Anatomy of a High-Profile Data Breach: Dissecting the Aftermath of a Crypto-Wallet Case

Svetlana Abramova and Rainer Böhme, *Universität Innsbruck*

https://www.usenix.org/conference/usenixsecurity23/presentation/abramova

This paper is included in the Proceedings of the 32nd USENIX Security Symposium.

August 9–11, 2023 • Anaheim, CA, USA

978-1-939133-37-3

Open access to the Proceedings of the 32nd USENIX Security Symposium is sponsored by USENIX.

Anatomy of a High-Profile Data Breach: Dissecting the Aftermath of a Crypto-Wallet Case

Svetlana Abramova
Universität Innsbruck

Rainer Böhme
Universität Innsbruck

Abstract

Media reports show an alarming increase of data breaches at providers of cybersecurity products and services. Since the exposed records may reveal security-relevant data, such incidents cause undue burden and create the risk of re-victimization to individuals whose personal data gets exposed. In pursuit of examining a broad spectrum of the downstream effects on victims, we surveyed 104 persons who purchased specialized devices for the secure storage of crypto-assets and later fell victim to a breach of customer data. Our case study reveals common nuisances (i. e., spam, scams, phishing e-mails) as well as previously unseen attack vectors (e. g., involving tampered devices), which are possibly tied to the breach. A few victims report losses of digital assets as a form of the harm. We find that our participants exhibit heightened safety concerns, appear skeptical about litigation efforts, and demonstrate the ability to differentiate between the quality of the security product and the circumstances of the breach. We derive implications for the cybersecurity industry at large, and point out methodological challenges in data breach research.

1 Introduction

Data breaches (i. e., the leakage or disclosure of sensitive, confidential, or otherwise protected information to unauthorized parties [58]) continue to plague businesses and customers around the globe. As these unwanted events become commonplace in some industries [70], it is worrying that more and more cybersecurity vendors and service providers appear in the headlines of reports announcing new leaks. Last year's notorious examples include Microsoft, with over two terabytes of disclosed business customer data [16], and the LastPass password manager, which suffered from an unauthorized access to its backup of customer vault data [40].

Breach cases of this kind are noteworthy for two reasons. First, they demonstrate that even the cybersecurity industry cannot effectively prevent breaches and thus must adopt a *'not if, but when'* security mindset. Security providers must prepare for potential breach events, so that their crisis communication can respond in a fully transparent, timely and instructive manner. Second, high-profile breaches may place a particularly high burden on the individuals whose personal data gets exposed, because these events can serve as stepping stones for future attacks. Customer data leaked by security providers may inadvertently supply threat actors with valuable facts about people and their security actions. As a result, victims become subject to increased and more targeted attacks, which force them to maintain high security vigilance.

To date, data breach research has not studied this kind of incidents and wider implications thereof. In fact, there is still a lack of knowledge and methods to estimate the downstream effects of conventional (i. e., lower-profile) breaches on customers and their post-breach behavior. Only recently, scholars began to study victims' reactions, such as attitude changes concerning risk and trust [7, 10, 72], emotional [8, 34, 52] and behavioral responses [12, 44, 50, 69, 72], monetary and psychological harms [37], and perceptions of litigation actions [26]. However, many studies either narrow their scope of analysis to a single type of an attack or behavior, or discuss hypothetical scenarios, for which customer-facing consequences are overly general or unfounded. Effects may take many forms and, therefore, neither approach yields an in-depth picture of the aftermath of a breach.

To address these research gaps, we present a case study of the data breach affecting Ledger customers, which happened in July 2020 [42]. This incident meets our criteria of being 'specific', because it affected one of the leading manufacturers of hardware wallet devices intended for secure offline storage of crypto-assets. Personal data (including names, postal and e-mail addresses, phone numbers) of a subset of the company's global customer base got publicly exposed, while the purchased security devices were not compromised.

We are interested in examining the effects of this particular event for several reasons. Crypto-asset owners are attractive targets for both amateur and professional fraudsters looking for ways to monetize the leaked data. Compared to the average internet user, crypto-asset owners who proactively pre-

fer specialized storage devices demonstrate higher security awareness and caution [2, 45]. Hence, they might be less susceptible to successful crimes and serious damages. Finally, security-aware victims of this incident may serve as a reliable and knowledgeable source of information when it comes to reporting about the aftermath of the breach.

Our study focuses on the three key research questions covering a broad spectrum of the potential effects of a data breach. First, we aim to elicit *which harms*, be they financial, emotional distress, or invasion of privacy, have been experienced, and in *what ways* this incident has impacted the affected persons (RQ1). We are guided by insights from prior research suggesting that leaked sensitive data results in an increased risk of online identity theft [59], account compromise [68], and phishing [55]. Furthermore, cryptocurrency markets themselves are deemed to be prone to fraud, thefts, and hacking [11, 30]. This suggests potentially heightened interest in the leaked data from perpetrators specializing in this domain. By collecting empirical evidence of the experienced threats and harms, we set to build a knowledge base, sought-after also in policy debates on data breach legislation [37, 58].

Concerning our second research question: in order to mitigate harm, victims are advised to adopt protective measures, such as changing passwords [25], adding two-factor authentication [19], and managing multiple e-mail aliases for account registration [50]. However, these advisable actions are rarely observable in practice, as past studies repeatedly show [50, 73]. Drawing on these results, we seek to investigate which responses and changes in individuals' security behaviors this breach has triggered (RQ2). The fact that crypto-assets (without consumer protection and little chance of legal recourse) are at stake, adds relevance to this research question.

Third, the analysis of a breach affecting a security provider would be incomplete without studying far-reaching impacts on customer attitudes, trust, and loyalty. To this end, we sought inspiration in the marketing literature [15, 28], which suggests separating an individual's corporate and product associations. With RQ3, we intend to explore to what extent the breach may have affected consumer attitudes to the company itself as compared with the impact on its security product.

To address these questions, we recruited 104 participants using e-mail addresses from the leaked dataset. This recruitment method makes us confident that many of our respondents are in fact victims of the breach. By implementing a series of risk mitigation measures outlined below in the paper, our study was approved by the Institutional Review Board (IRB).

The rest of this paper is organized as follows. We first present our case, including a timeline of relevant events (Section 2). Then, we review the state-of-the-art literature on data breaches (Section 3). Next, we thoroughly explain the methodology of our study (Section 4). We present the empirical results broken down by research question in Section 5. Finally, we discuss the implications of our results for practitioners and researchers, recall limitations, and conclude (Sections 6, 7).

2 The Case

Ledger,[1] headquartered in France, is a vendor of hardware crypto-wallets. Their products are portable physical devices purposefully designed for *offline storage* of cryptographic keys to crypto-assets. Hardware wallets are widely advertised as the most secure option for managing digital assets [6]. However, upon device compromise—or in this case a long recovery phrase being disclosed—an attacker can get access to all crypto-assets secured by the device and transfer them irrevocably to accounts under his control.

In this market, Ledger competes with a handful of vendors selling similar security products. The vendor offers its customers a choice of two products, with basic or advanced functions, in the price range between € 80 and € 150. Given their technical features and purchase price, hardware crypto-wallets are often recommended for long-term storage, large crypto-asset portfolios, or users interested in additional security [2, 53].

Figure 1 shows a timeline of the key events of the breach and our study. Our reconstruction of the case is mainly sourced from information published by the companies involved [43, 63], as independent and objective public information is rare. Ledger partnered with Shopify, a popular e-commerce platform, to manage its online sales. In spring 2020, rogue employees at Shopify reportedly gained unauthorized access to customer records of up to 200 merchants [63], including Ledger [43]. In July 2020, a bug bounty hunter notified Ledger of a breach of their e-commerce and marketing database, involving a third-party API key. Ledger patched this problem and initiated internal investigations to estimate risks and protect its customers [43]. A subset of Ledger's marketing database, alleged to be the contents stolen earlier, was dumped in plain text on Raidforum on December 20, 2020. This public database contained personal information (first and last names, postal and e-mail addresses, phone numbers) of approximately 272 000 customers [43].

We got notified of this leak via social media and downloaded the dump from the Intelligence X search engine and data archive[2] on December 21, 2020. Given our purpose of use, we erased all personal data contained in the dump except for e-mail addresses of customers. We started the study by seeking legal advice form our university's data protection officer. For this, we outlined a document explaining our study, its methods as well as potential issues of concern. The design of a questionnaire started in June 2021, followed by several rounds of peer revision and pretests in October 2021. In parallel, we applied for ethical approval from the IRB in summer 2021. The recruitment of survey participants started on November 9, 2021 and continued until February 22, 2022. The data analysis and reporting of results were started and finalized in the summer and winter of 2022, respectively.

[1] https://www.ledger.com/
[2] https://intelx.io/

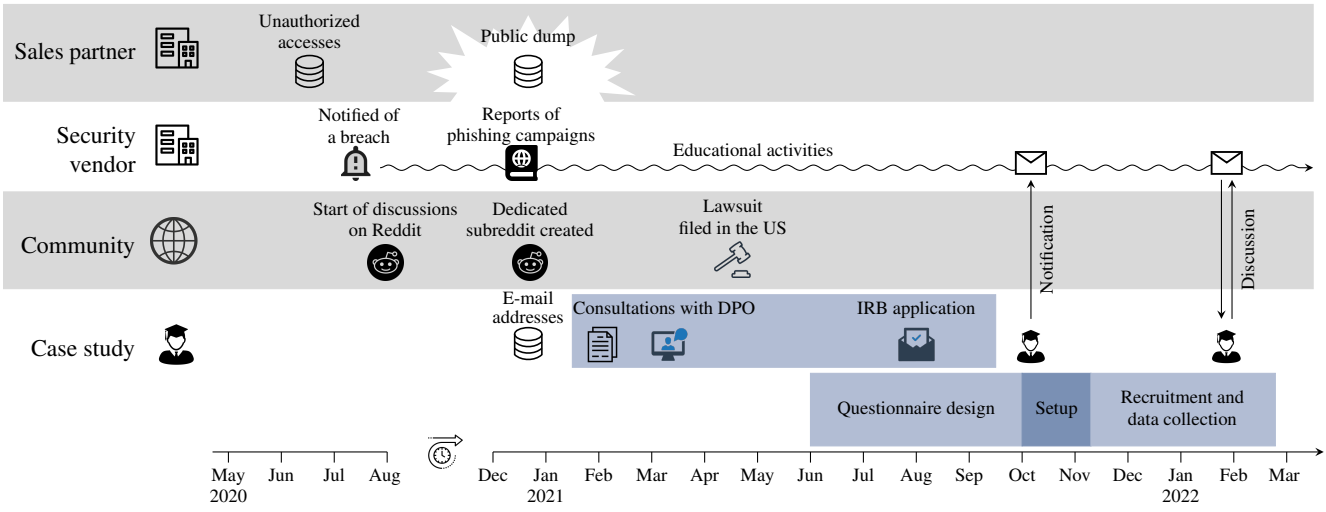


Figure 1: Timeline of key events in our case study

In the beginning of October 2021, before the start of the field phase, we notified Ledger about our study using the only e-mail address the security vendor provided on its website. The company's representatives established contact and discussed the study with us in January 2022, after they received reports about our running survey from the customers. We also shared a pre-print of the conditionally accepted paper in May 2023, allowing the company to prepare for potential customer reactions. Feedback received from the company at this stage, after cross-checking against independent sources, has helped us to clarify some facts. We did not interact with Shopify.

