12] R. E. Endeley *et al.*, "End-to-end encryption in messaging services and national security—case of whatsapp messenger," *Journal of Information Security*, vol. 9, no. 01, p. 95, 2018.
[13] P. Patel, W. P. Barr, P. Dutton, A. Little, and B. Blair. (2020) International statement: End-to-end encryption and public safety. [Online]. Available: https://www.justice.gov/opa/pr/international-statement-end-end-encryption-and-public-safety
[14] N. A. of Sciences Engineering & Medicine *et al.*, *Decrypting the Encryption Debate: A Framework for Decision Makers*. National Academies Press, 2018.
[15] R. Gorwa, R. Binns, and C. Katzenbach, "Algorithmic content moderation: Technical and political challenges in the automation of platform governance," *Big Data & Society*, vol. 7, no. 1, p. 2053951719897945, 2020.
[16] S. Gürses, A. Kundnani, and J. Van Hoboken, "Crypto and empire: The contradictions of counter-surveillance advocacy," *Media, Culture & Society*, vol. 38, no. 4, pp. 576–590, 2016.
[17] H. Abelson, R. Anderson, S. M. Bellovin, J. Benaloh, M. Blaze, J. Callas, W. Diffie, S. Landau, P. G. Neumann, R. L. Rivest *et al.*, "Bugs in our pockets: The risks of client-side scanning," *arXiv preprint arXiv:2110.07450*, 2021.
[18] R. Anderson, "Chat control or child protection?" 2022.
[19] S. Jain, A.-M. Crețu, and Y.-A. de Montjoye, "Adversarial detection avoidance attacks: Evaluating the robustness of perceptual hashing-based client-side scanning," in *31st USENIX Security Symposium (USENIX Security 22)*, 2022, pp. 2317–2334.
[20] L. Struppek, D. Hintersdorf, D. Neider, and K. Kersting, "Learning to break deep perceptual hashing: The use case neuralhash," in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 2022, pp. 58–69.
[21] I. Levy and C. Robinson, "Thoughts on child safety on commodity platforms," *arXiv preprint arXiv:2207.09506*, 2022.
[22] M. D. M. N. ACHIAGA, "Curbing the surge in online child abuse: The dual role of digital technology in fighting and facilitating its proliferation," 2020.
[23] (2020) Technical solutions to detect child sexual abuse in end-to-end encrypted communications: draft document. Commission of the European Union. [Online]. Available: https://www.politico.eu/wp-content/uploads/2020/09/SKM_C45820090717470-1_new.pdf
[24] S. Savage, "Lawful Device Access without Mass Surveillance Risk: A Technical Design Discussion," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. Toronto Canada: ACM, Oct. 2018, pp. 1761–1774. [Online]. Available: https://dl.acm.org/doi/10.1145/3243734.3243758
[25] "What do eu citizens think of the balance between online privacy and child protection?" Tech. Rep., 2021. [Online]. Available: https://www.ecpat.at/fileadmin/download/Externe_Studien/Summary_Report_Polling_Research.pdf
[26] H. Krasnova, S. Spiekermann, K. Koroleva, and T. Hildebrand, "Online social networks: Why we disclose," *Journal of information technology*, vol. 25, no. 2, pp. 109–125, 2010.
[27] H. Nissenbaum, "Privacy as contextual integrity," *Wash. L. Rev.*, vol. 79, p. 119, 2004.
[28] A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum, "Privacy and contextual integrity: Framework and applications," in *2006 IEEE symposium on security and privacy (S&P'06)*. IEEE, 2006, pp. 15–pp.
[29] N. Apthorpe, Y. Shvartzshnaider, A. Mathur, D. Reisman, and N. Feamster, "Discovering smart home internet of things privacy norms using contextual integrity," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 2, jul 2018. [Online]. Available: https://doi.org/10.1145/3214262
[30] A. Alqhatani and H. R. Lipford, "{"There} is nothing that i need to keep {secret"}: Sharing practices and concerns of wearable fitness data," in *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 2019, pp. 421–434.
[31] P. Shi, H. Xu, and Y. Chen, "Using contextual integrity to examine interpersonal information boundary on social network sites," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 35–38. [Online]. Available: https://doi.org/10.1145/2470654.2470660
[32] P. Wijesekera, J. Reardon, I. Reyes, L. Tsai, J.-W. Chen, N. Good, D. Wagner, K. Beznosov, and S. Egelman, "Contextualizing privacy decisions for better prediction (and protection)," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–13. [Online]. Available: https://doi.org/10.1145/3173574.3173842
[33] S. B. Green, "How many subjects does it take to do a regression analysis," *Multivariate behavioral research*, vol. 26, no. 3, pp. 499–510, 1991.
[34] (2022) Statistisches Bundesamt. [Online]. Available: https://www.destatis.de/
[35] M. Häring, E. Gerlitz, C. Tiefenau, M. Smith, D. Wermke, S. Fahl, and Y. Acar, "Never ever or no matter what: Investigating Adoption Intentions and Misconceptions about the Corona-Warn-App in Germany," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, Aug. 2021, pp. 77–98. [Online]. Available: https://www.usenix.org/conference/soups2021/presentation/acar
[36] S. Altmann, L. Milsom, H. Zillessen, R. Blasone, F. Gerdon, R. Bach, F. Kreuter, D. Nosenzo, S. Toussaert, and J. Abeler, "Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study," *JMIR mHealth and uHealth*, vol. 8, no. 8, p. e19857, 2020.
[37] S. Degli Esposti, K. Ball, and S. Dibb, "What's in it for us? benevolence, national security, and digital surveillance," *Public Administration Review*, vol. 81, no. 5, pp. 862–873, 2021.
[38] L. Zhang-Kennedy and S. Chiasson, ""whether it's moral is a whole other story": Consumer perspectives on privacy regulations and corporate data practices," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, 2021, pp. 197–216.
[39] S. Barth, M. D. de Jong, and M. Junger, "Lost in privacy? online privacy from a cybersecurity expert perspective," *Telematics and Informatics*, vol. 68, p. 101782, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736585322000156
[40] (2022). [Online]. Available: https://www.statista.com/statistics/1059426/social-media-usage-germany/
[41] (2022). [Online]. Available: https://gs.statcounter.com/os-market-share/all/germany

