Data Breaches: User Comprehension, Expectations, and Concerns with Handling Exposed Data

Sowmya Karunakaran, Kurt Thomas, Elie Bursztein, and Oxana Comanescu, *Google*

https://www.usenix.org/conference/soups2018/presentation/karunakaran

This paper is included in the Proceedings of the
Fourteenth Symposium on Usable Privacy and Security.

August 12–14, 2018 • Baltimore, MD, USA

ISBN 978-1-931971-45-4

Open access to the Proceedings of the
Fourteenth Symposium
on Usable Privacy and Security
is sponsored by USENIX.

Data Breaches: User Comprehension, Expectations, and Concerns with Handling Exposed Data

Sowmya Karunakaran Kurt Thomas Elie Bursztein Oxana Comanescu
Google Inc.
{sowmyakaru, kurtthomas, elieb, oxana}@google.com

ABSTRACT

Data exposed by breaches persist as a security and privacy threat for Internet users. Despite this, best practices for how companies should respond to breaches, or how to responsibly handle data after it is leaked, have yet to be identified. We bring users into this discussion through two surveys. In the first, we examine the comprehension of 551 participants on the risks of data breaches and their sentiment towards potential remediation steps. In the second survey, we ask 10,212 participants to rate their level of comfort towards eight different scenarios that capture real-world examples of security practitioners, researchers, journalists, and commercial entities investigating leaked data. Our findings indicate that users readily understand the risk of data breaches and have consistent expectations for technical and non-technical remediation steps. We also find that participants are comfortable with applications that examine leaked data—such as threat sharing or a "hacked or not" service—when the application has a direct, tangible security benefit. Our findings help to inform a broader discussion on responsible uses of data exposed by breaches.

1. INTRODUCTION

In recent years, data breaches have exposed the online credentials and personal data of billions of users across the Internet. In 2017 alone, news headlines announced that criminals had stolen usernames and passwords for 3 billion Yahoo users [16], the financial details of 143 million Americans collected by Equifax [10], and private data belonging to 57 million Uber users [17]. Once stolen, this data becomes readily accessible via black markets. Previous studies have identified over 3.3 billion credentials from breaches freely traded on the underground along with credit cards and other financial data [7, 25, 26]. Exposure puts victims at further risk of account takeover, financial theft, identity theft, or worse.

Despite repeated data leaks due to breaches, best practices for how companies should respond to incidents have yet to be formalized. One common remediation step—requested by both victims and increasingly by regulators [1, 3, 19, 22]—is that companies notify any affected victim. However, evidence that notifications influence user behavior is limited. For example, victims do not opt to switch to other, more secure services [1, 4, 14]. Moreover, companies do not always notify victims in a timely manner: Uber waited over a year before disclosing a $100,000 ransom payment in response to a breach [9].

At the same time, there are no clear boundaries for how one should responsibly handle data after it is leaked. Some security systems examine third-party breaches to protect victims from further harm: Google, Facebook, and Netflix automatically reset passwords for victims appearing in password dumps [2, 29]. Others provide information to victims, such as leak aggregation services that collect exposed credentials to help notify victims [13]. Exposed data also plays a role in the construction of password strength meters and investigations of underground market activity [5, 6, 18, 24, 28]. How victims weigh any potential security benefits against other concerns, including their privacy, remains uncertain.

In this paper, we bring users into the discussion of how companies should respond to breaches and how user data should be respected even after it finds its way on to black markets. We do this through two surveys. In the first, we asked 551 participants what actions a company should take upon learning of a data breach including technical solutions such as forcing password resets or enabling two-factor authentication. In the second part of our study, we surveyed 10,212 participants across six countries to assess their level of comfort and concerns towards eight scenarios capturing real-world situations where security practitioners, researchers, journalists, and commercial entities investigated data exposed by breaches. While we focus on lessons for the security community, we include these latter categories to act as a baseline comparison.

We frame our key findings as follows:

Data breaches feature prominently in the public's mind share: Over 93% of participants understood the meaning of a data breach. These participants cited identity theft (52%), the loss of personal information (25%), and monetary loss (9%) as their top concerns.

Notifications remain the most popularly requested remediation step: 83% of participants requested that companies affected by a breach send an immediate notifi-

Copyright is held by the author/owner. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee.
USENIX Symposium on Usable Privacy and Security (SOUPS) 2018.
August 12–14, 2018, Baltimore, MD, USA.

cation to victims. Other more technical requests included enabling two-factor authentication on accounts (63%) and resetting exposed passwords (61%).

Users are supportive of applications that consume exposed data if they provide a direct security benefit: Of the 8 scenarios we examined, users were most comfortable with proactive password resetting in the event of reuse and sharing information with other identity providers.

Past experience as a victim of a breach increases support for security use cases: We observed significant differences in a prior victim's vs. non-victim's level of comfort for security related use cases. For example, 44% of prior victims expressed comfort with proactive password resetting compared to 34% of non-victims.

Users are wary of interacting with criminals (such as purchasing exposed data), but recognize the potential security benefits: For non-security use cases, over 70% of participants negatively expressed that purchasing exposed data was unethical and incentivized criminal behavior. However, with security related use cases only 40–51% participants expressed similar negative concerns.

Support for security use cases is consistent across countries: Although we observed significant differences in the absolute comfort levels between countries, every country consistently weighed security use cases over non-security use cases in terms of comfort.

2. RELATED WORK

Our work builds upon prior research into the experiences of victims of data breaches. In a study similar to ours, Ablon et al. surveyed 6,000 participants from the United States in 2015 and found 44% reported having received a data breach notification [1]. Credit card details topped the list of exposed data (49%), but health information (21%), social security numbers (17%), and account details (13%) also featured prominently. Reactions from participants to breaches were varied: only 11% of those surveyed stopped interacting with the affected company. More commonly, victims changed their password or PIN (51%) or switched to a new account (24%), while another 22% of participants did nothing at all. Other studies have also found that users rarely switch to another service or stop interacting with a company even upon receiving a breach notification [4, 14]. Our study examines in greater detail the expectations of breach victims and the technical remedies they most strongly prefer.

While financial theft features prominently in the concerns of victims, account takeover is also a significant risk. A survey by Shay et al. found that 15.6% of 1,502 survey participants self-reported having their account taken over [23]. A similar study by Rainie et al. found 21% of 1,002 adults experienced a social network or email account being hijacked [21]. These common experiences stem from billions of usernames and passwords exposed due to data breaches, with Thomas et al. estimating that data breach victims are 11.6x more likely to fall victim to account takeover than a random sample of users [25]. The prevalence of account takeover heavily influences the design of our study scenarios.

3. METHODOLOGY

We conducted two online surveys to evaluate user comprehension, attitudes, and expectations around data breaches. We describe each survey in detail. We refer readers to the Appendix for the full structure and text of both surveys.

3.1 Survey on responding to breaches (N=551)

Our first survey gauged user perceptions of risk surrounding breaches and how users would want a company to respond if their data had been exposed. We recruited participants via Amazon Mechanical Turk in July 2017 and administered the survey through Google Forms. Participants were asked to take a "simple task and experience survey." We avoided using the term "data breach" to prevent nonresponse bias. The survey took approximately 3 minutes to complete and participants were each compensated $0.50, including the screened out participants. In total, we received 604 responses, of which 551 feature in our final analysis.

3.1.1 Survey structure

We began the survey with a single screener question with three possible definitions of a data breach. The ordering of these options was randomized.

- Public exposure of usernames and passwords of millions of users of an online system. (N=564)
- Using large sets of data to aid robots to solve a problem that humans cannot solve. (N=13)
- Web page that is unable to load due to too much data on the page. (N=27)

Overall, 564 of 604 participants chose the correct definition and were allowed to continue through the rest of the survey. We dropped the remaining 40 participants from any further questions.

Following the screener, we asked participants to select the single most important "harm" that might arise from their password being exposed through a data breach and what remediation steps a company should take to protect the participant's account. Finally, we asked participants to rate their level of comfort with six potential actions a company could take in response to a breach. For each action, in addition to rating their level of comfort, participants provided an open-ended reason for their rating.

Outside these core questions, we asked whether participants had ever been the victim of a data breach. We also included two quality control questions, and six demographic questions. In total, we eliminated 13 inattentive responses where participants answered both quality control questions incorrectly, leaving a total of N=551 responses.

We reviewed a small sample (N = 50) of open-ended responses and developed codes. The rest of the open ended responses were then assigned codes through manual inspection. Responses that did not fall into any of the coding buckets were categorized under 'Other'. Roughly 3% of responses were blank which we did not categorize. The researcher not involved in the coding process conducted the quality checks by independently reviewing a sub-sample. The agreement rate was about 90%.