The broader community of crypto-asset owners reacted to the breach on Reddit almost immediately after the vendor announced this fact on its website and in newsletter e-mails. When the dump got exposed to the public, the community created a subreddit forum dedicated to this incident. Also, a US-based law firm reportedly filed a class action complaint against the vendor in April 2021 [46]. The vendor has made efforts to support its customers by expanding their educational materials on security practices, shutting down phishing websites, and tracking scammers' transactions [43].

Note that we decided against using the community term 'Ledger breach' in this paper. We are interested in studying the effects of this breach, but do not want to discredit any of the involved companies, which in this case reportedly fell victim to cybercriminals themselves.

3 Related Work

Data breach research spans various disciplines, topics and methods. Historically, the most popular topics of study were breached organizations, their preventive and response measures [5, 29], costs [57, 62], stock market reactions [71], and workforce decisions [9]. Schlackl et al. [60] present a systematic review of more than 120 articles published in 1990–2021 on a range of the factors preceding and following a breach. One key takeaway of this work—underpinning our own view—is that this emerging research field lacks empirical studies measuring the effects of breaches on individuals, chiefly how much real harm was incurred.

We conducted our own systematic search of potentially relevant articles in the ACM, IEEE, and information systems digital libraries, using a combination of key terms searched in titles or abstracts. For greater coverage, we also included a few academic venues which may solicit works related to data breaches. In total, we identified 92 papers of interest, out of which 20 can be found in the sample reviewed in [60]. Our own search reconfirms the ongoing evolution and growth of this research area. We summarize the most relevant findings here and refer the reader to [60] for a more detailed overview.

Topics of growing interest in the recent years include increased concerns [35], perceptions [27, 31, 39, 50, 72], damages of different nature [8], post-breach reactions and behaviors of victims [12, 50, 69]. Methodologically, some studies survey general population panels to sample respondents with self-reported or verified breach experience [1, 50]. Other works use hypothetical scenarios to frame potential effects of a breach and inquire individuals' behavioral reactions to it, although the respondents may never have experienced a breach event [35]. A number of studies, like us, adopt a case-based approach to examining consumers' perspective on breaches. Selected examples include Equifax [72], the US Office of Personnel Management [8], Ashley Madison [21], Home Depot [64], and Canva [51]. However, we could not find any study examining a breach affecting a security vendor. To the best of our knowledge, the present work is novel in this regard.

The results of prior studies indicate that individuals affected by a breach not always act on notifications or heed enclosed security advice [33, 72, 73]. For example, only half

of the population sample surveyed in [1] reported having changed passwords or personal identification numbers after being notified about an incident. Recent data on behavioral reactions to breaches look slightly better, with 88% of the breach victims reporting to have changed their passwords in [50]. However, the scope of this study spans a multitude of breaches of varying size and magnitude, which adds noise to the signal. Turning to explanatory factors, the general reluctance to make mitigation efforts can be explained partly by high perceived costs associated with protective actions [72], as well as low levels of concern for specific accounts or breached data items [1, 50]. Against this backdrop, our work adds to the expanding body of knowledge by looking at a case involving victims with heightened security awareness [2, 45] and concerns about their digital assets. Moreover, our study stands out for its analysis of more targeted and previously unseen threats.

4 Method

Surveying a sample of crypto-asset owners is known to be a challenging task in research [2]. The main obstacles include a globally disperse and heterogeneous population, the use of digital pseudonyms as user identifiers, privacy concerns, and the ensuing non-response bias. While commercial crowdsourcing recruitment platforms extended their pre-screening filters with cryptocurrency-related fields and are being tested by scholars [2, 48], they still lack reliable and non-intrusive measures for the detection of fake self-reports of crypto-asset ownership. Against this backdrop, the leaked e-mail addresses offered us a unique research opportunity to recruit victims of the high-profile breach who are very likely to be genuine crypto-asset owners.

With ethical approval from our IRB and data protection officer, we sent out *one-off*, *non-personalized* e-mail invitations to a randomly chosen subset of the leaked addresses. Given a contentious nature of our recruitment approach and being guided by the principle of research transparency, we devote the next subsection to the discussion of legal and ethical aspects of our study. As we aim to demonstrate full transparency on our course of action, we specify technical and non-technical measures taken to minimize (additional) harm to the concerned parties. We then describe our recruitment and survey design processes in more detail (Sections 4.2 and 4.3), comment on the data quality (Section 4.4), and describe the socio-demographics of our survey respondents (Section 4.5).

4.1 Legal and Ethical Aspects

The use of personal data—including nonconsensually acquired or leaked—has turned into a debated topic in the scientific community over the recent years. Many scholars, including us, concur that there is a scarcity of uniform guidelines, standards, or frameworks for a consistent assessment of legal and ethical risks of potentially problematic studies [13, 18, 23, 32]. This also holds true for security research, in which clear and enforceable codes of conduct are not developed yet [47, 61]. The Menlo Report, a reference document intended to guide computer science research [22], suggests to consult with a research ethics board when reusing existing data. The status quo assumes that scientists apply key principles for ethical research and identify benefits and potential harms of using data of illicit origin [22], while ethics committees expose those studies to heightened scrutiny [13].

Our study design needs to be examined in relation to both ethical and legal aspects. Under the EU General Data Protection Regulation (GDPR), e-mail addresses are personal data, the processing of which requires a legal basis. Consent is one possible legal basis. The nature of our study, however, rendered it infeasible to obtain prior consent from the victims for processing their e-mail addresses. Therefore, we used this personal data item of the affected persons on the legal grounds of the public interest (Article 6, EU GDPR; with the additional safeguards detailed in Article 89, EU GDPR) and the freedom of scientific research (cf. Article 13, EU Charter of Fundamental Rights). An argument in favor of using public interest as legal basis is the absence of alternative methods to contact a significant number of cryptocurrency owners, who are simultaneously customers of a security provider having suffered a data breach. This, in turn, is the prerequisite for our research that serves the public interest.

We use consent as a legal basis for all items collected after the initial contact, thereby adhering closely to the Menlo Report's guidelines [22]. The first page of our survey asked for explicit consent to participation and personal data collection (see Appendix C). This page also included the purpose of this research, intended use, instructions, and contact details. The participants were informed of the anonymous nature of the data collection and an option to withdraw from the survey at any time during its completion. With this procedure, the results and anonymized comments of the participants reported in this paper are collected in compliance with the EU GDPR.

Our review of potential ethical issues started with a search of guidelines, frameworks, and recommendations in the computer science literature [22, 47, 61, 67]. It became evident that most ethical discussions focus on the use case of analyzing data of illicit origin (e. g., user behavioral data) and inferring scientific knowledge directly from it. However, our intended purpose of use was to recruit survey participants by sending unsolicited invitations to the leaked e-mail addresses. This method of participant recruitment is not yet covered in the examined literature. We therefore followed the standard practice and did a risk–benefit analysis [22].

This study concerns the following groups of stakeholders: the victims themselves, the vendor, we as lead researchers, and the affiliated university. We report potential risks individually for each group, followed by mitigation strategies we put in place to minimize harm. With respect to the victims, the

complete leaked dataset contained other sensitive personal information. Following the principle of data minimization, we permanently erased all data fields but the e-mail address. Furthermore, we accessed and stored the dataset in accordance with the EU guidelines for accessing confidential data for scientific purposes. We deleted the dataset and cleaned up our project mail account after the data collection had finished. All these security measures were meant to reduce the risk of another data leakage on our side. The victims could also have experienced additional distress or discomfort when receiving our one-off e-mail. Against this backdrop, we looked for conventional guidelines in the survey invitation design on how to reduce respondent burden [14, 24, 36], and elaborated the subject line and invitation text with succinct, but informative enough details about the purpose and nature of our study. Victims who volunteered to respond to our survey have invested time and cognitive effort, without being compensated financially. We refrained from doing this to mitigate the risk of multiple participation,[3] and the associated biases and response errors, in an anonymous online survey that intentionally did not track responses to invitations.

Our study could have harmed the vendor's reputation if it had found compromising results. To mitigate this potential harm, we informed the vendor of our research initiative before the start of participant recruitment (see Figure 1), and offered a communication channel. When presenting results, we refrain from exposing the vendor unnecessarily. Note that we deliberately avoided a partnership with the vendor to ensure the independence and objectivity of this study. While the cooperation may have alleviated some ethical concerns, it would not have changed the legal basis of data processing as the breached organization had not obtained consent for the purpose of scientific research. More importantly, conducting an independent study allowed us to reduce the risk of uncontrollable response biases caused by potentially reduced consumer trust in the vendor. By doing so, we were also able to maintain the freedom of designing our own questionnaire and controlling the collected data.

As a final stakeholder, we or our university could be confronted with negative reactions or requests, including via e-mail communication or on social media platforms. We also identified the risk that our university domain could be blacklisted by providers abroad who do not share (or do not care about) our legal assessment of the e-mail distribution and thus (falsely) classify it as spam, or blindly follow takedown requests. Throughout our study, we pursued the principle of being transparent and diligent when responding to any incoming or forwarded inquiry via our e-mail account, telephone, or other university staff members [22]. We kept track of a few social media posts discussing the study, however considered our direct intervention unnecessary.