APPENDIX A
LEAKED MEMO

Team Apple,

I wanted to share a note of encouragement to say that everyone at NCMEC is SO PROUD of each of you and the incredible decisions you have made in the name of prioritizing child protection.

It's been invigorating for our entire team to see (and play a small role in) what you unveiled today.

I know it's been a long day and that many of you probably haven't slept in 24 hours. We know that the days to come will be filled with the screeching voices of the minority.

Our voices will be louder.

Our commitment to lift up kids who have lived through the most unimaginable abuse and victimizations will be stronger.

During these long days and sleepless nights, I hope you take solace in knowing that because of you many thousands of sexually exploited victimized children will be rescued, and will get a chance at healing and the childhood they deserve.

Thank you for finding a path forward for child protection while preserving privacy.

source: https://9to5mac.com/2021/08/06/apple-internal-memo-icloud-photo-scanning-concerns/

APPENDIX B
SURVEY

Demographics

- **Q1** Please specify your gender: [Male/ Female/ Non binary/ I would like to describe myself: *(Free text)*/ I would rather not say]
- **Q2** How old are you? [*(Free text)*]
- **Q3** Which state do you live in? [Baden-Württemberg/ Bavaria/ Berlin/ Brandenburg/ Bremen/ Hamburg/ Hesse/ Mecklenburg-Vorpommern/ Lower Saxony/ North Rhine-Westphalia/ Rhineland-Palatinate/ Saarland/ Saxony-Anhalt/ Saxony/ Schleswig-Holstein/ Thuringia]

Scenarios

- **Q4** How much privacy do you think there is on the Internet right now? [There is much privacy/ There is some privacy/ There is not much privacy/ There is no privacy at all/ I do not know/ I would rather not say]

Neutral (all groups): For the following part of the questionnaire, please imagine the following scenario: There is software that allows mass automated search for files relevant to criminal activities on devices such as smartphones or laptops. If the software finds suspicious material, then the law enforcement authorities are informed.

Positive priming (only PG): Proponents of this software point out that serious crimes such as child abuse or terrorism can be solved or even prevented with this technology. This is particularly relevant in the fight against organized crime and crimes where victims are sought on the Internet.

Negative priming (only NG): Critics of this software point out that this represents mass surveillance and that what is considered criminally relevant can be too easily expanded and abused. They also point out that errors in the software can cause innocent people to be suspected and then bear negative consequences, such as interrogations and blocked accounts.