3.1.2 Survey development

Prior to running the survey, we conducted an initial pilot (N=34) where the single most important harm was left as an open-ended question. We then codified the most popular responses, selecting eight possible options for the final survey. We also expanded the list of remediation steps to include new, incorrect steps (e.g., buying a new computer) to gauge comprehension. We also switched from strictly asking each participant's comfort towards certain responses to also requesting their reasoning. We then ran a second pilot (N=31). We used the open-ended responses to clarify the six actions a company might take in response to a breach. Finally, we added a demographic question related to whether participants had ever been a victim of a previous breach.

3.1.3 Participant demographics

For the 551 participants, 52% identified as male, 47% identified as female, and 1% preferred not to answer. Roughly 12% were 18–24, 45% 25–34, 22% 35-44, 13% 45–54, 6% 55–64, and 2% older than 65 or preferred not to say. In terms of education, 47% had a bachelors degree, 19% a masters degree or higher, and 17% some college education. Participants predominantly resided in the United States—69%—with another 23% residing in India and 8% in other countries. In terms of employment, 80% were had some form of employment (53% full-time, 17% self-employed, and 10% part time), 8% were students, and 12% were unemployed, retired, or looking for work.

3.1.4 Limitations

In terms of the study sample, although the user population on Mechanical Turk is relatively diverse for an Internet sample, there is still a bias. For example, the Mechanical Turk workers are considered WEIRD (Western, educated, industrialized, rich, and democratic) [15]. To reduce the effect of this bias, we opened the survey to residents of all countries, not just United States residents. However, the underlying demographics of workers still skews towards the United States and India.

3.2 Survey on breach data use cases (N=10,212)

Our second survey examined user comfort towards a spectrum of use cases that handle data exposed by breaches. We recruited participants through an international panel provider that recruits through online communities, social networks and the web. The panel provider also enforced strict quality controls such as digital fingerprinting to identify duplicate participants and pattern recognition to flag fraudulent responses. As such, we do not embed any quality control questions in the survey questions. We specifically stratified our sample to participants from the United States, Canada, United Kingdom, Australia, India, and Germany. We administered the survey using the online survey platform and panel provider Qualtrics. We paid $6 per response to our panel provider, a portion of which was paid to the participants as incentive.

3.2.1 Survey structure

We used a scenario based survey to frame eight potential use cases of data exposed by breaches. To minimize fatigue, each survey was structured to included only two scenarios randomly selected from our pool of eight. When considering a scenario, we asked users to rate their level of comfort if they knew the data had been purchased from criminals via a black market; to explain their rating in an open-ended question; and finally whether their level of comfort would change if they knew the data was freely available. We also included six demographic questions and one question on whether the participant had previously been a victim of a breach. We outline each scenario and highlight real-world equivalents. In total, we received 10,212 responses, with over 400 responses per scenario and per country.

Security research (S1): In the first scenario, we framed whether it was acceptable for a researcher at a university to use data exposed by a breach to study how users select passwords. Examples of such research in practice include studies of password reuse [5] and the development of better password strength meters from existing, exposed data [28, 6, 18].

Hacked or not service (S2): We asked participants whether it was acceptable for a company to provide a paid service where anyone could query for a "username" to determine whether their data was exposed due to a breach. A multitude of such services currently exist, such as *haveibeenpwned.com*, *breachalarm.com*, and *leakedsources.com*. In practice, some of these services operate on donations and only reveal whether an account was present in a breach. Others require a monthly fee and allow a subscriber to look up any username and its associated passwords, at times running afoul of law enforcement [20].

Threat sharing, finance and social (S3, S4): For two scenarios, we asked whether participants were comfortable with a breached company sharing the email addresses of victims with third-party services to protect against lateral attacks. We offered two, independent scenarios for the third-party service involved: a financial institution and an online social network. These scenarios mimic emerging threat exchange services where companies share information on ongoing attacks.

Proactive password resetting (S5): We asked participants whether they were comfortable with a service finding usernames and passwords exposed in third-party breaches to proactively re-secure the participant's account if they reused an exposed password. This scenario matches how Google, Facebook, and Netflix currently reset passwords for victims appearing in third-party breaches [29, 2].

Journalist, tax fraud (S6): We framed whether participants were comfortable with a journalist writing an article on tax evasion that sourced their materials from private emails exposed due to a breach. Rough equivalents include the Panama Papers [11] and Paradise Papers [8] that exposed millions of email records detailing the financial dealings of offshore investments and entities.

Journalist, dating site (S7): We examined whether it was appropriate for a journalist to use personal information from breached data profiles as source material for an article. Recent examples include the leak of Ashley Madison users, which media outlets used to expose the activities of registered members.

Competitor (S8): We framed whether it was appropriate for a non-breached company to contact victims in order to advertise switching services. For example, after the Equifax breach, one identity theft provider created ads and press released to announce how it could help victims [12].

3.2.2 *Survey development*

Prior to running our survey, we conducted two pilots. The first involved user researchers at our institution who provided feedback on the framing text of the scenarios. The second pilot involved a small sample of participants (N=40). Based on the responses, we added a follow-up question for every scenario to understand whether a participant's comfort would change if data was freely available.

3.2.3 *Participant demographics*

For the 10,212 participants, 51% identified as male, 48% female, and 1% preferred not to answer. In terms of age, 11% were 18–24, 28% 25–34, 19% 35-44, 17% 45–54, 13% 55-64, and 9% older than 65. Participants were equally distributed across six countries: 16% in Australia, 18% in Canada, 17% in Germany, 14% in India, 16% in the United Kingdom, and 15% in the United States. 46% indicated to be employed full-time, 13% employed part-time, 13% retired, 5% students, 7% self-employed, 7% home makers, 5% unemployed and 4% other. In terms of education, 5% indicated receiving less than high school education, 17% High School, 18% Some college no degree, 15% Associate's degree, 28% Bachelor's degree, 12% Master's degree, 1% Ph.D and 3% Other.

3.2.4 *Limitations*

Our surveys were spread across several weeks, however we could not control for respondent's exposure to external information such as news stories and press articles on data breaches. In addition, given that our approach relies on scenarios based assessment, one can argue the presence of availability bias. Availability heuristic is a mental shortcut that relies on immediate examples that come to a given person's mind when evaluating a specific topic, concept, method or decision [27]. In reality, users would have access to many other pieces of input about the scenario at hand which may also play a role in influencing their level of comfort. Gut reactions and framing may also influence the perceived acceptability of the scenarios we explored. Likewise, privacy enhancing technologies might help to allay user concerns with respect to data sharing.

4. RESPONDING TO A DATA BREACH

We report on the results of our first survey, which explored the familiarity of participants with data breaches as both a concept and a personal experience. We present how participants perceived the risk of breaches, what actions they felt companies should take in response to a breach to protect victims, and finally their level of comfort with companies engaging with the press, government, criminals, and other companies as part of remediation.

Comprehension: As a first step towards interpreting our results, we examined whether participants were familiar with data breaches and their accompanying risk. The vast majority of participants (N=564, 93%) correctly identified the definition of a breach from one of three choices. As shown in Table 1, their top concerns of what harm might arise included identity theft (N=287, 52%) and the leak of personal information (N=138, 25%). Monetary loss was a distant third (N=50, 9%), possibly due to our framing of data breaches as relating to usernames and passwords. Impossible harms, such as a participant's computer being infected with a virus, were selected by only 17 participants (3%). More than a hypothetical experience, 232 participants (42%) reported having had their data exposed by a prior breach while 65 participants (12%) reported not knowing. These results suggest that participants are both familiar with the concept of a data breach and the resulting consequences.

Table 1: Top harm that results from a data breach.

Potential harm	Breakdown	N
Identity theft	52%	287
Leak of personal information	25%	138
Monetary loss	9%	50
Loss of access to personal information	5%	28
Phone being monitored by hackers	3%	17
Computer being infected with virus	3%	17
Spam being sent out from your account	2%	11
Other	1%	4
No harm	< 1%	2

Table 2: Ranking of remediation steps companies should take in response to a breach.