We continue the ethical analysis with a list of potential benefits. First, the dumped data provided us with a unique opportunity to contact the victims and collect their valuable responses on the aftermath and response actions. In contrast to many other breaches, this incident predominantly concerned active or former cryptocurrency users. These persons are deemed to be a hard-to-reach community in behavioral security research [2]. Therefore, neglecting this method of reaching out to potential research subjects would be a missed opportunity in contributing empirical knowledge in this yet unexplored field. Second, we study the effects and behavioral changes caused by the breach exposure among the group of security-literate and aware individuals. Their responses will shed light on the scope and nature of harms even competent users cannot evade after a breach. By eliciting these empirical insights, we seek to add to the social good and advance user-centered breach research. Third, in the absence of financial compensation, this study will likely attract those participants who are interested in contributing to research and thus, willing to provide valuable responses.

Overall, while we could not completely eliminate the risk of causing additional harm to the involved stakeholders, we implemented a series of adequate risk mitigation measures and safeguards for storing and processing e-mail addresses. We submitted the presented risk–benefit analysis as well as the designed questionnaire to the university's IRB and obtained a positive decision after one follow-up request. We kept in close contact with the IRB throughout the recruitment phase and informed it of the progress and received feedback.

4.2 Participant Recruitment

As already mentioned, we made use of the leaked data dump and extracted e-mail addresses to contact victims. Note that we refer to this reduced dataset of leaked addresses in the rest of this paper. The dataset was stored on an external secure USB flash drive kept in a safe vault. Access to it was granted to authorized members of the research group only. We connected the flash drive to a computer for querying the next batch of e-mail addresses in an offline mode only.

Our recruitment strategy relied on *non-personalized*, *automated* e-mail invites which we distributed in batches to randomly selected addresses. Each batch included 2 500 e-mails with a varying proportion of Gmail and non-Gmail accounts. We made this distinction in our e-mail sampling method after distributing the first batch and noticing that Gmail's built-in spam filters marked our invite as spam. We adjusted the header parameters in our mailing script (as we surmise that those most likely triggered anti-spam filters) and increased the number of Gmail accounts stepwise from 200 to 1250 in each following batch. We added four Gmail accounts under our control (in the beginning and at the end of each batch) in order to check for potential spam marking. In total, we initi-

[3] The risk of multiple participation is significant in this target group. A blockchain analysis following the transaction that compensated 961 participants of [38] in bitcoins suggests that the total number of distinct entities is just above 600. One entity has submitted more than 50 responses.

ated the distribution of 13 batches and sent out approximately 31 632 e-mail invites (11.7% of the leaked records).

Our approach to sample is a measure to reduce harm without severely compromising data quality. Without knowledge of the response rate, we targeted a maximum sample size of 500 in the IRB request. The actual collection was terminated when we had reached a number of complete responses that would guarantee at least 100 cases after data cleaning. This number can be justified with a statistical power analysis.[4]

The e-mail invite, reproduced in Appendix A, contained a succinct description and purpose of the study. Following the GDPR requirements, it included information about the processing and storage of an e-mail address as well as the contact details of the principal investigator (PI) and data protection officer. To reassure all contacted persons that the survey links are not personalized, we hosted a static webpage on our research group's website, which contained a short note explaining the purpose of this intermediary page as well as a link to the questionnaire generated by the host platform Qualtrics. Appendix B reports the content of the static webpage.

Figure 2 presents the recruitment workflow and response statistics in greater detail. In particular, we kept track of the number of undeliverable messages which were bounced back to our university account due to abandoned, expired, full, or specially configured mailboxes. In total, we received 943 automatic non-delivery reports, equivalent to about 3% of the total number of e-mails sent. Some of our e-mail invites may have ended up in spam or non-active inboxes, which we were not able to account for. According to the report provided by Qualtrics, 193 contacted persons proceeded to the introduction page of the questionnaire, out of which 186 consented and started the survey; 7 did not consent.

We received a number of follow-up messages requiring further information or an action on our side. A handful of the contacted individuals attempted to authenticate the lead researcher and used a range of creative strategies for that (e. g., via telephone calls or e-mails to a PI's business mailbox, e-mails to the research group's head or university staff members). Twelve persons inquired about the source from which the leaked addresses were collected, (neutrally) questioned the benevolent intentions or recruitment method of the study, or expressed their interest in the final results. There were 4 messages which explicitly provided positive feedback to our research initiative. By contrast, we received 6 and 11 requests exercising the right to erase and object to processing of the e-mail addresses, respectively, and 3 scam attempt messages.

[4] Specifically, 97 or more measurements are needed to have a confidence level of 95% that the true value is within ± 10%-pts. of the estimated value. This is relevant for point estimates of averages. Turning to rare events, with 90 or more samples we have 99% probability to observe experiences at least once that apply to only 5% of the population. By principle, power analysis is limited to the sampling error. The total measurement error is likely higher due to the coverage and non-response errors.

4.3 Survey Instrument

The process of designing a survey instrument was informed by multiple posts from affected victims on a dedicated Reddit discussion forum, which emerged after this incident. A non-systematic exploratory review of the posted messages revealed a set of attack vectors, response actions, and reported examples of harm, which were taken into account when phrasing questions and answer options. While being guided by our primary research goals, we also borrowed some ideas and measurement scales from marketing, risk management, and other data breach studies [1, 49, 50, 54].

The questionnaire consisted of five parts.[5]

Part 1. Customer Experience The survey started with an open-ended question eliciting when and where a participant had learned about the data breach. The next block of questions addressed the purchase, use of, and satisfaction with a device, as well as the customers' expectations and attitudes toward the vendor and the product at the time before and after learning about the breach. We also asked participants whether they are aware of other hardware wallet vendors or even own their products.

Part 2. Harm and Losses This part focused on the breach's aftermath in terms of the harm caused to respondents, experienced attacks with monetary or other types of losses, and elevated concerns. Since some effects may not necessarily have been caused by the data breach or in relation to it, we additionally asked respondents to evaluate their level of confidence in the causal relationship between this incident and the reported crime. We also asked whether a respondent had fallen victim to another breach since January 2020.

Part 3. Individual Response Given our objective to explore individuals' security behaviors, we included a series of questions about precaution or protection actions the surveyed persons may have taken in response to this incident. In order to capture a potential triggering effect of this event on a behavioral change, we explicitly asked in a number of questions whether a respondent had adopted a certain practice *before* or *after* the breach (or not at all). This part also included questions on the discontinuance of using the leaked e-mail address or phone number, the purchased device, or any relationship with the vendor.

Part 4. Corporate Response This part covered respondents' general expectations about corporate response strategies to a breach or theft of customer data. We also attempted to discern customers' (dis)satisfaction with the actions taken by the vendor. Since this part is related to the specifics of organizational responses and communication strategies, it is left out of scope of this paper.

[5] The complete questionnaire is available in an arXiv version of this paper.

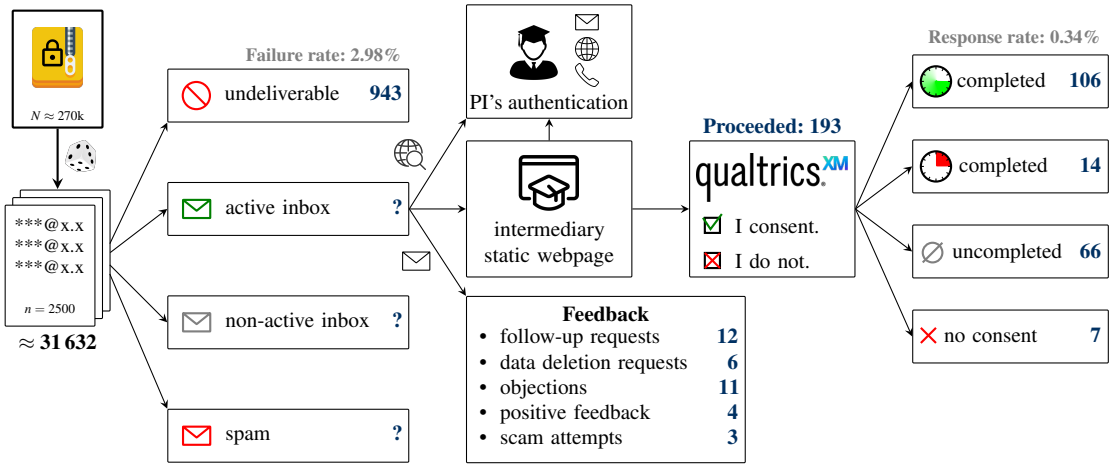


Figure 2: Visualization of the participant recruitment workflow (numbers refer to cases in each branch)

Part 5. Demographics In this final part, we inquired about a respondent's age, gender, age, education, place of residence, and occupation. In addition to the socio-demographic variables, we wished to enrich a descriptive profile of respondents with facts about their level of experience with crypto-assets measured in years and the approximate market value of crypto-assets owned.

The online survey was hosted on the licensed Qualtrics survey platform. In order to protect participants' privacy, we disabled the registration of a respondent's browser, operating system, IP address, or location information in the survey-specific settings. The survey could be accessed by anyone with a valid link. The estimated duration to complete the questionnaire was 25 minutes.

4.4 Data Quality

Several data quality measures were applied to ensure accurate and reliable responses. First, we reduced the dataset of completed responses ($N = 120$) to 106 records which took 15 minutes or longer to complete. We set this time threshold to provide a reasonable balance between fast and slow respondents given the estimated completion time of 25 minutes. Then, we manually reviewed 7 data records which took between 14 and 15 minutes to complete and evaluated their quality. Out of these, we added 5 additional records based on our subjective judgment about their reliability. Given the relatively low number of responses, we also manually inspected open-ended text entry questions as well as response patterns for any presence of the non-differentiation, neutral answer selection, or primacy biases. As a result of this quality check, 7 records were excluded from the backbone set of 106 responses due to suspicious or non-meaningful answers. After having applied these measures, the final dataset resulted into 104 responses, which are analyzed in this paper.