Introduction (all groups): In the next part of the questionnaire, you will be presented with various proposals for implementing the software described previously. Please indicate whether you approve or disapprove of each approach. The measures might only be slightly different from each other. Therefore, please read them carefully. Differences are highlighted in **bold**.

- **Q-Scenarios**: A pre-installed software from the manufacturer automatically searches the **[images/text messages]** in/on your **[device (e.g. smartphone, laptop...)/ cloud storage]**. In case of suspicion, an investigation is carried out by a **[law enforcement agency/ private company (e.g. Apple, Google, ...)]**. Would you approve or disapprove of this measure if the software was implemented to fight the following crimes?
  - Drug consumption [I disapprove/ I somewhat disapprove/ I neither approve nor disapprove/ I somewhat approve/ I approve]
  - Terrorism [I disapprove/ I somewhat disapprove/ I neither approve nor disapprove/ I somewhat approve/ I approve]
  - Child abuse [I disapprove/ I somewhat disapprove/ I neither approve nor disapprove/ I somewhat approve/ I approve]
  - Drug trafficking [I disapprove/ I somewhat disapprove/ I neither approve nor disapprove/ I somewhat approve/ I approve]
  - Tax fraud [I disapprove/ I somewhat disapprove/ I neither approve nor disapprove/ I somewhat approve/ I approve]
- **Q5.1** You have just indicated that you support one or more of the measures just presented. Are you concerned about possible misuse of the technology? [Yes/ Somewhat yes/ Undecided/ Somewhat no/ No]
- **Q5.2** You have just indicated that you oppose one or more of the measures just presented. How would you justify your attitude? [*(Free text)*]

Background

- **Q6** How many people live in your household in total, i.e. including you: [1 person/ 2 people/ 3 people/ 4 people or more/ I would rather not say]
- **Q7** Are any children under the age of 18 currently living in your household? [Yes/ No/ I would rather not say]
- **Q8** How many children under the age of 18 are currently living in your household? [1 child/ 2 children/ 3 children/ 4 children or more/ I would rather not say]
- **Q9** What age is this child/ are these children? [Under 2 years/ 2-3 years/ 4-5 years/ 6-11 years/ 12-13 years/ 14-17 years/ I would rather not say]
- **Q10** Does one or more of the following roles describe your relationship with a child under the age of 18? I am a... [Grandparent/ Parent/ Sister or Brother/ Aunt or Uncle/ Cousin/ Caregiver/ Teacher/ Youth Advisor/ Other relationship: *(Free text)*/ I stand in no relation to a child under 18 years of age/ I would rather not say]
- **Q11** What is your highest level of education? [Trade, technical or vocational training/ High school graduate, diploma or the equivalent (for example: GED)/ Some college credit, no degree/ Bachelor's degree/ Master's

degree/ Doctorate degree/ Other degree and namely: *(Free text)*/ No schooling completed/ I would rather not say]

- **Q12** What is your monthly net household income? [under 1300 Eur/ 1300 to 1699 Eur/ 1700 to 2599 Eur/ 2600 to 3599 Eur/ 3600 to 5000 Eur/ over 5000 Eur/ I would rather not say]
- **Q13** What is your professional status? [Student/ College student/ Employee/ Self-employed/ Freelancer/ Job seeker/ Retired/ Other: *(Free text)*/ I would rather not say]
- **Q14** Do you have specific computer skills such as: system administration, programming, IT security, tech support, power user, etc.? [Yes/ No/ I would rather not say]