Remediation step	Breakdown	N
Send you an immediate notification	83%	457
Enable two-factor authentication	63%	347
Reset your password	61%	336
Provide credit monitoring	56%	309
Issue a refund	39%	215
Give you a new account	32%	176
Change your username	31%	171
Pay users a consolation bonus for breaking their trust	29%	160
Upgrade your web browser	15%	83
Company buys you a new computer	5%	28

Preferred remediation steps: Table 2 provides a breakdown of the remediation steps participants selected as the best ways companies could protect their account in the event of a breach. Participants most frequently requested that companies send an immediate notification to affected users (N=457, 83%). This was followed by more technical account protections such enabling two-factor authentication (N=347, 63%) and resetting an account's password (N=336, 61%). Some of these actions mirror steps that victims self-report taking in response to a breach, such as 51% of victims changing their password or PIN [1]. However, the same is not true for two-factor authentication: fewer than 3% of hijacking victims adopt two-factor authentication after learning their account was compromised [25]. This suggests a disconnect between understanding the protections two-factor authentication provides and actual adoption.

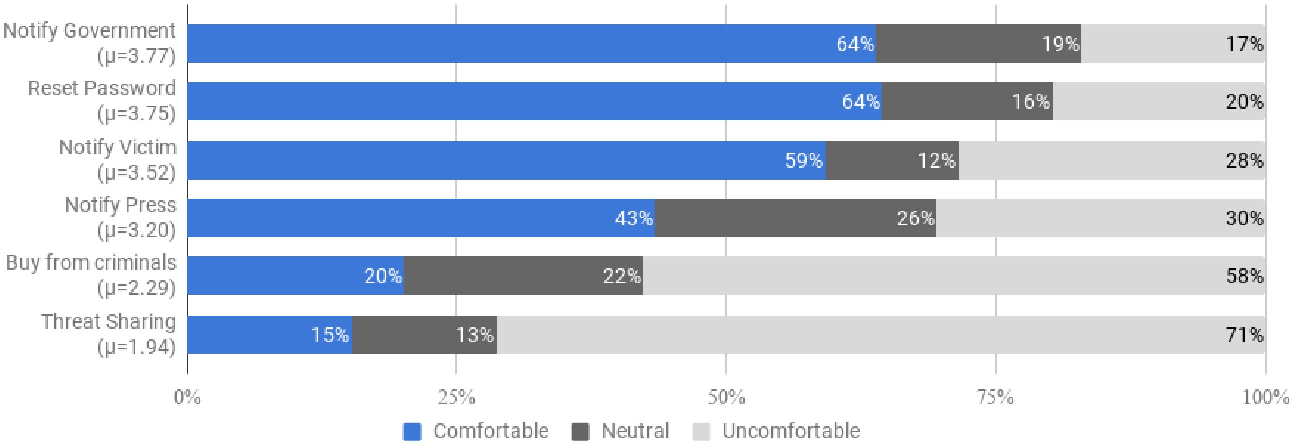

Figure 1: Comfort of participants towards breached companies dealing with victims, criminals, the press, the government, and other companies. We binned ratings of 1 or 2 as uncomfortable, 3 as neutral, and 4 or 5 as comfortable.

Account security measures and communication outranked financial protections, such as credit monitoring (N=309, 56%) or companies issuing a refund (N=215, 39%). This mirrors participants' perception of harm, where monetary loss ranked lower than identity theft or data loss. A small but not insignificant group of participants selected ineffective remediation steps that would provide no security benefit in the context of data breaches. These actions included changing usernames (N=171, 31%) or upgrading web browsers (N=83, 15%). The latter action suggests that users may conflate general security best practices such as keeping software up to date with something that might protect them from a breach.

Remediation and the wider ecosystem: Beyond user-centric remediation steps, we asked participants to rate how comfortable they were with companies taking a range of actions such as communicating with criminals, the press, the government, and other companies in response to a breach. We measured comfort on a scale of 1 to 5, with 1 indicating "Not at all comfortable" and 5 indicating "Very comfortable." We relied on two user-centric actions, namely resetting passwords and notifying victims, as a baseline comparison. Figure 1 shows the spectrum of ratings participants selected.

In the case of notifying victims of the breach, participants in aggregate rated the action with an average comfort level of $\mu = 3.52$. Common themes that correlated with a positive level of comfort—surfaced in the coded open-ended questions—included an obligation on the part of the company to be transparent (N=194, 36%) and that such a notification would allow participants to reset their password (N=46, 8%). Conversely, participants that were uncomfortable frequently cited that notifications made them feel insecure (N=63, 12%) and that it did nothing to make up for the loss of data (N=47, 9%). Neutral participants often cited that companies needed to do something more (N=45, 8%). For example:

> P474: "The notification is important, however, the company must also inform about the corrective measures it intends to take."

In comparison, participants were more favorable with notifying the government ($\mu = 3.77$), though less favorable of notifying the press ($\mu = 3.20$). The positive affinity towards government activity relates to prosecuting criminals and holding companies responsible (N=224, 41%):

> P355: "In order to prevent other breaches I think the government should be involved at helping catch the criminals responsible."

Unique concerns for reaching out to the press included feeling that victims should be contacted directly (N=77, 14%) and that headlines might attract criminals to take advantage of the exposed data (N=28, 5%).

> P406: "...making it too public may inspire others to try and take advantage of the breach".

Beyond notifications, a majority of participants expressed discomfort ($\mu = 2.29$) with companies reaching out to criminals to buy a copy of the leaked data to know what was exposed. Participants commonly cited that it was unethical to deal with criminals (N=89, 16%) and that it would incentivize further attacks (N=110, 20%):

> P24: "That shows the hackers that that company can be bullied, making them future targets for hacks."

Surprisingly, participants rated the prospect of companies sharing exposed usernames and passwords with other identity providers as the least comfortable action a company could take, lower even than dealing with criminals ($\mu = 1.94$). Common concerns included a violation of the participant's trust (N=205, 38%) and feeling it exacerbated the problem by exposing private information further (N=99, 18%):

> P338: "OMG no. I don't want my info shared!!!"
> P399: "The company has no permission to share my data, even if it was already stolen."

Table 3: Comparison of the level of comfort for past breach victims and non-victims. We note statistically significant differences with an astericks.

Remediation step	Comfort (victim)	Comfort (non-victim)	p-value
Notify government	3.90	3.65	0.018**
Reset password	3.79	3.73	0.443
Notify user	3.67	3.37	0.047**
Notify press	3.35	3.06	0.012**
Buy from criminals	2.23	2.35	0.217
Threat sharing	1.88	2.00	0.178

Taken as a whole, our findings indicate that participants are comfortable with actions that lead to better protections or even catching the criminals involved. However, participants expressed a strong degree of discomfort for actions that might further distribute exposed data or encourage future criminal activity. These concerns heavily influenced the design of our second survey (Section 5).

Influence of breach experiences on comfort: As an added dimension, we examined how prior experience with a data breach influenced a participant's level of comfort towards various actions. For our analysis, we treat participants that reported as being unsure if they had been part of a previous data breach as non-victims. Table 3 presents our results. Overall, victims reported a higher level of comfort for notifying users, the press, and the government than non-victims, while actions beyond notification saw no statistically significant difference.

5. HANDLING EXPOSED DATA

Turning to our second survey, we report how participants valued security applications built from exposed data and the trade-offs they perceived. We also examine how demographic variations and past experience with a breach influence a participant's level of comfort.

5.1 Scenarios, in depth

We provide a ranking of participant comfort to all eight scenarios in Figure 2. Participants were most comfortable with scenarios that helped to directly protect them from further risk, such as resetting reused passwords and working with other identity providers to prevent lateral attacks. In contrast, security protections that might help in the abstract, such as a "hacked or not" service or research in password security were rated lower. We explore each scenario (in order of comfort level) and the top concerns that participants surfaced through our open-ended questioner.

Threat Sharing, Finance: Participants were most comfortable when presented with a scenario of a breached company working with another identity provider—in this case a financial institution–to share threat intelligence of victims ($\mu = 2.94$). The stated goal of this sharing was to enable password resetting at the financial institution to protect victims from financial fraud. Based on our coded responses, participants most frequently expressed a lingering fear their financial assets remained at risk (56%) and skepticism resetting a password would dissuade criminals (19%). For one participant, this was an intimate experience:

> P[8920]: "[the breached company] owes explanation how my email got hacked in the first place and why they didn't protect me. This exact scenario happened to me with Yahoo and Paypal and somebody got into my account, took my Paypal credit card number and charged thousands of dollars at Walmart on it."

Despite these concerns, participants still remained neutral or positive on threat sharing as a minimum step towards responding to a breach. For example:

> P[7429]: "They are doing something to help fix a problem and partnering with a trusted company, so I have no objections to their being proactive."