While clearly not representative for any meaningful population, we deem the collected data is of high reliability, since many respondents put extra effort into providing lengthy responses to open-ended questions. Respondents have spent an average of 34 minutes and a median of 26 minutes to fill out the survey. At the same time, we observe a high fraction of incomplete or explicit nondisclosure responses to the socio-demographic and profile questions. This confirms that crypto-asset owners tend to be privacy-conscious individuals who are not willing to disclose too many personal details about themselves.

4.5 Profile of Survey Respondents

Table 1 summarizes the demographics of our sample. Most respondents are men in the age of 25 to 54 years, with a university degree and an employment status. Almost half of the sample have 3–4 years of experience with crypto-assets, while 40% have been owners of crypto-assets for even longer. It comes at no surprise that almost half of the respondents prefer not to reveal the value of their investments. The ones who report (about one third of the sample) say that they hold crypto-assets worth between several thousands up to one million US dollars. (Using exchange rates at the time of completing the survey, not at the time of the breach.) This confirms prior research on a broader sample, which revealed a highly significant relationship between owning more than $10,000 in crypto-assets and the use of a hardware wallet [2, Table 9].

A total of 28 (26.9%) participants report having been a victim of another data breach since January 2020 (43 no; and 33 not aware). The most popular reasons for purchasing a hardware wallet were (i) to store crypto-assets for personal use (87.5%), (ii) to give away as a gift (17.3%), (iii) to review, experiment with and test security features (6.7%), and (iv) to use for business purposes (5.8%). More than half of the

Table 1: Demographics of the survey participants

	Abs.	%		Abs.	%
Gender			**Age**		
Men	88	84.6	24 or younger	2	1.9
Women	2	1.9	Between 25 and 34	20	19.2
Non-binary	2	1.9	Between 35 and 44	24	23.1
I prefer not to answer.	12	11.5	Between 45 and 54	26	25.0
			Between 55 and 64	10	9.6
			65 or older	7	6.7
			I prefer not to answer.	15	14.4
Current occupation			**Formal education**		
Student	0	0.0	Less than high school	0	0.0
Skilled manual worker	2	1.9	High school incomplete	5	4.8
Employed in a service job	6	5.8	High school graduate	11	10.6
Self-employed	14	13.5	College or assoc. degree	14	13.5
Unemployed	2	1.9	Bachelor's degree	25	24.0
Retired / sickness leave	9	8.7	Master's degree	29	27.9
Employed professional	56	53.8	Doctoral degree	6	5.8
Other	1	1.0	Other professional degree	4	3.8
I prefer not to answer.	14	13.5	I prefer not to answer.	10	9.6
Experience with crypto-assets			**Ownership of crypto-assets**		
Less than 1 year	0	0.0	Less than $1,000	7	6.7
Between 1 and 2 years	12	11.5	$1,000 – $5,000	5	4.8
Between 3 and 4 years	48	46.2	$5,000 – $10,000	4	3.8
Between 5 and 6 years	18	17.3	$10,000 – $100,000	18	17.3
More than 6 years	24	23.1	$100,000 – $1,000,000	15	14.4
No answer	2	1.9	More than $1,000,000	4	3.8
			I prefer not to answer.	51	49.0

sample (56.7%) report that they continue to use a purchased device, as opposed to those who never used it (11.5%), or discontinued their usage before (3.8%) or after (17.3%) the breach. Out of those 12 respondents who never used the product, 6 purchased it in the hope of using it for personal needs, 3 as a giveaway or gift, and one for research.

Besides Ledger, there are competing producers of hardware wallet devices (e. g., Trezor, Ellipal, or KeepKey). The majority of respondents (76%) report not using similar products of other brands. The most common factors which respondents said influenced their decision to buy the Ledger product were (i) the vendor's reputation (75%), (ii) product's technical security features (54.8%), (iii) the number and type of supported crypto-assets and services (44.2%), and (iv) recommendations from members of the community (39.4%). These results suggest that Ledger's products are cherished and recognized by the global customer base for their security attributes and a broad and ever-growing range of supported crypto-assets.

5 Empirical Results

We structure the presentation of our main results along the research questions.

5.1 Experienced Harms and Attacks (RQ1)

In general, victims of a data breach may anticipate or experience harms of different kinds. While acknowledging that beliefs about harm severity are often subjective and contingent on individual circumstances [50], we differentiate between serious damages and distress, such as account takeover attacks, physical threats, or money thefts; and less serious or negligible repercussions in form of scam, marketing, or nuisance e-mails, text messages, or phone calls. Figure 3 presents the frequency of experienced harms reported by the respondents. It should be emphasized that retrospective reports of the victims should be interpreted with caution due to potential recall bias and confounding factors, such as a respondent's involvement in another data breach or limited causal reasoning.

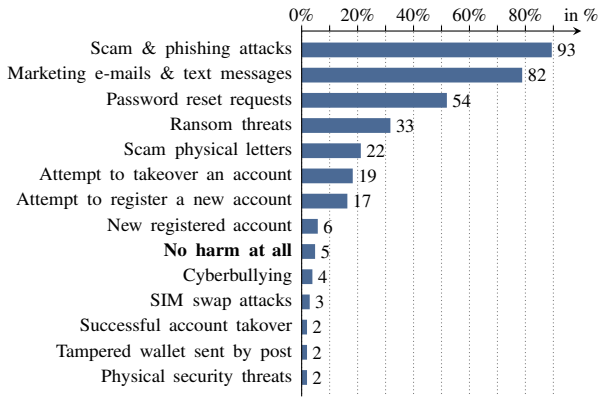


Figure 3: Reported harms and attacks (bars visualize percentages; numbers refer to absolute frequencies)

In line with prior research (e. g., [50]), the most prevalent forms of harm were scams, phishing attacks, marketing messages, and phone calls. Frequently, criminals were impersonating the vendor in an attempt to trick their target victims into sharing recovery phrases, updating devices, or installing malware on victim's computers. The victims were annoyed by targeted phishing e-mails and unsolicited marketing calls promoting crypto-asset investment opportunities. Almost half of the sample received password reset requests for accounts registered with a leaked e-mail. Criminal attempts to take over the control of online accounts were reported by 19 victims (18%). There are also 33 (32%) reports of blackmailing and threats to disclose allegedly available personal data.

More serious forms of harm and advanced attack vectors seem to be rarer. Two instances of the successful account takeover were reported, while four persons received cyberbullying or physical threats. Two respondents reported receiving an authentically-looking parcel with a tampered crypto-wallet inside it, which was allegedly sent by the vendor with the security advice to transfer crypto-assets to the new device. While the numbers are too small for quantification attempts, we interpret this as existential evidence for resourceful attacks, requiring both technical expertise and upfront investments.

Out of the 104 responses, 10 reported financial losses which were incurred after July 2020. Figure 4 visualizes the magnitude of loss in a cryptocurrency, as reported by each participant. For the sake of comparability, we converted each number to its equivalent in US dollar by taking the average

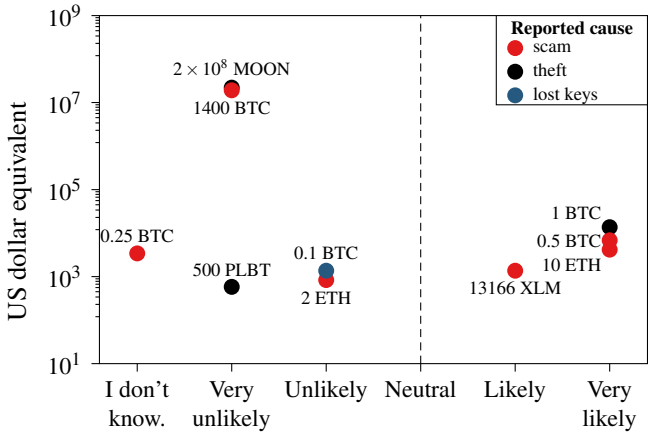


Figure 4: Reported financial losses

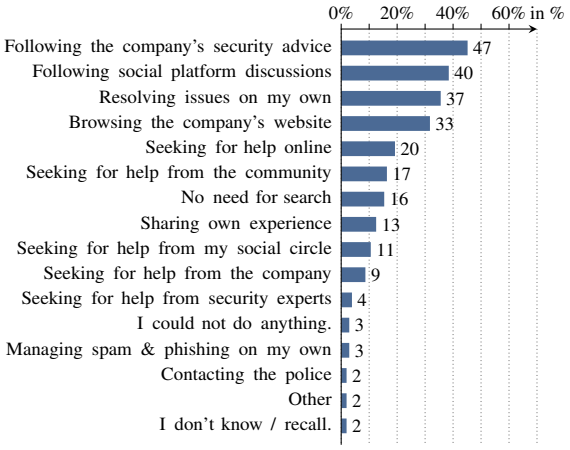


Figure 5: Information search strategies used (bars visualize percentages; numbers refer to absolute frequencies)

market price of a cryptocurrency in the second half of 2020. This approximation is dictated by the lack of exact timing information on the loss event. As visible in Figure 4, most losses were attributed to scams and thefts (6 and 3 cases, respectively) and affected bitcoin and ether savings. However, according to the respondents' self-reports, only 4 events with losses ranging from one thousand up to several thousand US dollars are 'likely' or 'very likely' associated with the data breach. So, with some level of certainty, there were sporadic cases of successful cryptocurrency-related scams and thefts. Although the breach had no direct impact on the wallets and concerned sales information only, it provided criminals with channels to contact and take advantage of vulnerable victims.

Besides asset security, physical safety risks (e. g., burglaries, assaults or kidnapping) could be of concern in the community of crypto-asset owners. Our data includes two reported instances of exposure to a physical threat. In response to another question about individual's concerns, 30 respondents reported having moderate to extreme concerns about their physical safety. Quoting one respondent: *"I used a temporary (exclusive to this vendor) e-mail and phone number for my order, but a postal address was genuine and thus presents the biggest risk for me and others working there."*

To sum up, the majority of the reported effects resembles those of typical data breaches and we find a few instances of threats which are characteristic to this case (e. g., fake physical devices, perceived safety hazard). At the same time, 5 respondents stated not having experienced any harm.