Other attitudes

- **Q15** To what extent do you agree or disagree with the following statement: "In general, I trust the government to do the right thing." [I agree/I somewhat agree/ I neither agree nor disagree/ I somewhat disagree/ I disagree]
- **Q16** To what extent do you agree or disagree with the following statement: "In general, I trust government agencies to obey the law." [I agree/I somewhat agree/ I neither agree nor disagree/ I somewhat disagree/ I disagree]
- **Q17** To what extent do you agree or disagree with the following statement: "In general, I trust companies like Google, Apple, etc. to obey the law." [I agree/I somewhat agree/ I neither agree nor disagree/ I somewhat disagree/ I disagree]
- **Q18** To what extent do you agree or disagree with the following statement: "In general, I trust that law enforcement officials would treat me fairly." [I agree/I somewhat agree/ I neither agree nor disagree/ I somewhat disagree/ I disagree]
- **Q19** Do you use cloud storage (e.g. iCloud, Dropbox, Google Drive, OneDrive)? [Yes/ No/ I would rather not say]
- **Q20** Do you use a smartphone? [Yes, an Android/ Yes, an iPhone/ Yes, another smartphone: *(Free text)*/ Yes, but I do not know which/ No/ I would rather not say]
- **Q21** On which side of the debate between surveillance for law enforcement vs. privacy do you see yourself? [Rather pro law enforcement/ Rather pro privacy/ Neither side/ I am not familiar with the debate/ I would rather not say]

APPENDIX C
DEMOGRAPHICS BY SCENARIO

Tables VI and VII contain a detailed breakdown of the demographics by scenario.

TABLE VI
OVERVIEW OVER DEMOGRAPHICS, SPLIT BY SCENARIO (TABLE 1 OF 2)

Scenario		Overall (n=1062)	Image/ Personal Device/ Law enforcement agency (n=799)	Image/ Personal Device/ Private company (n=802)	Image/ Cloud storage/ Law enforcement agency (n=799)	Image/ Cloud storage/ Private company (n=795)
Gender	Female:	51.1%	51.1%	51.2%	51.8%	52.5%
	Male:	48.7%	48.7%	48.6%	48.1%	47.4%
	Diverse:	0.2%	0.3%	0.1%	0.1%	0.1%
Age	under 20:	0.8%	0.9%	0.7%	0.6%	0.8%
	20 to 39:	30.7%	30.8%	30.7%	31.3%	29.8%
	40 to 59:	35.3%	36.4%	35.2%	34.4%	35.1%
	60 to 79:	32.2%	30.9%	32.5%	32.7%	33.3%
	80 to 99:	1.0%	1.0%	0.9%	1.0%	1.0%
State of residency	Baden-Württemberg:	12.4%	12.5%	12.5%	12.3%	13.1%
	Bavaria:	15.7%	15.8%	15.8%	15.1%	16.0%
	Berlin:	4.5%	4.8%	4.6%	4.5%	4.4%
	Brandenburg:	2.9%	2.8%	2.9%	2.8%	2.6%
	Bremen:	0.7%	0.6%	0.6%	0.6%	0.6%
	Hamburg:	2.4%	2.6%	2.4%	2.5%	2.0%
	Hesse:	7.7%	7.9%	8.1%	7.9%	7.5%
	Lower Saxony:	9.3%	9.3%	9.2%	9.3%	9.2%
	Mecklenburg-Vorpommern:	2.4%	2.1%	2.2%	2.6%	2.1%
	North Rhine-Westphalia:	21.0%	19.9%	20.4%	22.4%	21.4%
	Rhineland-Palatinate:	4.9%	4.9%	4.6%	4.0%	4.8%
	Saarland:	1.3%	1.8%	1.5%	1.4%	1.1%
	Saxony:	8.0%	8.1%	8.6%	8.6%	7.9%
	Schleswig-Holstein:	3.5%	3.8%	3.2%	3.3%	3.8%
	Thuringia:	2.7%	2.8%	2.6%	2.3%	2.8%
	n.a.:	0.6%	0.5%	0.6%	0.5%	0.6%
Monthly income	below €1300:	13.3%	12.1%	14.0%	13.1%	12.7%
	€1300-1699:	9.7%	10.4%	9.5%	9.3%	9.9%
	€1700-2599:	21.2%	22.0%	20.3%	20.7%	21.8%
	€2600-3599:	19.6%	18.9%	19.6%	19.9%	19.0%
	€3600-5000:	19.0%	19.6%	18.5%	19.0%	19.1%
	above €5000:	8.2%	8.3%	8.4%	8.6%	7.9%
	n.a.:	9.0%	8.6%	9.9%	9.4%	9.6%
Degree	College:	8.1%	8.5%	8.7%	8.4%	6.9%
	College or certified engineer:	5.5%	5.4%	5.2%	5.9%	5.2%
	Graduate:	21.5%	20.9%	21.6%	22.2%	22.3%
	High school:	12.6%	12.5%	12.7%	12.8%	12.1%
	No vocational education:	2.9%	2.9%	3.4%	3.3%	2.6%
	Technical school:	5.3%	5.3%	5.1%	4.5%	5.9%
	Vocational diploma:	13.1%	13.1%	11.6%	13.6%	13.6%
	Vocational training:	28.7%	29.2%	29.2%	27.5%	29.2%
	other:	1.3%	1.4%	1.5%	0.9%	1.4%
	n.a.:	1.0%	0.9%	1.0%	1.0%	0.9%