Threat Sharing, Social Network: Similar to the previous threat sharing scenario, participants reported the second highest level of comfort when a social network was the recipient of threat intelligence ($\mu = 2.92$). Overall, participants most frequently cited privacy as their top concern (43%). Others felt that the security benefit outweighed any privacy concerns (20%) or welcomed the extra level of protection (22%):

> P[385]: "Of course which [sic] is also an invasion of my "privacy", but I find it a justified and proper engagement in order to protect other accounts before a hacker attack."
> P[7668]: "A proactive approach on the part of [the breached company] is likely the best means of blocking fraudulent activity and instituting counter-measures."

Proactive Password Resetting: When asked about a company purchasing third-party credential dumps from criminals to proactively protect against password reuse, 35% of participants reported being comfortable with such an activity ($\mu = 2.85$). Of participants, 53% stated this would enhance their security and another 16% that it was good to see proactive activity.

> P[9615]: This is a proactive step from [the company], and one that they are not actually obligated to do. This makes me feel like the company cares about protecting my identity.

However, another 25% of participants were concerned with the legality of such activity even if were beneficial, or whether it might encourage criminals:

> P[8941]: They're paying people who obtained the information illegally. This seems a bit odd, almost like they're encouraging people to hack sites.

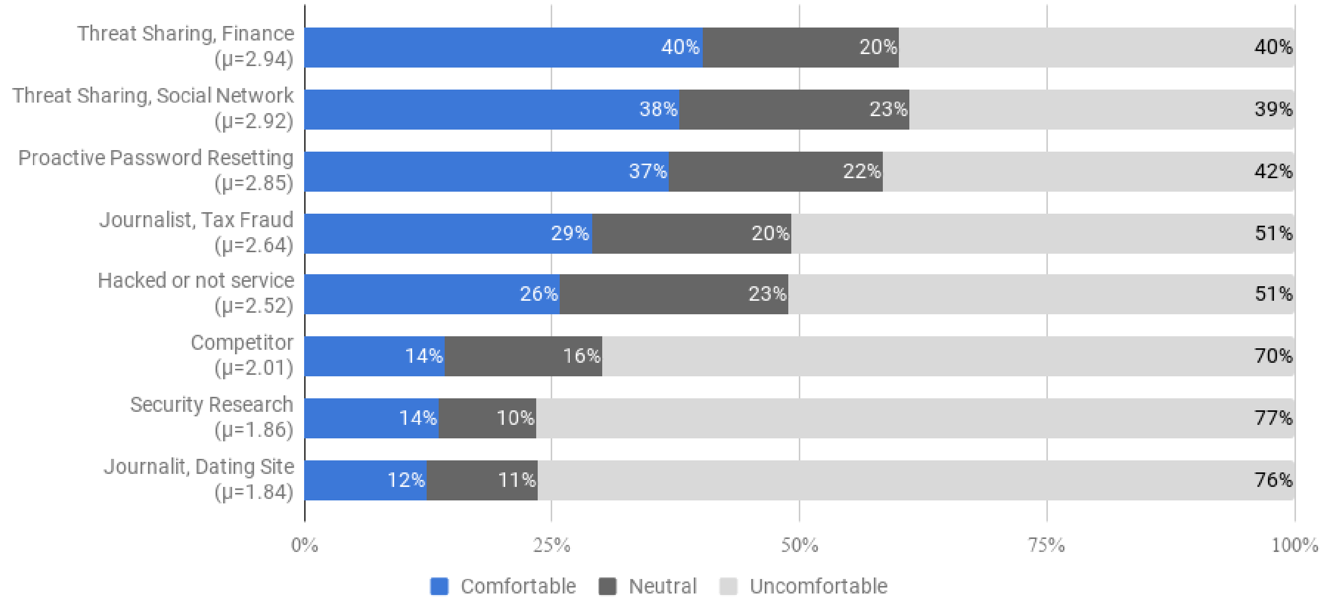

Figure 2: Participant comfort towards the eight scenarios involving purchasing (or where the source of data was not applicable). Participants ranked scenarios that provided direct security benefits higher than all other scenarios.

Less frequent, 4% of participants highlighted ethical concerns with any purchasing of data from criminals:

> P[575]: "I believe it's unacceptable for any company, whether their motives are good, to purchase or otherwise obtain illegally-gained data, especially personal information. ... It's a blatant disregard for people's privacy."

Surprisingly, participants were less comfortable ($\mu = 2.73$) when the data was freely available, a phenomenon also observed with the "hacked or not" service scenario as shown in Figure 3. We did not collect open-ended follow ups in conjunction with asking participants about freely available exposed data, so we cannot definitively state why this is the case. One hypothesis is that participants may have felt the damage is already done if credentials become freely available.

Journalist, Tax fraud: As a source of comparison, we asked participants their level of comfort towards journalists using data exposed by a breach to investigate fraud. Roughly 30% of participants reported being comfortable purchasing data from criminals to conduct such an investigation ($\mu = 2.64$). More participants expressed comfort when the data was freely available ($\mu = 2.77$). Participants frequently raised concerns about the legality of such behavior (56%):

> P[5414]: "Obtaining the information illegally doesn't make me feel comfortable. If it was handed to him for free, this feels a little less immoral."
> P[4607]: "It's important that the information be obtained and revealed, but [the journalist] has done so by potentially breaching the privacy of innocent individuals."

Others supported the journalist's actions, with the ends justifying the means:

> [7830]: "Even though the method is unethical, he is exposing a corruption. I will have to trade my uncomfortableness."
> P[5102]: "Whilst I wouldn't necessarily condone the hacking element, it is now a fact of modern society that these methods of information gathering are available. ... Publishing what was found through that means is in the public interest."

As with purchasing credential dumps, participants fall into a spectrum of ethical frameworks. For some, there is never a justification for using private data. For others, the value extracted from exposed data can override privacy concerns.

Hacked or not service: When asked to rate their level of comfort towards a service aggregating breaches to provide a "hacked or not" service, 25% of participants reported being comfortable ($\mu = 2.52$). As with proactive password resetting, comfort dropped when data was freely available ($\mu = 2.43$). Participants frequently cited the trustworthiness of the "hacked or not" service operator as their top concern (58%). Participants also felt any purchase would encourage criminals. In the words of participants:

> P[7550]: "It makes me fear that they work with the hackers and may not be trustworthy."

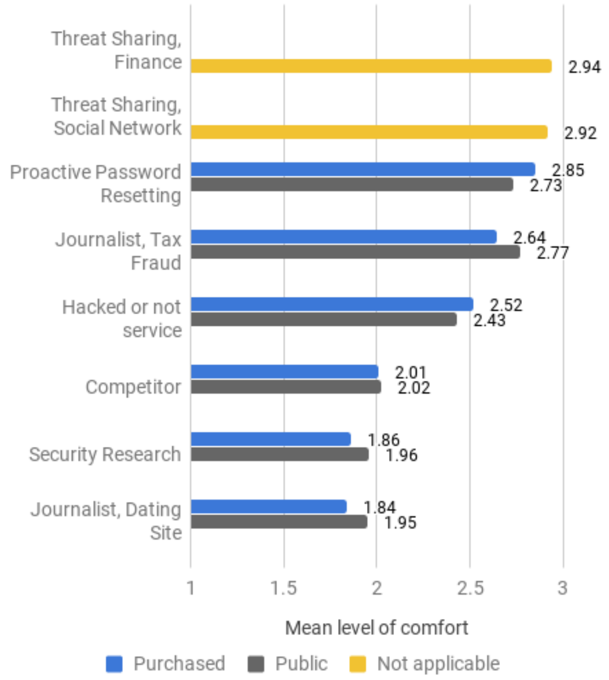

Figure 3: Comparison of comfort towards scenarios that involved purchasing data from the black market versus scenarios where data was freely available. All differences are statistically significant.

> P[4868]: "The fact that they are buying it from the hackers themselves is of concern....how do they know them? How are they getting the info? Do they have a relationship with the hackers?"

Despite these concerns, participants remained neutral on the prospect of such a service. A minority of participants (9%) highlighted the benefit of such a service, given a lack of clear notifications:

> P[10188]: "It is a good idea to see whether your account has been hacked. How else would you find out?"

Competitor: Only 14% of participants reported being comfortable with a competitor purchasing data from criminals in order to identify victims and offer for them to switch services ($\mu = 2.01$). There was little change in comfort if the data was freely available ($\mu = 2.02$). Participants frequently cited ethical concerns (44%) and the illegality of such behavior (39%).

> P[6145]: "This seems very unethical since they plan to gain profit from data that has been stolen."

However, 17% of participants expressed they would be better off in the end:

> P[622]: "This is a cheap shot to get consumers but at this stage I would probably go with [the competitor] as I know they have the proper software to avoid hackers."