5.2 Response Actions (RQ2)

As for the analysis of response actions, one should keep in mind that crypto-asset owners tend to be security-conscious, risk-aware, and proactive in taking protective actions [45]. This is, however, balanced by the growth of the global community of crypto-investors, which naturally becomes more heterogeneous [2]. Looking at the socio-demographics of our sample (Table 1), we observe that the majority of the surveyed victims have multiple years of experience with crypto-assets. It is safe to assume that they are familiar with basic security risks and protective measures relating to crypto-assets.

To set the context, it is instructive to note that the large majority of victims in our sample has heard about the breach before suffering harm. Almost half of the sample learned about it through the notification e-mail from the vendor or its newsletter, another 40% from social media and online media; only 7 individuals detected dubious activities first, and one person reportedly was hit unprepared and confronted with an empty wallet. This means most response actions were post-breach, but still pro-active in relation to follow-up attacks.

Next we analyse from which sources of information the victims acquired help in relation to protective response actions. As shown in Figure 5, the survey participants largely took the vendor's security advice (45%) or followed suggestions from online peer discussions (38%). This evidence hints at the importance of timely releases of advice on response strategies by breached organizations. It also suggests that opinion leaders in the crypto-asset community enjoy quite some trust.

Figure 6 presents a list of the adopted protection measures ordered by the observed frequency. These represent actions the victims (reportedly) took in response to this breach. Some are of a general nature (e. g., using or extending the use of multi-factor authentication (MFA) or password managers), while others apply specifically to crypto-assets (e. g., changing passwords for crypto accounts or relocating backups of recovery seeds to a safer place). Overall, the most common practices were the increased use of MFA (52%) or specialized authenticator apps (44%), and the change of passwords for e-mail (47%) and crypto-related accounts (41%). In relation to the heightened safety concerns mentioned earlier, 18

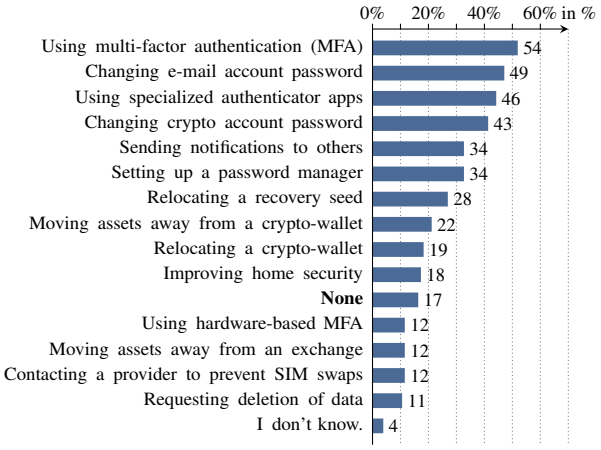


Figure 6: Adopted protection measures (bars visualize percentages; numbers refer to absolute frequencies)

respondents reported improving their home security.

Interestingly, 22 respondents reported moving crypto-assets away from their crypto-wallets. This is a thought-provoking insight given that crypto-wallets are generally supposed to be secure and, more specifically, were not compromised by the breach. Unfortunately, our survey did not include a follow-up question on the nature and security level of the alternative storage options chosen by the victims who moved their funds.

As a lesson learned from this case, the affected victims might have revised their security behaviors and adopted—or intended to adopt—measures which may reduce the probability of experiencing harm and its magnitude as an effect of a similar breach. For instance, one may allocate an e-mail address exclusive to crypto or financial operations, or use one-time addresses, postal boxes, or delivery forwarding services for online orders. To study these breach-specific implications, we showed respondents a list of 8 security practices and asked which they had practiced before or in response to the breach. If a person has not adopted a certain practice yet, we inquired the likelihood of doing so on a 5-point rating scale. Figure 7 summarizes the responses.

Probably because of its ease of implementation, the most-adopted action involves using dedicated e-mail addresses for any crypto-related business. In fact, managing multiple e-mail aliases and separating accounts for critical and general matters are what several respondents emphasized retrospectively: *"I am sad that I did not use a special e-mail address when purchasing the device."* This is followed by the use of fake phone numbers or "burner" e-mail addresses for online orders. The use of fake names, decoy storage devices in case of a physical extortion of crypto-assets, or post office boxes for the delivery of online orders are less common, arguably due to practical limitations or inaccessibility. Finally, using postal forwarding services for the delivery of online orders or dedicated SIM cards for any business with crypto-assets were least adopted

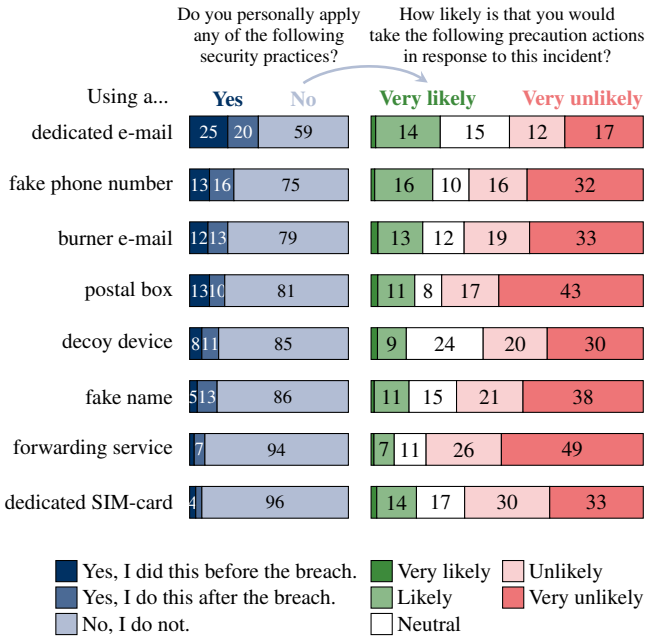


Figure 7: Security practices and adoption intentions (numbers refer to absolute frequencies)

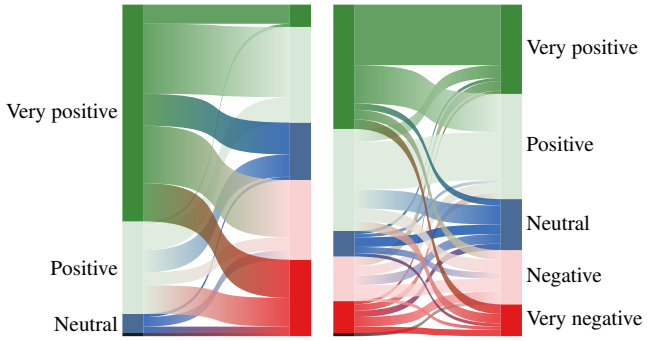
by the victims. When looking at the adoption intentions of the respondents who did not apply these measures yet, the overall patterns remain largely consistent with those of (self-reported) actual behaviors.

Seeking for legal protection and compensation is another possible response action to a breach. Two persons in our sample reported being part of a class-action lawsuit against the security vendor.[6] Many open-ended comments on this matter were along the lines of *"too much effort with little perspective of success"*, *"too expensive"*, *"time and money needed would not hold up against potential benefits"*, or *"harm is already done."* So, victims seem to be rather skeptical about pursuing litigation against breached entities. Specifically to our case, this skepticism could have been caused by difficulties of proving asset losses or suffering from cognizable damage as a result of the breach.

5.3 Consumer Attitudes (RQ3)

Our review of Reddit messages has shown a salient divergence of the post-breach attitudes of victims towards the security vendor as compared to its products. There were cases of negative reactions to the incident, blaming the vendor for failing to protect customer data. At the same time, some consumers publicly admitted that while they may no longer trust the vendor, the purchased product continues to meet their needs and

[6] We are not aware of class actions against the sales partner and did not offer this answer option. Future survey instruments could improve by considering all involved parties as potential defendants.



(a) Attitudes toward the vendor (b) Attitudes toward the product

Figure 8: Consumer attitudes *before* and *after* the breach

original expectations: *"[...] I have no issue with the devices and consider them to be solid for the purpose I use them."*[7]

A potential divergence in attitudes towards a company's societal obligations and product expertise has been a topic of interest in the marketing literature for many years [15]. Specifically, individuals' corporate image associations are empirically found to influence—sometimes counter-intuitively—product evaluations [15, 28]. Being inspired by this firm-versus-product level of abstraction, we asked respondents to rate their attitudes towards the vendor and, separately, the security device using 5-point rating scales (1 – very negative, 5 – very positive). In order to track any attitudinal changes triggered by the breach, the question wording explicitly invited the participants to recollect their attitudes at the time before and after the breach.

Figure 8 summarizes the observed patterns in two Sankey diagrams, showing the apparent changes in individuals' attitudes. Not very surprisingly, almost half of our sample reports *negative* or *very negative* post-breach attitudes towards the vendor. Likewise, only 6% out of 65% of those who recollect to have a very positive attitude prior to the incident remained loyal in their ex-post evaluation. On the other hand, the breach had a moderate effect on the consumers' attitudes towards the product: there are no drastic changes between the different levels of the attitude scale. Statistically, a Wilcoxon rank sum test confirms that the differences in attitudes towards the vendor are highly significant ($p < 0.001$), whereas the attitudes towards the product do not differ significantly ($p = 0.095$).

6 Discussion

This case study sheds light on a high-profile data breach which exposed a population of affluent crypto-asset owners as potentially attractive targets for opportunistic profit-oriented crimes. In line with prior work [35, 50, 72], our findings show that scam, phishing and attempted identity thefts characterize the aftermaths of breaches. This specific case reveals a physical attack vector previously unseen in data breach research (to the best of our knowledge). It could explain our finding of elevated safety concerns among the victims whose residence addresses were leaked. As some respondents pointed out: changing a place of living is not as easy as switching to a new e-mail address.

In Section 6.1, we discuss the implications of our study and its key findings for the cybersecurity industry and related policy initiatives. Next, we share our lessons learned regarding the methodological challenges in data breach research (Section 6.2). Section 6.3 addresses potential limitations.