TABLE VII
OVERVIEW OVER DEMOGRAPHICS, SPLIT BY SCENARIO (TABLE 2 OF 2)

Scenario		Overall (n=1062)	Text message/ Personal Device/ Law enforcement agency (n=794)	Text message/ Personal Device/ Private company (n=795)	Text message/ Cloud storage/ Law enforcement agency (n=798)	Text message/ Cloud storage/ Private company (n=790)
Gender	Female:	51.1%	51.3%	50.3%	50.6%	50.3%
	Male:	48.7%	48.6%	49.4%	49.1%	49.5%
	Diverse:	0.2%	0.1%	0.3%	0.3%	0.3%
Age	under 20:	0.8%	0.8%	0.8%	0.6%	0.9%
	20 to 39:	30.7%	30.5%	30.8%	31.1%	30.6%
	40 to 59:	35.3%	35.3%	35.8%	34.6%	35.7%
	60 to 79:	32.2%	32.6%	31.4%	32.6%	31.5%
	80 to 99:	1.0%	0.9%	1.1%	1.1%	1.3%
State of residency	Baden-Württemberg:	12.4%	12.7%	13.0%	11.4%	12.0%
	Bavaria:	15.7%	16.0%	15.7%	14.9%	16.5%
	Berlin:	4.5%	4.4%	4.3%	4.5%	4.7%
	Brandenburg:	2.9%	3.0%	2.6%	3.6%	3.0%
	Bremen:	0.7%	0.8%	0.6%	0.9%	0.5%
	Hamburg:	2.4%	2.4%	2.6%	2.5%	1.8%
	Hesse:	7.7%	7.8%	6.9%	7.9%	7.7%
	Lower Saxony:	9.3%	9.6%	9.4%	10.0%	8.6%
	Mecklenburg-Vorpommern:	2.4%	1.9%	2.4%	2.8%	2.7%
	North Rhine-Westphalia:	21.0%	20.3%	21.4%	21.6%	20.6%
	Rhineland-Palatinate:	4.9%	5.7%	5.3%	4.4%	5.6%
	Saarland:	1.3%	1.3%	0.9%	1.1%	1.5%
	Saxony:	8.0%	7.9%	7.5%	7.5%	7.7%
	Schleswig-Holstein:	3.5%	3.0%	3.6%	3.1%	4.1%
	Thuringia:	2.7%	2.8%	2.9%	3.3%	2.5%
	n.a.:	0.6%	0.5%	0.8%	0.5%	0.5%
Monthly income	below €1300:	13.3%	13.5%	14.6%	13.9%	12.3%
	€1300-1699:	9.7%	10.7%	8.6%	9.0%	10.3%
	€1700-2599:	21.2%	21.5%	21.9%	20.6%	20.8%
	€2600-3599:	19.6%	18.6%	20.1%	20.9%	19.6%
	€3600-5000:	19.0%	18.9%	18.4%	19.7%	19.0%
	above €5000:	8.2%	8.3%	7.8%	8.3%	8.0%
	n.a.:	9.0%	8.4%	8.7%	7.6%	10.1%
Degree	College:	8.1%	7.7%	7.8%	8.3%	8.5%
	College or certified engineer:	5.5%	5.7%	4.9%	6.0%	5.4%
	Graduate:	21.5%	20.8%	22.5%	20.9%	20.6%
	High school:	12.6%	13.7%	12.8%	12.8%	11.5%
	No vocational education:	2.9%	2.5%	2.9%	2.4%	3.4%
	Technical school:	5.3%	5.9%	5.2%	5.1%	5.2%
	Vocational diploma:	13.1%	13.6%	13.1%	13.2%	12.9%
	Vocational training:	28.7%	27.7%	28.4%	28.8%	29.7%
	other:	1.3%	1.3%	1.3%	1.5%	1.4%
	n.a.:	1.0%	1.1%	1.1%	1.0%	1.3%