Security Research: Faced with the prospect of researchers purchasing stolen credentials to study, participants reported the second lowest level of comfort compared to other scenarios ($\mu = 1.86$)—behind using stolen data for advertising. This comfort increased slightly when researchers obtained credentials from a free source ($\mu = 1.96$). Based on our coded responses, participants' top concern was the legality of the researchers actions (45%):

> P[9507]: "This may help his research and the result may help millions of internet users but the way he acquires the data is illegal and without the permission of account owners."

Other negative reactions included breaching the privacy of the victim (12%) and a sentiment that it was unethical to deal with criminals (9%):

> P[9307]: "It's incredibly unethical for [the researcher] to buy passwords from hackers. It's no different than someone buying a car that was stolen."

Another 23% of participants felt the value of research outweighed other concerns:

> P[7953]: "I have faith that this action will ultimately contribute to research that will make the general population less vulnerable in the long run."

These results suggest that, in the absence of a tangible security benefit, the privacy and ethical concerns of participants outweigh any potential justification. Surprisingly, research is viewed in even lower light than a scenario of a competitor advertising to victims of a breach, yet the latter still provides the prospect of a tangible benefit.

Journalist, Dating site:: A mere 13% of participants reported being comfortable with journalists using exposed dating profiles purchased from criminals to reveal the private lives of entities involved ($\mu = 1.84$). Comfort increased when data was freely available ($\mu = 1.95$). Most participants cited this was illegal (50%), a breach of privacy (27%), and unethical (11%).

> P[562]: "I find it very disconcerting that someone thinks they have a right to invade my space in any way without permission. It makes me want to withdraw from computors [sic] FULLSTOP. ... Makes for an unsafe unstable unfair dog eats dog world. Where has human respect, honesty and compassion gone."

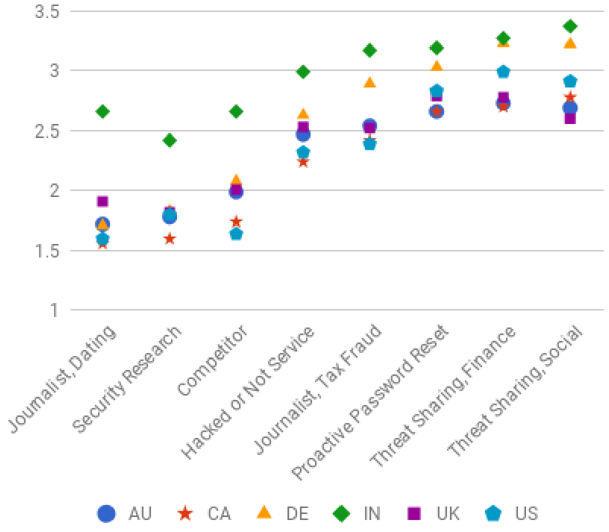

Figure 4: Mean level of comfort per scenario, broken down by country. Where applicable, comfort reflects the sub-scenario of purchasing data from a breach. Participants consistently rank security scenarios as the most comfortable compared to other scenarios.

Some participants put a positive spin on the activity, stating it would help expose the threat of data breaches (6%). Others emphasized freedom of speech above all else (3%):

> P[1613]: "Concrete examples of how the hack has affected the lives of ordinary people makes the story more relatable."
> P[3878]: "Freedom of the press is essential in a well-functioning democracy. Reporters must come in some way to their publications."

5.2 Demographic variations

Table 4 provides a detailed summary of how the level of comfort, measured as a mean, compared across genders, age-groups and countries alongside the results from tests for statistical significance ($p = 0.05$). We corrected for multiple testing for Age-groups and Countries, using Bonferroni correction (adj.p = 0.008).

Differences across age groups: Across age groups, younger participants had a higher level of comfort with handling exposed data than older participants. The difference in comfort ranged between 0.14–0.68 for all scenarios, with the exception of the scenario of journalists reporting tax fraud via a public source. With research suggesting that younger adults are more likely to be victims of breaches [30], this may reflect a greater desire for solutions to a common experience.

Differences across country: Among the six countries we surveyed, participants from India and Germany indicated the highest level of comfort towards scenarios where entities purchased exposed data. Canada and Australia expressed the lowest levels of comfort (Figure 4). The difference in comfort for each scenario ranged between 0.53–1.10 per country. Despite absolute differences in comfort, participants universally rated security applications as more comfortable relative to other scenarios.

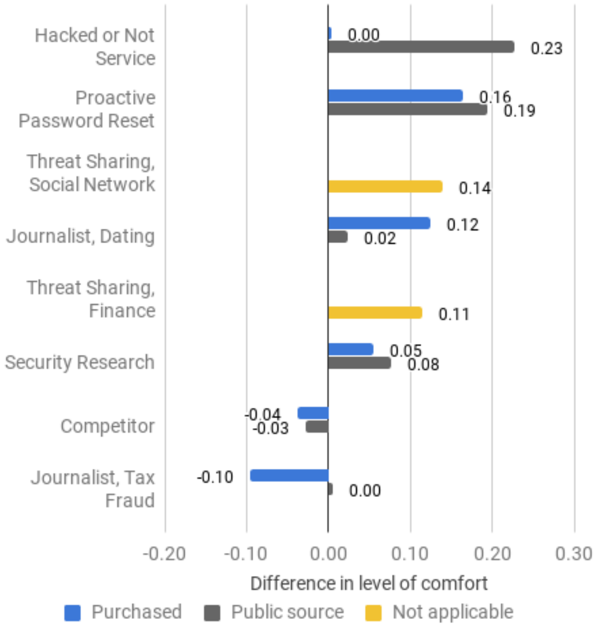

Figure 5: Difference in the comfort between prior victims of breaches and non-victims. Victims are consistently more likely to support security scenarios. All differences are statistically significant.

Differences based on breach experience: Of the respondents we surveyed, 24% self-reported having experienced a data breach. Among countries, 40% of US respondents reported experiencing a breach. Between genders, we did not see a significant difference between men (25%) and women (23%). Among age-groups we observed that more respondents between the ages 25-44 years reported experiencing a data breach (28%).

Figure 5 shows the difference in the level of comfort for prior victims of breaches and non-victims for each scenario. A positive value indicates victims were more comfortable than non-victims. Overall, victims reported consistently higher levels of comfort for all scenarios, including purchasing from criminals. The only exceptions were the competitor and tax fraud scenarios. The highest change in comfort related to a "Hacked or not" service. One explanation is that victims are more familiar with the harms that can result from a breach and are thus more supportive of security applications.

Differences across genders: Across genders, men had a higher level of comfort than women ($\mu = 2.57$ vs. $\mu = 2.41$). This was true in all scenarios, other than a "Hacked or Not" service where men had the same comfort level as women.

Table 4: Comparison of mean level of comfort across demographics: Gender, Age-group, and Country.