6.1 Implications for Industry and Policy

The victims in our sample were exposed to a range of crimes exploiting almost all leaked data items. Most reported attacks were targeted, domain-specific, and plausibly related to the breach event, even though our method cannot establish causality. This led us to the insight that breaches affecting security vendors (and by extension security service providers) deserve special attention in data breach research. We call these breaches 'high-profile' and note their potentially far-reaching effects because they enable targeted follow-up crimes. In this sense, the identification of affluent crypto-asset holders as users of a Ledger wallet is qualitatively similar to the exposure of website user accounts in the recent LastPass case.

Beyond doubt, breaches will continue to plague the world. Against this background, it is essential like never before that the cybersecurity industry serves as a role model of effective information security and crisis management. After all, companies whose mission is to provide excellence in security products and services should stand by their own promises. In the age of digital transformation, these expectations add a new form of corporate social responsibility, which, as this case has shown, extends beyond the boundaries of the security vendor. In particular security companies should exercise extreme caution when choosing business partners and avoid sales channels not designed for dealing with the risk of a high-profile breach. Several of our findings reveal a potential disconnect between victims' concerns and material harm suffered. This underscores the importance of effective corporate crisis communication and the management of reputation threats, in particular on social media platforms [65].

When looking at the anatomy of the case at large, we conjecture that more harm was prevented through perceptions and mitigating actions of victims and by the vendor. Owners of crypto-wallets are reputedly the most mistrustful among the already security-concerned crowd of crypto-asset owners. This trait, which we also observed during recruitment, further connects to a found ability of the surveyed participants to evaluate the security properties of products independent of a vendor's reputation. While this may be specific to the target

[7] https://www.reddit.com/r/ledgerwalletleak/comments/kpzm4x/is_anyone_still_using_ledger_had_a_close_call/

group of our study, it may call for a critical (and empirical) reassessment of a commonly accepted fact in the security economics literature, namely that consumers cannot evaluate security properties when making purchase decisions [4].

Stricter vendor liability is often postulated as a cure for market failures and as a means to incentivize security practices and investments [3, 41]. For example, the 2023 US National Cybersecurity Strategy sets out to "reshape laws that govern liability for data losses and harm caused by cybersecurity errors" [66, p. 19]. However, the liability channel is prone to fail if victims keep skeptical about its success and do not support it wholeheartedly, as our results have demonstrated (albeit under the current lax liability regime). This calls for studying the subjective causes of perceived skepticism towards litigation and a potential reform of cybersecurity law [37].

Prior work [50] puts forward an idea of integrating the automated generation of unique e-mail aliases into account registration workflows. While this practice has its own merits and interest from the surveyed participants, our results hint at limitations of this approach, too. One-time emails may function as double-edged swords, as they may make it harder for victims to learn about a breach notification. In addition, their utility is uncertain in the light of a permanent shipping address with the ensuing—perceived or real—risk of burglary or physical extortion in the worst case. In this regard, future research may explore workable design options for account registration and order checkout processes which would anticipate potential data leakages and adverse effects thereof.

6.2 Implications for Breach Research

In terms of methodology, our study has taught us some unanticipated lessons. As we started recruitment, we were surprised to see that some victims made quite some efforts to convince themselves of the authenticity of our invitation as well as of the researcher's identity. For example, they contacted other university staff members through various channels. While this is generally a positive sign of high risk awareness, our ethical analysis missed these activities as a form of additional (though very minor) cost on victims, researchers, and researchers' colleagues. It is unimaginable what would have happened had we used a non-existing persona as sender; an idea we contemplated to protect the lead researcher, but dismissed eventually. Related to this, at the time of applying for IRB approval, we were unaware of the vendor's post-breach initiative to crowd-source reports of suspicious phishing campaigns from the community. Therefore, our study might have inadvertently created additional work for the vendor's customer support agents processing those notices. We tried to inquire the cost after the study has ended to inform future ethics assessments, but we did not receive an estimate. Likewise, we were unaware of the vendor's take-down efforts, which might have increased the risk of blacklisting and related collateral damage to other users of our university's infrastructure. In hindsight, this appears obvious given the size of their business and the availability of professional services. We encourage other researchers to take these contingent actions into account in their ethics analyses.

Finally, we observe that many forms of experienced harm repeat across case studies of breaches [55, 59, 68]. Therefore, data breach research could benefit from the development of standard survey questions and response scales. This includes the refinement of measures of harm, an incredibly difficult concept with high relevance also in the field of breach litigation [37]. More standardized measurement instruments would allow for more harmonized reporting of empirical results, improve comparability between cases, and pave the way for meaningful quantitative meta-analyses in the future.

6.3 Limitations

Our work has a number of limitations inherent to survey-based cybercrime studies. It is evident that findings from a sample of volunteers among victims affected by a high-profile breach should be interpreted with caution. Known sources of coverage error include the receipt of the invitation, which is correlated with the e-mail provider's spam filtering practices. Moreover, we could not reach victims who implemented certain security precautions, such as using one-time e-mail addresses. Each individual decision to participate in our survey may be correlated with victim experience, for example the salience of perceived or actual harm; as well as victim's aptitude to report on sensitive topics in English. Other causes of non-response are victims' subjective expectations towards receiving personalized invitations and reminders of pending survey invitations, which we refrained from sending for ethical considerations. All this explains the low response rate [20]. Yet the sample size appears decent given that we have confirmed victims of a high-profile breach (cf. Footnote 4). Other studies interview thousands in order to obtain a sample of a few hundred victims of everyday cybercrime [56].

Turning to item reliability and validity, our survey instrument lacks effective mechanisms to verify and prevent inaccurate reporting. With self-reported statements, we cannot rule out response errors, strategic responses in hope for redress, and social desirability bias [72]. Also, recall bias and the capability of causal reasoning may vary between individuals. While post-breach attacks can occur with some delay, individuals' ability to accurately recollect past events tends to degrade. This makes it tricky to optimize the timing of surveys which collect data on reciprocal causal effects. We hope that our choice of one year gave appropriate time for the influencing factors and effect variables to manifest, while important episodes are still in memory. A longitudinal design could in principle overcome some of these shortcomings, however it is even harder to implement in an ethical manner. Moreover, recollections of effects of this data breach may be confounded by other breaches and attacks. Finally, the survey

instrument may be prone to unknown ceiling effects resulting from the chosen question wording and response scales [17]. We had to make some compromises given the scarcity of standardized and validated item batteries in our domain. Closing this gap—certainly not on high-profile cases at first—could catalyze the development of breach research.

7 Conclusion

We document one of the first empirical case studies on high-profile data breaches affecting vendors of critical security products. While in this case the security of the product itself was not affected, the disclosed sales information has enabled or facilitated new attack vectors. At the same time, it enabled us, researchers, to ethically, legally, technically, and practically explore the practice of sampling an otherwise hard-to-reach victim population from leaked contact information for the purpose of breach research. We have reported existential and (limited) quantitative evidence on the case, derived lessons learned, and demonstrated the feasibility of the sampling technique. The effort is tremendous and researchers have to exercise a lot of patience before getting hold of data or results. Time will tell if our approach becomes more commonly accepted and eventually easier to pursue. What we can state with more certainty is that the security community is well-advised in learning from high-profile breaches.

Acknowledgments

The authors are grateful to the anonymous reviewers and the paper shepherd for their valuable critical comments on the manuscript. We also thank Jérémie Glossi for his technical support and valuable contributions to the study, our colleagues Nicole Krismer-Stern and Paulina Jo Pesch for legal consultation, Leonid Risteski for his help in producing Figure 1, and Daniel W. Woods for very thoughtful comments on a draft of this paper.

References

[1] Lillian Ablon, Paul Heaton, Diana Catherine Lavery, and Sasha Romanosky. Consumer Attitudes Toward Data Breach Notifications and Loss of Personal Information. Technical report, Santa Monica, CA, USA, 2016.

[2] Svetlana Abramova, Artemij Voskobojnikov, Konstantin Beznosov, and Rainer Böhme. Bits Under the Mattress: Understanding Different Risk Perceptions and Security Behaviors of Crypto-Asset Users. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery. doi: 10.1145/3411764.3445679.

[3] Ross Anderson. Liability and Computer Security: Nine Principles. In *Proceedings of the Third European Symposium on Research in Computer Security (ESORICS)*, pages 231–245, Brighton, United Kingdom, 1994. Springer.

[4] Ross Anderson. Why Information Security is Hard—An Economic Perspective. In *Proceedings of the Computer Security Applications Conference*, pages 358–365. IEEE, 2001.

[5] Corey M. Angst, Emily S. Block, John D'Arcy, and Ken Kelley. When Do IT Security Investments Matter? Accounting for the Influence of Institutional Factors in the Context of Healthcare Data Breaches. *MIS Quarterly*, 41(3):1–24, 2017.

[6] Andreas M. Antonopoulos. *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. O'Reilly Media, Inc., Sebastopol, CA, 2014.

[7] Emmanuel W. Ayaburi and Daniel N. Treku. Effect of Penitence on Social Media Trust and Privacy Concerns: The Case of Facebook. *International Journal of Information Management*, 50:171–181, 2020.

[8] Eric Bachura, Rohit Valecha, Rui Chen, and H Raghav Rao. The OPM Data Breach: An Investigation of Shared Emotional Reactions on Twitter. *MIS Quarterly*, 46(2), 2022.

[9] Sarah Bana, Erik Brynjolfsson, Wang Jin, Sebastian Steffen, and Xiupeng Wang. Human Capital Acquisition in Response to Data Breaches. In *The 21st Workshop on the Economics of Information Security (WEIS 2022)*, Tulsa, Oklahoma, USA, 2022.

[10] Gaurav Bansal and Merrill Warkentin. Do You Still Trust? The Role of Age, Gender, and Privacy Concern on Trust after Insider Data Breaches. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 52(4):9–44, 2021.

[11] Massimo Bartoletti, Stefano Lande, Andrea Loddo, Livio Pompianu, and Sergio Serusi. Cryptocurrency Scams: Analysis and Perspectives. *IEEE Access*, 9: 148353–148373, 2021.

[12] Sruti Bhagavatula, Lujo Bauer, and Apu Kapadia. (How) Do People Change Their Passwords After a Breach? arXiv preprint 2010.09853, 2020. URL https://arxiv.org/abs/2010.09853. [Accessed: 2023-02-06].