Metric	Gender		Age-Group						Country					
	Male	Female	18-24	25-34	35-44	45-54	55-64	65+	AU	CA	DE	IN	UK	US
			Research (Purchased)											
Mean	1.96	1.75	2.06	2.08	1.75	1.74	1.76	1.54	1.78	1.60	1.83	2.42	1.82	1.80
Test Statistic	$p < 0.001$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 64.8$						$p < 0.001$; KW test; $\chi^2 = 87.62$					
			Research (Public)											
Mean	2.09	1.83	2.21	2.19	1.93	1.88	1.76	1.53	1.84	1.80	1.92	2.44	1.89	1.96
Test Statistic	$p < 0.001$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 79.62$						$p < 0.001$; KW test; $\chi^2 = 55.61$					
			Hacked or Not (Purchased)											
Mean	2.51	2.52	2.68	2.82	2.50	2.37	2.25	2.11	2.47	2.24	2.63	2.99	2.53	2.32
Test Statistic	$p = 0.597$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 89.27$						$p < 0.001$; KW test; $\chi^2 = 80.77$					
			Hacked or Not (Public)											
Mean	2.48	2.38	2.40	2.68	2.41	2.36	2.30	2.05	2.33	2.24	2.39	2.73	2.40	2.47
Test Statistic	$p = 0.122$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 51.53$						$p < 0.001$; KW test; $\chi^2 = 23.71$					
			Threat Sharing, Finance											
Mean	3.00	2.88	3.05	3.10	2.99	2.93	2.76	2.49	2.73	2.70	3.23	3.27	2.78	2.99
Test Statistic	$p = 0.023$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 48.19$						$p < 0.001$; KW test; $\chi^2 = 80.47$					
			Threat Sharing, Social Network											
Mean	2.98	2.87	3.05	3.07	2.97	2.89	2.69	2.67	2.69	2.78	3.22	3.37	2.60	2.91
Test Statistic	$p = 0.035$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 32.69$						$p < 0.001$; KW test; $\chi^2 = 114.13$					
			Proactive Password Reset (Purchased)											
Mean	2.91	2.80	3.13	3.04	2.70	2.83	2.69	2.56	2.66	2.66	3.03	3.19	2.79	2.83
Test Statistic	$p < 0.001$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 51.58$						$p < 0.001$; KW test;$\chi^2 = 53.97$					
			Proactive Password Reset (Public)											
Mean	2.83	2.63	2.92	2.82	2.66	2.80	2.61	2.40	2.60	2.61	2.89	2.77	2.57	2.95
Test Statistic	$p = 0.122$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 29.7$						$p < 0.001$; KW test; $\chi^2 = 29.5$					
			Journalist, Tax Fraud (Purchased)											
Mean	2.76	2.52	2.69	2.84	2.59	2.58	2.43	2.55	2.47	2.24	2.63	2.99	2.53	2.32
Test Statistic	$p < 0.001$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 36.05$						$p = 0.001$; KW test; $\chi^2 = 112.6$					
			Journalist, Tax Fraud (Public)											
Mean	2.86	2.68	2.78	2.88	2.84	2.73	2.55	2.78	2.33	2.30	2.39	2.73	2.40	2.47
Test Statistic	$p = 0.001$; Mann-U		$p = 0.003$; KW test; $\chi^2 = 18.22$						$p = 0.001$; KW test; $\chi^2 = 21.6$					
			Journalist, Dating Site (Purchased)											
Mean	1.97	1.69	1.98	2.12	1.86	1.62	1.60	1.51	1.72	1.56	1.71	2.66	1.91	1.60
Test Statistic	$p < 0.001$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 87.60$						$p < 0.001$; KW test; $\chi^2 = 214$					
			Journalist, Dating Site (Public)											
Mean	2.08	1.81	2.19	2.16	2.01	1.78	1.72	1.57	1.88	1.78	1.81	2.60	2.01	1.74
Test Statistic	$p < 0.001$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 97.2$						$p < 0.001$; KW test; $\chi^2 = 127.8$					
			Competitor (Purchased)											
Mean	2.06	1.94	2.28	2.23	2.05	1.87	1.72	1.68	1.99	1.74	2.08	2.66	2.01	1.64
Test Statistic	$p = 0.008$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 92.30$						$p < 0.001$; KW test; $\chi^2 = 160.20$					
			Competitor (Public)											
Mean	2.06	1.98	2.30	2.24	2.00	1.92	1.77	1.69	1.90	1.80	2.08	2.53	2.04	1.85
Test Statistic	$p = 0.185$; Mann-U		$p < 0.001$; KW test; $\chi^2 = 79.56$						$p < 0.001$; KW test; $\chi^2 = 85.04$					

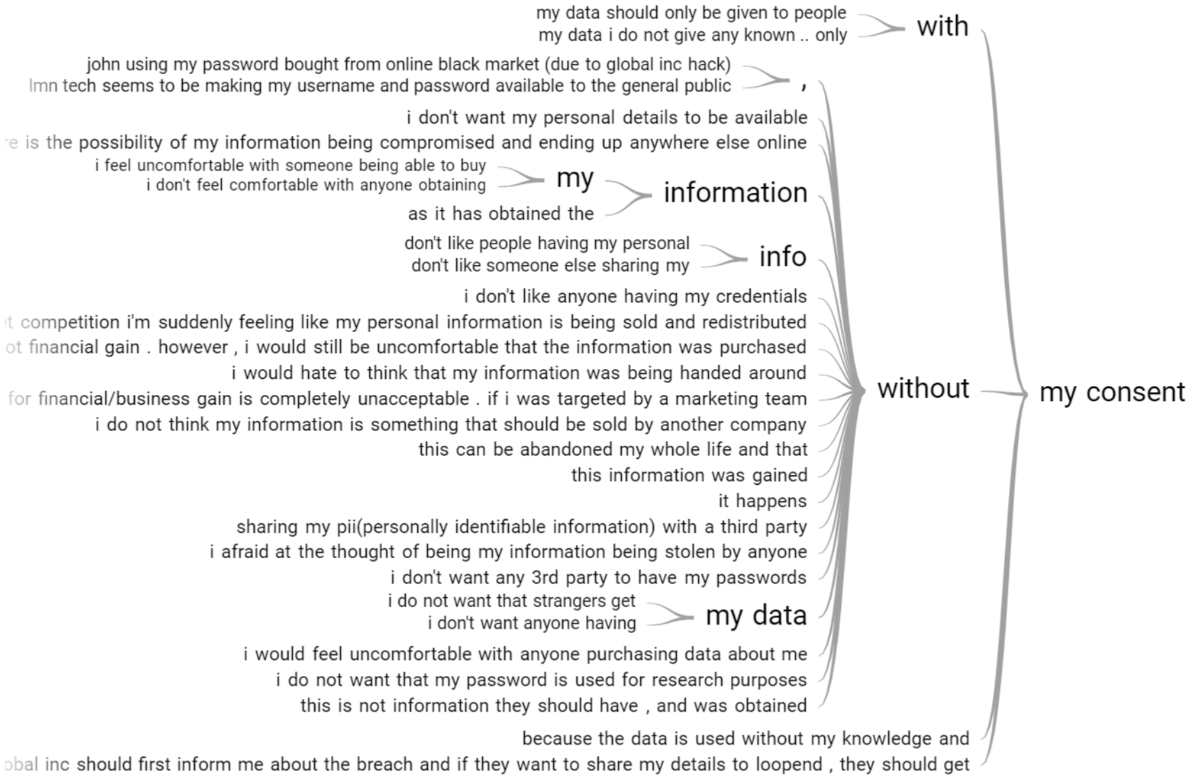
my data should only be given to people
my data i do not give any known .. only
with
john using my password bought from online black market (due to global inc hack)
lmn tech seems to be making my username and password available to the general public
,
i don't want my personal details to be available
e is the possibility of my information being compromised and ending up anywhere else online
i feel uncomfortable with someone being able to buy
i don't feel comfortable with anyone obtaining
my
information
as it has obtained the
don't like people having my personal
don't like someone else sharing my
info
i don't like anyone having my credentials
t competition i'm suddenly feeling like my personal information is being sold and redistributed
ot financial gain . however , i would still be uncomfortable that the information was purchased
i would hate to think that my information was being handed around
for financial/business gain is completely unacceptable . if i was targeted by a marketing team
i do not think my information is something that should be sold by another company
this can be abandoned my whole life and that
this information was gained
it happens
sharing my pii(personally identifiable information) with a third party
i afraid at the thought of being my information being stolen by anyone
i don't want any 3rd party to have my passwords
i do not want that strangers get
i don't want anyone having
my data
i would feel uncomfortable with anyone purchasing data about me
i do not want that my password is used for research purposes
this is not information they should have , and was obtained
without
my consent
because the data is used without my knowledge and
obal inc should first inform me about the breach and if they want to share my details to loopend , they should get

Figure 6: Tree map of open-ended responses from second study suffixed with the term "my consent".

6. DISCUSSION

User expectations after a breach: Over 40% of participants from the United States and 25% across the United Kingdom, Germany, Australia, Canada, and India reported being former victims of a breach. Our results indicate participants have strong expectations of being notified when their data is exposed. In the case of credentials, participants expressed this allows them to take precautionary measures such as resetting their password for all their affected accounts. Participants also emphasized that transparency remained important, even when a notification alone was viewed as an insufficient response. Other proactive measures included force resetting passwords or otherwise hardening security around accounts, such as with two-factor authentication. Outside of a company's responsibilities to users, participants also strongly favored interacting with government authorities as a means of holding criminals accountable—as well as the breached company.

Community responses to a breach: Digital identity is not an island; a breach at one company may allow criminals to access other resources due to reused passwords or recovery questions. Our results indicate that participants are supportive, or at least neutral towards, emerging security strategies by identity providers. Between 37–40% of participants expressed comfort towards threat sharing between identity providers as well as proactively resetting passwords exposed by third-party breaches. Another 20–23% of participants expressed a neutral opinion towards these activities. Lingering concerns for participants fell into two categories: skepticism any such actions would help secure their accounts, and strong expectations about privacy and ethical behavior—even when companies can acquire data without engaging with black markets.

Conversely, participants expressed a greater degree of concern towards "Hacked or not" services, and even more concern for research based on data exposed from breaches. Consent was a consistent theme, with common strains of feedback shown in Figure 6. Here, participants valued their privacy over abstract security benefits, though victims of prior breaches reported higher levels of comfort. A key takeaway is that security professionals and researchers need to articulate how any services or investigations can provide a direct benefit to the victims of breaches given the sensitive nature of the data involved.

Responsibly handling exposed data: Our study provides a perspective of user expectations and concerns with respecting breached data. However, before establishing a line in the sand for best practices of responsibly handling exposed data, it is also vital to consider the views of journalists, security experts, and researchers. We leave building such a broad perspective to future work.

7. CONCLUSION

In this work, we presented the results of two surveys that gauged user comprehension, expectations, and concerns with both responding to breaches and how one handles exposed data. Our results indicate that data breaches are a highly topical issue in the minds of participants. As such, participants have clear expectations for remediation steps: Breached companies need to be transparent and notify victims; proactively reset passwords and lock down accounts from further damage; and engage with law enforcement to identify the criminals involved.

Zooming out to the wider community, our results show that participants are supportive of emerging security practices including threat sharing between companies in the event of a breach, as well as proactively resetting reused passwords found in password dumps. Other use cases that have less direct or tangible benefits to users, such as research or "hacked or not" services, were viewed less favorably due to overriding privacy concerns. Our findings also help to inform a broader discussion within the community of how to responsibly handle exposed data while respecting concerns around privacy and consent.

8. REFERENCES

[1] L. Ablon, P. Heaton, D. C. Lavery, and S. Romanosky. Consumer attitudes toward data breach notifications and loss of personal information. In *Proceedings of the Workshop on the Economics of Information Security*, 2016.

[2] B. Benko, E. Bursztein, T. Pietraszek, and M. Risher. Cleaning up after password dumps. `https://security.googleblog.com/2014/09/cleaning-up-after-password-dumps.html`, 2014.

[3] Council of the European Union. Notification of a personal data breach to the supervisory authority. `https://gdpr-info.eu/art-33-gdpr/`, 2017.

[4] L. F. Cranor. Giving notice: why privacy policies and security breach notifications aren't enough. *IEEE Communications Magazine*, 43(8):18–19, 2005.

[5] A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. Wang. The tangled web of password reuse. In *Proceedings of the Network and Distributed System Security Symposium*, 2014.

[6] X. D. C. De Carnavalet, M. Mannan, et al. From very weak to very strong: Analyzing password-strength meters. In *Proceedings of the Network and Distributed System Security Symposium*, 2014.

[7] J. Franklin, A. Perrig, V. Paxson, and S. Savage. An inquiry into the nature and causes of the wealth of internet miscreants. In *Proceedings of the Conference on Computer and Communications Security*, 2007.

[8] J. Garside. Paradise papers leak reveals secrets of the world elite's hidden wealth. `https://www.theguardian.com/news/2017/nov/05/paradise-papers-leak-reveals-secrets-of-world-elites-hidden-wealth`, 2017.

[9] A. Greenberg. Hack brief: Uber paid off hackers to hide a 57-million user data breach. `https://www.wired.com/story/uber-paid-off-hackers-to-hide-a-57-million-user-data-breach/`, 2017.

[10] S. Gressin. The equifax data breach: What to do. `https://www.consumer.ftc.gov/blog/2017/09/equifax-data-breach-what-do`, 2017.

[11] L. Harding. What are the panama papers? a guide to history's biggest data leak. `https://www.theguardian.com/news/2016/apr/03/what-you-need-to-know-about-the-panama-papers`, 2016.

[12] M. Hiltzik. Did TransUnion Increase Cost Of Credit Monitoring In Wake Of Equifax Breach? `http://beta.latimes.com/business/hiltzik/la-fi-hiltzik-lifelock-equifax-20170918-story.html`, 2017. [Online; accessed 20-January-2018].

[13] T. Hunt. The impact of "Have I been pwned" on the data breach marketplace. `https://www.troyhunt.com/the-impact-of-have-i-been-pwned-on-data/`, 2016.

[14] R. Janakiraman, J. H. Lim, and R. Rishika. The effect of data breach announcement on customer behavior: Evidence from a multichannel retailer. *Journal of Marketing*, 2017.

[15] M. G. Keith and P. D. Harms. Is mechanical turk the answer to our sampling woes? *Industrial and Organizational Psychology*, 9(1):162–167, 2016.

[16] S. Larson. Every single yahoo account was hacked - 3 billion in all. `http://money.cnn.com/2017/10/03/technology/business/yahoo-breach-3-billion-accounts/index.html`, 2017.

[17] S. Larson. Uber's massive hack: What we know. `http://money.cnn.com/2017/11/22/technology/uber-hack-consequences-cover-up/index.html`, 2017.

[18] W. Melicher, B. Ur, S. M. Segreti, S. Komanduri, L. Bauer, N. Christin, and L. F. Cranor. Fast, lean, and accurate: Modeling password guessability using neural networks. In *Proceedings of the USENIX Security Symposium*, 2016.

[19] R. M. Peters. So you've been notified, now what: The problem with current data-breach notification laws. *Ariz. L. Rev.*, 56:1171, 2014.

[20] R. Price. A site that tracked massive hacks has disappeared after being allegedly raided by the cops. `http://www.businessinsider.com/hack-tracking-site-leakedsource-disappears-allegedly-raided-law-enforcement-2017-1`, 2017.

[21] L. Rainie, S. Kiesler, R. Kang, M. Madden, M. Duggan, S. Brown, and L. Dabbish. Anonymity, privacy, and security online. *Pew Research Center*, 2013.

[22] P. M. Schwartz and E. J. Janger. Notification of data security breaches. *Michigan Law Review*, pages 913–984, 2007.

[23] R. Shay, I. Ion, R. W. Reeder, and S. Consolvo. My religious aunt asked why i was trying to sell her viagra: experiences with account hijacking. In *Proceedings of the Conference on Human Factors in Computing Systems*, 2014.

[24] D. R. Thomas, S. Pastrana, A. Hutchings, R. Clayton, and A. R. Beresford. Ethical issues in research using datasets of illicit origin. In *Proceedings of the Internet Measurement Conference*, 2017.

[25] K. Thomas, F. Li, A. Zand, J. Barrett, J. Ranieri, L. Invernizzi, Y. Markov, O. Comanescu, V. Eranti, A. Moscicki, et al. Data breaches, phishing, or malware?: Understanding the risks of stolen credentials. In *Proceedings of the Conference on Computer and Communications Security*, 2017.

[26] K. Thomas and A. Moscicki. New research:

Understanding the root cause of account takeover. `https://security.googleblog.com/2017/11/new-research-understanding-root-cause.html`, 2017.

[27] A. Tversky and D. Kahneman. Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232, 1973.

[28] D. L. Wheeler. zxcvbn: Low-budget password strength estimation. In *Proceedings of the USENIX Security Symposium*, 2016.

[29] V. Woollaston. Facebook and netflix reset passwords after data breaches. `http://www.wired.co.uk/article/facebook-netflix-password-reset`, 2016.

[30] K. Zickuhr. Generations and their gadgets. *Pew Internet & American Life Project*, 20, 2011.

Appendix

Study 1: Questionnaire:

Which of the following according to you is a data breach?
o Public exposure of usernames and passwords of millions of users of an online system *[screened in]*
o Web page that is unable to load due to too much data on the page *[screened out]*
o Using large sets of data to aid robots to solve a problem that humans cannot solve *[screened out]*

According to you, what is the most important harm that can happen due to a data breach?
o Spam being sent out from your account / Receiving Spam
o Your computer will be infected by a virus
o Identity theft
o Monetary Loss
o No harm
o Loss of access to personal information
o Leak of personal information
o Your phone will be monitored by hackers
o Other, Please specify

In response to being hacked, which of the following actions should a company take to protect your account? (Choose all that Apply)
o Upgrade your web browser
o Change your username
o Issue a refund
o Send you an immediate notification about the breach
o Enable two-factor authentication
o Provide credit monitoring
o Pay users a consolation bonus for breaking their trust
o Reset your password
o Give you a new account
o Company buys you a new computer
o Other, Please specify

What is the shape of a red ball?
o Red o Blue o Square o Round

How comfortable would you be with a company taking the following actions after the company had a data breach?