[13] Anne E. Boustead and Trey Herr. Analyzing the Ethical Implications of Research Using Leaked Data. *PS: Political Science and Politics*, 53(3):505–509, 2020. doi: 10.1017/S1049096520000323.

[14] Philip .S Brenner, Carol Cosenza, and Floyd J. Fowler Jr. Which Subject Lines and Messages Improve Response to E-mail Invitations to Web Surveys? *Field Methods*, 32(4):365–382, 2020.

[15] Tom J. Brown and Peter A. Dacin. The company and the Product: Corporate Associations and Consumer Product Responses. *Journal of Marketing*, 61(1):68–84, 1997.

[16] Microsoft Security Response Center. Investigation Regarding Misconfigured Microsoft Storage Location, 2022. URL https://msrc-blog.microsoft.com/2022/10/19/investigation-regarding-misconfigured-microsoft-storage-location-2/. [Accessed: 2023-03-06].

[17] Seung Youn Chyung, Douglas Hutchinson, and Jennifer A. Shamsy. Evidence-Based Survey Design: Ceiling Effects Associated with Response Scales. *Performance Improvement*, 59(6):6–13, 2020.

[18] Karin Clark, Matt Duckham, Marilys Guillemin, Assunta Hunter, Jodie McVernon, Christine O'Keefe, Cathy Pitkin, Steven Prawer, Richard Sinnott, Deborah Warr, and Jenny Waycott. Advancing the Ethical Use of Digital Data in Human Research: Challenges and Strategies to Promote Ethical Practice. *Ethics and Information Technology*, 21(1):59–73, 2019. doi: 10.1007/s10676-018-9490-4.

[19] Jessica Colnago, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Lorrie Cranor, and Nicolas Christin. "It's Not Actually That Horrible": Exploring Adoption of Two-Factor Authentication at a University. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–11, New York, NY, USA, 2018. Association for Computing Machinery. doi: 10.1145/3173574.3174030.

[20] Colleen Cook, Fred Heath, and Russel L. Thompson. A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. *Educational and Psychological Measurement*, 60(6):821–836, 2000.

[21] Cassandra Cross, Megan Parker, and Daniel Sansom. Media Discourses Surrounding 'Non-Ideal' Victims: The Case of the Ashley Madison Data Breach. *International Review of Victimology*, 25(1):53–69, 2019.

[22] David Dittrich and Erin Kenneally. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. https://catalog.caida.org/paper/2012_menlo_report_actual_formatted. [Accessed: 2023-02-06].

[23] Serge Egelman, Joseph Bonneau, Sonia Chiasson, David Dittrich, and Stuart Schechter. It's Not Stealing If You Need It: A Panel on the Ethics of Performing Research Using Public Data of Illicit Origin. In Jim Blyth, Sven Dietrich, and L. Jean Camp, editors, *Financial Cryptography and Data Security*, pages 124–132, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[24] Hafsteinn Einarsson, Alexandru Cernat, and Natalie Shlomo. Reducing Respondent Burden with Efficient Survey Invitation Design. *Survey Research Methods*, 15 (3):207–233, 2021.

[25] Michael Fagan and Mohammad Maifi Hasan Khan. Why Do They Do What They Do?: A Study of What Motivates Users to (Not) Follow Computer Security Advice. In *Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 59–75. Denver, CO, USA, 2016.

[26] Rejikumar G, Gopikumar V, K.G.Sofi Dinesh, Aswathy Asokan-Ajitha, and Ajay Jose. Privacy breach perceptions and litigation intentions: Evidence from e-commerce customers. *IIMB Management Review*, 33 (4):322–336, 2021. doi: https://doi.org/10.1016/j.iimb.2021.11.001.

[27] Yixing Lisa Gao, Lu Zhang, and Wei Wei. The Effect of Perceived Error Stability, Brand Perception, and Relationship Norms on Consumer Reaction to Data Breaches. *International Journal of Hospitality Management*, 94, 2021. doi: https://doi.org/10.1016/j.ijhm.2020.102802.

[28] Zeynep Gürhan-Canli and Rajeev Batra. When Corporate Image Affects Product Evaluations: The Moderating Role of Perceived Risk. *Journal of Marketing Research*, 41(2):197–205, 2004.

[29] Kholekile L. Gwebu, Jing Wang, and Li Wang. The Role of Corporate Reputation and Crisis Response Strategies in Data Breach Management. *Journal of Management Information Systems*, 35(2):683–714, 2018.

[30] J.T. Hamrick, Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek. An Examination of the Cryptocurrency Pump-and-Dump Ecosystem. *Information Processing & Management*, 58 (4), 2021.

[31] Zahra Hassanzadeh, Robert Biddle, and Sky Marsen. User Perception of Data Breaches. *IEEE Transactions on Professional Communication*, 64(4):374–389, 2021.

[32] Marcello Ienca and Effy Vayena. Is It Ethical to Use Hacked Data in Scientific Research? *Available at SSRN 3843733*, 2021. doi: http://dx.doi.org/10.2139/ssrn.3843733.

[33] Ponemon Institute. The Aftermath of a Data Breach: Consumer Sentiment, 2014. Technical Report.

[34] Koteswara Ivaturi and Akshay Bhagwatwar. Mapping Sentiments to Themes of Customer Reactions on Social Media During a Security Hack: A Justice Theory Perspective. *Information & Management*, 57(4):1–13, 2020.

[35] Sowmya Karunakaran, Kurt Thomas, Elie Bursztein, and Oxana Comanescu. Data Breaches: User Comprehension, Expectations, and Concerns with Handling Exposed Data. In *Proceedings of the Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 217–234, Baltimore, MD, USA, 2018.

[36] Mareile Kaufmann and Meropi Tzanetakis. Doing Internet Research with Hard-to-Reach Communities: Methodological Reflections on Gaining Meaningful Access. *Qualitative Research*, 20(6):927–944, 2020.

[37] Ido Kilovaty. Psychological Data Breach Harms. *North Carolina Journal of Law & Technology*, 23(1):1–66, 2021.

[38] Katharina Krombholz, Aljosha Judmayer, Matthias Gusenbauer, and Edgar Weippl. The other side of the coin: User experiences with Bitcoin security and privacy. In *Financial Cryptography and Data Security*, volume 9603 of *Lecture Notes in Computer Science*, pages 555–580, Berlin Heidelberg, 2017. Springer.

[39] Thomas Kude, Hartmut Hoehle, and Tracy Ann Sykes. Big Data Breaches and Customer Compensation Strategies: Personality Traits and Social Influence as Antecedents of Perceived Compensation. *International Journal of Operations & Production Management*, 37 (1):56–74, 2017.

[40] LastPass. Notice of Recent Security Incident, 2022. URL https://blog.lastpass.com/2022/12/notice-of-recent-security-incident/. [Accessed: 2023-03-06].

[41] Aron Laszka, Mark Felegyhazi, and Levente Buttyan. A survey of interdependent information security games. *ACM Computing Surveys (CSUR)*, 47(2):1–38, 2014.

[42] Ledger. Addressing the July 2020 E-Commerce and Marketing Data Breach – A Message from Ledger's Leadership, 2020. URL https://www.ledger.com/addressing-the-july-2020-e-commerce-and-marketing-data-breach. [Accessed: 2023-03-06].

[43] Ledger. Beta Testers' AMA: Demystifying the Data Breach with Ledger's CISO, 2021. URL https://www.ledger.com/blog/beta-testers-ama-demystifying-the-data-breach-with-matt-johnson. [Accessed: 2023-03-06].

[44] Yangfan Liang and Rahul Telang. Customer Response to Adverse Security Events: An Empirical Study. *Available at SSRN 3523788*, 2020. doi: https://dx.doi.org/10.2139/ssrn.3523788.

[45] Gunnar Lindqvist, Joakim Kävrestad, Dennis Modig, and Ali Mohammad Padyab. How do Bitcoin Users Manage Their Private Keys? In *7th International Workshop on Socio-Technical Perspective in IS Development (STPIS 2021)*, pages 11–21, Trento, Italy, (online), 2021.

[46] Schneider Wallace Cottrell Konecky LLP. Ledger and Shopify Class Action Over Allegations of a Data Breach Cover Up , 2021. URL https://blog.lastpass.com/2022/12/notice-of-recent-security-incident/. [Accessed: 2023-03-06].

[47] Kevin Macnish and Jeroen van der Ham. Ethics in Cybersecurity Research and Practice. *Technology in Society*, 63:1–10, 2020. doi: https://doi.org/10.1016/j.techsoc.2020.101382.

[48] Easwar Vivek Mangipudi, Udit Desai, Mohsen Minaei, Mainack Mondal, and Aniket Kate. Uncovering Impact of Mental Models towards Adoption of Multi-device Crypto-Wallets. Cryptology ePrint Archive, Paper 2022/075, 2022. URL https://eprint.iacr.org/2022/075. [Accessed: 2023-06-02].

[49] Kristin Masuch, Maike Greve, Jens Cyrenius, Benjamin Wimmel, and Simon Trang. Do I Get What I Expect? An Experimental Investigation of Different Data Breach Recovery Actions. In *Proceedings of the Twenty-Eighth European Conference on Information Systems (ECIS 2020)*, pages 1–18, Marrakesh, Morocco, 2020.

[50] Peter Mayer, Yixin Zou, Florian Schaub, and Adam J. Aviv. "Now I'm a bit angry:" Individuals' Awareness, Perception, and Responses to Data Breaches that Affected Them. In *Proceedings of the 30th USENIX Security Symposium (USENIX Security 21)*, pages 393–410. USENIX Association, 2021.

[51] Minh Hieu Nguyen Ba, Jacob Bennett, Michael Gallagher, and Suman Bhunia. A Case Study of Credential Stuffing Attack: Canva Data Breach. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 735–740, 2021. doi: 10.1109/CSCI54926.2021.00187.

[52] Alison N. Novak and M. Olguta Vilceanu. "The internet is not pleased": Twitter and the 2017 Equifax Data Breach. *The Communication Review*, 22(3):196–221, 2019.

[53] Ehsan Nowroozi, Seyedsadra Seyedshoari, Yassine Mekdad, Erkay Savaş, and Mauro Conti. *Cryptocurrency*

Wallets: Assessment and Security, pages 1–19. Springer International Publishing, Cham, 2023. doi: 10.1007/978-3-031-25506-9_1.