ACTION 1: Company that was hacked and experienced the data breach notifies you that your password was stolen
Not at all Comfortable | 1 2 3 4 5 | Very Comfortable
Please explain your rating.
open ended response

ACTION 2: Company that was hacked and experienced the data breach notifies the press about the incident
Not at all Comfortable | 1 2 3 4 5 | Very Comfortable
Please explain your rating.
open ended response

ACTION 3: Company that was hacked and experienced the data breach reports the incident to the government
Not at all Comfortable | 1 2 3 4 5 | Very Comfortable
Please explain your rating
open ended response

ACTION 4: Company that was hacked and experienced the data breach buys a copy of stolen usernames and passwords from the hacker to know what was exposed
Not at all Comfortable | 1 2 3 4 5 | Very Comfortable
Please explain your rating
open ended response

ACTION 5: Company that was hacked and experienced the data breach resets your password

Not at all Comfortable | 1 2 3 4 5 | Very Comfortable
Please explain your rating
open ended response

ACTION 6: Company that was hacked and experienced the data breach shares a copy of all stolen usernames and passwords with other companies to protect your other accounts where you may have reused your username/password
Not at all Comfortable | 1 2 3 4 5 | Very Comfortable
Please explain your rating
open ended response

Have you ever been a victim of data breach?
o Yes o No o Don't know

What is your gender?
o Female o Male o Transgender o I prefer not to answer o Other:

What is your age-group?
o 18-24 years old o 25-34 o 35-44 o 45-54 o 55-64 o 65 or older o I prefer not to answer

Which country do you live in?
drop-down with list of countries

What is the highest degree or level of school that you have completed?
o Professional doctorate (for example, MD, JD, DDS, DVM, LLB) o Doctoral degree (for example, PhD, EdD) o Masters degree (for example, MS, MBA, MEng, MA, MEd, MSW) o Bachelors degree (for example, BS, BA) o Associates degree (for example, AS, AA) o Some college, no degree o Technical/Trade school o Regular high school diploma o GED or alternative credential o Some high school o I prefer not to answer o Other, Please Specify

Which of the following describes your current employment status?
o Employed full-time o Employed part-time o Self-employed o Care-provider o Homemaker o Retired o Student - Undergraduate o Student - Masters o Student - Doctoral o Looking for work / Unemployed o Other, Please specify

What is the color of a red ball?
o Red o Blue o Square o Round

Study 2: Questionnaire:

WELCOME
The goal of this study is to get your feedback on a few scenarios that are detailed in the next set of screens. You will be presented with two hypothetical scenarios. For each of these scenarios, you will be asked to rate your level of comfort.

Your contribution to this study would be of great value to us as we are always looking for ways to improve the experience of internet users like yourself.

The study will take approximately 5-10 minutes to complete.

When you are ready to proceed, please click ■.

Introduction: Global Inc is a leading online social network that provides services such as Email, Chat, Blogs, and Profile pages with over 100 million users using its services everyday. Global Inc just suffered a data breach resulting in hackers gaining access to the data of every user, including their username and password. The hackers responsible for the attack are now selling the stolen data online for a price.

(Randomly show each respondent 2 out of the 8 questions below)

Security Research
John is a researcher from the Southern University investigating online security. John‘s research focuses on identifying how internet users select passwords, including the most commonly selected passwords. When John hears that he can buy millions of passwords exposed by the Global Inc hack via the online black market, he decides to buy a copy to use for his research.

Q1: Imagine you are one of the users of Global Inc who was affected by the breach. Rate your level of comfort with John buying the hacked data which may contain your credentials too?
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
– open ended response –

Q3:If the hacked data was publicly available on the internet for free download, rate your level of comfort with John downloading and using the hacked data for his research.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

---

Hacked or not service
LMN Tech is an online security service provider. They provide a paid service where anyone can look up their username to see whether it was exposed as part of a major data breach. Their service relies on archives of data collected from a variety of data breach incidents. LMN Tech hears about Global Inc's recent data breach and is planning to buy the hacked data from the hackers to be able to add it to their huge archive of breached datasets.

Q1: Imagine you are one of the users of Global Inc. Rate your level of comfort with LMN Tech's purchase.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Q3: If the hacked data was publicly available on the internet for free download, what will be your level of comfort with LMN Tech downloading and using them to support their service?
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

---

Threat sharing: Finance
PayPool is an industry leading online payments provider that supports making online purchases. Global Inc knows that its users often logged into other online services such as PayPool, with their Global Inc email address. The Security team at Global Inc suspect that the hackers will use the hacked data to further hack accounts of PayPool users to commit financial fraud. Global Inc believes that sharing a list of hacked email ids that are definitely linked to PayPool accounts will enable PayPool to guard those user's account on PayPool through proactive password resets.

Q1: Imagine you are one of the users of Global Inc and you use your Global Inc email to login to PayPool for online purchases, rate your level of comfort with the security team's plan.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Threat Sharing: Social
LoopedIn is a popular professional networking site where job seekers post their CVs and employers post jobs. Global Inc knows that its users often logged into services such as LoopedIn, with their Global Inc email address. The Security team at Global Inc suspect that the hackers will use the hacked data to further hack accounts of LoopedIn users and may leak their job search information. Global Inc believes that sharing a list of hacked email ids that are definitely linked to LoopedIn accounts will enable LoopedIn guard those user's account on LoopedIn through proactive password resets.

Q1: Imagine you are one of the users of Global Inc and you use your Global Inc email to login to LoopedIn for professional networking. Rate you level of comfort with the security team's plan.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Proactive Password Reset
Doodle is a large internet company with billions of users. Many of Doodle's users reuse their passwords on third party services. In an attempt to proactively protect users from someone breaking into their account, Doodle regularly buys hacked datasets that are sold on the black market to scan and re-secure accounts of users who have had their password exposed on other third party services.

Q1: Imagine you are a Doodle user. Rate your level of comfort with Doodle‘s proactive security measure.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Q3: If the hacked data was publicly available on the internet for free download, rate your level of comfort with Doodle downloading and using them as a proactive security measure?
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

---

Journalist, tax fraud
Jerry is a prominent journalist working at That‘s Correct Media and Publishing company. After Global Inc‘s hack Jerry buys the hacked data via the online blackmarket and gets access to personal emails of some of Global Inc‘s users. These emails reveal a major public scam involving several public officials using offshore financial centers to avoid taxes. Jerry publishes a news article to disclose the tax avoidance scam, using the hacked email as proof.

Q1: Imagine you are one of the users of Global Inc who was affected by the breach. Rate your level of comfort with John buying the hacked data which may contain your credentials too?
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Q3: If the hacked data was publicly available on the internet for free download, rate your level of comfort with John downloading and using the hacked data for his research.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

---

Journalist, dating site
Mark a journalist at TownNews Today learns about a recent data breach of a dating site GoDate.com. He decides to purchase the hacked data to look up names of people from his town and publish information about their private dating profiles. Mark feels that the profiles would make up interesting subject matter for his articles and is planning to publish some articles based on these profiles.

Q1: Imagine you are reading JerryâĂŹs article, rate your level of comfort with JerryâĂŹs source of proof.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Q3: If the hacked data was publicly available on the internet for free download, rate your level of comfort with Jerry downloading and using it as a source of proof for his article.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

---

Competitor
Moon Inc is a competitor of Global Inc. After hearing about Global Inc‘s data breach, the marketing team at Moon Inc plans to buy the hacked data via the online blackmarket to learn about Global Inc‘s users who were hacked. The marketing team plans to target these Global Inc users with an offer to switch services.

Q1: Imagine you are one of the users of Global Inc who was affected by the breach. Rate your level of comfort with being approached by Moon Inc offering you to switch to their service.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

Q2: Please explain your rating.
open ended response

Q3: If the hacked data was publicly available on the internet for free download, rate your level of comfort with Moon Inc downloading and using them to approach hacked Global Inc users to switch to their service.
Extremely Uncomfortable | Somewhat Uncomfortable | Neither comfortable nor Uncomfortable | Somewhat Comfortable | Extremely Comfortable

---

Please specify your gender: o Male o Female o Other o Prefer not to say

Please select your age group: o 18-24 o 25-34 o 35-44 o 45-54 o 55-64 o 65+ o Prefer not to say

What is the highest degree or level of school that you have completed? o Less than high school o High school graduate (includes equivalency) o Some college, no degree o Associate‘s degree o Bachelor‘s degree o Master‘s degree o Ph.D. o Other, Please specify

Please specify your current employment status o Employed Full-time o Employed Part-time o Self-employed o Care-provider o Homemaker o Retired o Student - Undergraduate o Student - Masters o Student - Doctoral o Looking for work / Unemployed o Other, Please Specify

Have you ever been a victim of data breach? o Definitely yes o Probably yes o Might or might not o Probably not o Definitely not

---