[54] The Regents of the University of Michigan. American Customer Satisfaction Index. Methodology Report. Technical report, Ann Arbor, MI, USA, 2008.

[55] Peng Peng, Chao Xu, Luke Quinn, Hang Hu, Bimal Viswanath, and Gang Wang. What Happens After You Leak Your Password: Understanding Credential Sharing on Phishing Sites. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, Asia CCS '19, pages 181–192, New York, NY, USA, 2019. Association for Computing Machinery. doi: 10.1145/3321705.3329818.

[56] Markus Riek and Rainer Böhme. The Costs of Consumer-facing Cybercrime: An Empirical Exploration of Measurement Issues and Estimates. *Journal of Cybersecurity*, 4(1), 2018.

[57] Sasha Romanosky. Examining the Costs and Causes of Cyber Incidents. *Journal of Cybersecurity*, 2(2):121–135, 2016.

[58] Sasha Romanosky, Rahul Telang, and Alessandro Acquisti. Do data breach disclosure laws reduce identity theft? *Journal of Policy Analysis and Management*, 30 (2):256–286, 2011.

[59] Sasha Romanosky, David Hoffman, and Alessandro Acquisti. Empirical Analysis of Data Breach Litigation. *Journal of Empirical Legal Studies*, 11(1):74–104, 2014.

[60] Frederic Schlackl, Nico Link, and Hartmut Hoehle. Antecedents and Consequences of Data Breaches: A Systematic Review. *Information & Management*, page 103638, 2022.

[61] Sebastian Schrittwieser, Martin Mulazzani, and Edgar Weippl. Ethics in Security Research Which Lines Should Not Be Crossed? In *2013 IEEE Security and Privacy Workshops*, pages 1–4, San Francisco, CA, USA, 2013. doi: 10.1109/spw.2013.6914700.

[62] Annika Selzer, Daniel W. Woods, and Rainer Böhme. An Economic Analysis of Appropriateness under Article 32 GDPR. *European Data Protection Law Review*, 7(3):456–470, 2021.

[63] Shopify. Incident update, 2020. URL https://community.shopify.com/c/shopify-discussions/incident-update/m-p/888971/highlight/true#M197487. [Accessed: 2023-05-25].

[64] Romilla Syed. Enterprise Reputation Threats on Social Media: A Case of Data Breach Framing. *The Journal of Strategic Information Systems*, 28(3):257–274, 2019. doi: https://doi.org/10.1016/j.jsis.2018.12.001.

[65] Romilla Syed. Enterprise Reputation Threats on Social Media: A Case of Data Breach Framing. *The Journal of Strategic Information Systems*, 28(3):257–274, 2019.

[66] The White House. *National Cybersecurity Strategy*. Washington, DC, March 2023.

[67] Daniel R. Thomas, Sergio Pastrana, Alice Hutchings, Richard Clayton, and Alastair R. Beresford. Ethical Issues in Research Using Datasets of Illicit Origin. In *Proceedings of the 2017 Internet Measurement Conference*, IMC '17, pages 445–462, New York, NY, USA, 2017. doi: 10.1145/3131365.3131389.

[68] Kurt Thomas, Frank Li, Ali Zand, Jacob Barrett, Juri Ranieri, Luca Invernizzi, Yarik Markov, Oxana Comanescu, Vijay Eranti, Angelika Moscicki, Daniel Margolis, Vern Paxson, and Elie Bursztein. Data Breaches, Phishing, or Malware? Understanding the Risks of Stolen Credentials. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1421–1434, Dallas, Texas, USA, 2017.

[69] Dana Turjeman and Fred M Feinberg. When The Data Are Out: Measuring Behavioral Changes Following a Data Dreach. *Available at SSRN 3427254*, 2019. doi: https://dx.doi.org/10.2139/ssrn.3427254.

[70] Josephine Wolff and William Lehr. Ex-Post Mitigation Strategies for Breaches of Non-Financial Data. In *The 44th Research Conference on Communication, Information and Internet Policy*, Arlington, VA, USA, 2016. doi: http://dx.doi.org/10.2139/ssrn.2756842.

[71] Daniel W. Woods and Rainer Böhme. SoK: Quantifying Cyber Risk. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 909–926, San Francisco, CA, USA, May 2021. doi: https://doi.ieeecomputersociety.org/10.1109/SP40001.2021.00053.

[72] Yixin Zou, Abraham H. Mhaidli, Austin McCall, and Florian Schaub. "I've Got Nothing to Lose": Consumers' Risk Perceptions and Protective Actions after the Equifax Data Breach. In *Proceedings of the Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 197–216, Baltimore, MD, USA, 2018.

[73] Yixin Zou, Kevin Roundy, Acar Tamersoy, Saurabh Shintre, Johann Roturier, and Florian Schaub. Examining the Adoption and Abandonment of Security, Privacy, and Identity Theft Protection Practices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–15, New York, NY, USA, 2020. doi: 10.1145/3313831.3376570.

A E-Mail Invite

Subject line: Follow-up on the Ledger data breach

Hello,

I am [*first author's name*], a security researcher at the [*researchers' affiliation*]. I am contacting you regarding the data breach of Ledger's customer database last year. I sincerely regret that personal information had been compromised and leaked to the public. Our team of security researchers would like to follow up on this incident and estimate the scope of effects and potential harm it caused to the victims. With this initiative, we strive to collect valuable first-hand insights from you. This will help us to make recommendations for better protection and response in order to avoid similar cases in the future. If you support this research initiative, I would like to invite you to participate in our **anonymous online survey** available at [*link to the static webpage*].

Please note that your participation in this study is **voluntary**. You are free to not participate at all or to withdraw from the survey at any time. Completing the survey will take about **25 minutes** of your time. Your anonymous responses will be analyzed and used by the members of our team for research purposes only. The results will be published in aggregate form in academic journals and conferences.

This research study has been approved by the Research Board for Ethical Issues of the [*researchers' affiliation*] under the principles of freedom of research and beneficial societal impact. We, the researchers in charge, aim to collect opinions, experiences, and sentiments of the users affected by the data breach in order to derive policy recommendations and managerial implications for firms offering secure storage devices. The data controller of any personal data related to this study is the [*researchers' affiliation*]. Processing takes place under the General Data Protection Regulation (GDPR) of the European Union.

The only piece of personal data processed is your **e-mail address**. The processing is justified by the public interest of the conducted research (Article 6, p. 1(e) GDPR). Your e-mail address will be processed until the end of the collection of 500 reliable response sets or until the expiration of 6 months since the start of the research study on November, 8th 2021. Your e-mail address is processed only for the purpose of inviting you to participate in this study and will not be shared with other parties. It will be permanently deleted after the end of the term of processing (on May 8th, 2022) or immediately after your written request to [*contact e-mail address*].

We take the following security measures to protect your e-mail address:

- Your e-mail address is stored in an encrypted form on an external USB flash drive, which is kept in a secure physical vault. The access to the vault and the drive is limited to authorized staff of the research group.
- This invitation is sent from an e-mail account purposefully created and used for this study only. The account as well as all the entire communication will be permanently deleted after completion of the data collection phase.

In matters related to data processing, you may also contact the data protection coordinator of the [*researchers' affiliation*] at [*DPO e-mail address*]. You have the right to lodge a complaint with the data protection supervisory authority if you believe that the processing of your e-mail address for the purpose of this research study infringes the provisions of the GDPR.

The anonymous online survey is accessible at [*link to the static webpage*]. I sincerely thank you for your collaboration.

Sincerely,
[*first author's name*]
[*researchers' affiliation*]

B Static Webpage

Welcome to Our Survey on the Ledger Data Breach

Thank you for your interest in our study!

Access to the Questionnaire

Please click here to access the anonymous online survey. Completing the survey will take about 25 minutes of your time.

The questionnaire is hosted by Qualtrics, a professional company offering survey tools to researchers. The company is bound by law and contracts to follow our high data protection standards. We are using this static webpage, hosted by the researchers, between the invitation link in the e-mail you received and the actual questionnaire to reassure all respondents that the links are not personalized. **This research respects your privacy**.

Contacts

If you have any questions or concerns about the project, please do not hesitate to contact us at [*contact e-mail address*].

In matters related to data processing, you may also contact the data protection coordinator of the [*researchers' affiliation*] at [*DPO e-mail address*]. You also have the right to lodge a complaint with the data protection supervisory authority if you believe that the processing of your e-mail address for the purpose of this research study infringes the provisions of the GDPR.

C Questionnaire[8]

Welcome and thank you for supporting our research initiative!

[8] The complete questionnaire is available in an arXiv version of this paper.

This research respects your privacy. The survey does not collect personal data in the form of your name or contact details and we ask you not to provide this information in your open text responses. The survey is intended to be anonymous. If you disclose information concerning harm to yourself or another person within this survey, the researchers will not be able to take any action.

What is the purpose of this research? We use this questionnaire to learn more about the effects and potential harm resulting from the Ledger data breach in 2020, which may have impacted you. We would love to hear your personal story of this incident. These insights will help us estimate the aggregate consequences of this data breach as well as understand your attitudes, security behaviors, and coping strategies.

Who are we? We are researchers of the [*researchers' affiliation*]. We conduct this research independently of the Ledger company. This research is funded from the research budget of the [*researchers' affiliation*].

Your participation is voluntary. You do not have to participate in this survey, and you can leave it at any point. Completing the survey will take about 25 minutes of your time.

Your responses are anonymous. We assure that we will remove any information which would make you identifiable. Your anonymous responses will be analyzed and reported in aggregate form and will be used for research purposes only.

There are potential benefits from participating. You have the opportunity to share your story and thoughts about the data breach with us. Your valuable input will contribute to advance scientific research, improve companies' security practices, and inform policy making in this domain.

There are potential risks from participating. Responding to our questions may remind you of your experience as a victim of crime and any pain you may have experienced.

How to contact us? If you have any questions or would like to get further information about this study, you may contact the principal investigator [*first author's name*] [*contact e-mail address*].

Consent. By clicking "Yes", you confirm to be over 18 years of age and consent to us collecting information about your attitudes, experiences, and demographic information. Detailed information regarding this project and your participation has been explained to you.