Consumer Beware!
Exploring Data Brokers' CCPA Compliance

Elina van Kempen, Isita Bagayatkar, Pavel Frolikov, Chloe Georgiou, and Gene Tsudik
University of California, Irvine
{evankemp, ibagayat, pavel, cgeorgio, gtsudik}@uci.edu

***Abstract*—Data brokers collect and sell the personal information of millions of individuals, often without their knowledge or consent. The California Consumer Privacy Act (CCPA) grants consumers the legal right to request access to, or deletion of, their data. To facilitate these requests, California maintains an official registry of data brokers. However, the extent to which these entities comply with the law is unclear.**

This paper presents the first large-scale, systematic study of CCPA compliance of all 543 officially registered data brokers. Data access requests were manually submitted to each broker, followed by in-depth analyses of their responses (or lack thereof). Above 40% failed to respond at all, in an apparent violation of the CCPA. Data brokers that responded requested personal information as part of their identity verification process, including details they had not previously collected. Paradoxically, this means that exercising one's privacy rights under CCPA introduces new privacy risks.

Our findings reveal rampant non-compliance and lack of standardization of the data access request process. These issues highlight an urgent need for stronger enforcement, clearer guidelines, and standardized, periodic compliance checks to enhance consumers' privacy protections and improve data broker accountability.

***Index Terms*—CCPA, privacy compliance, data brokers, privacy law, data access rights, privacy risks**

1. Introduction

Data brokers ($\mathcal{DBR}$s) operate largely hidden from public view: collecting, aggregating, and selling personal information (PI) of consumers without their knowledge or consent. These entities systematically harvest data from various sources: public records, online activities, social media profiles, and even other $\mathcal{DBR}$s. They routinely analyze this collected data to determine sensitive information such as purchasing behavior, financial status, and health conditions. Then, $\mathcal{DBR}$s monetize this information by selling it to various third parties, including companies, government agencies, and individuals. These sales are often completed without informing the consumer. One notorious and all-too-familiar class of $\mathcal{DBR}$s corresponds to so-called "people search" websites: online platforms that aggregate and market individual consumers' PI. Often, some PI is available for free, and additional information is available for gradually increasing prices. Usually, there is a prominent pitch to pay a fee in order to see detailed PI, with variable pricing schemes. Thus, any malicious actor can gain access to, and misuse, consumer PI to mount identity theft, fraud, or phishing attacks.

$\mathcal{DBR}$s operate without current relationships with consumers whose data they process, unlike traditional data aggregators (such as credit reporting agencies) with which consumers knowingly share their data to obtain presumably useful services. This fundamental difference results in consumers being unaware that their PI is being collected, analyzed, and made available for sale. Actual implications of such unmitigated access to consumers' PI remain poorly understood by the general public.

To remedy the situation, several jurisdictions enacted data protection laws, e.g., General Data Protection Regulation (GDPR) [1] in the European Union, California Consumer Privacy Act (CCPA) [2] in California, and Lei Geral de Proteção de Dados (LGPD) [3] in Brazil. Though specific provisions differ among them, these laws aim to give individuals greater control over their PI. Notably, they grant consumers the rights to access, correct, and delete PI held by either businesses or other organizations. Under CCPA, consumers exercise these rights by making a *Verifiable Consumer Request (VCR)*:

> "a request made by a consumer ..., and that the business can verify, using commercially reasonable methods, ... to be the consumer about whom the business has collected personal information" (1798.140(ak)) [2].

The California Privacy Protection Agency (CPPA) specifically targets $\mathcal{DBR}$s. Besides CCPA compliance, $\mathcal{DBR}$s must abide by the Data Broker Registration law, which requires each $\mathcal{DBR}$ to register annually with the CPPA. The resulting registry is made publicly available on the CPPA website [4].

Composing and submitting a VCR is burdensome for consumers because there is no standard process for doing so. Each $\mathcal{DBR}$ has its own method for consumers to submit a VCR, such as by filling out a form on the website, sending an email, making a phone-call, or even having to go through a multi-step process. The necessary steps are typically (supposed to be) contained within a $\mathcal{DBR}$s' privacy policy.

Identity verification adds another layer of complexity. A $\mathcal{DBR}$ must verify the requesting consumer's identity before releasing the data in order to prevent data breaches. However, this verification process is not standardized and is taxing for the average consumer.

Accessing consumer data is especially challenging when interacting with $\mathcal{DBR}$s. Unlike other business entities, which users expect to collect PI, there is no way of knowing beforehand whether a $\mathcal{DBR}$ collected PI of a particular consumer. Thus, the consumer is forced to *blindly* make VCRs to a long list of $\mathcal{DBR}$s. The identity verification process is also somewhat Kafkaesque:

> How can a consumer prove their identity to a $\mathcal{DBR}$ that may, or may *not*, have their PI?

Throughout this paper, we distinguish between: (1) Personal Information (**PI**), i.e., any information a $\mathcal{DBR}$ may have about a consumer, and (2) Personally Identifiable Information (**PII**), which is a subset of PI that specifically identifies a consumer, e.g., name, address, or Social Security Number (SSN). This distinction is important for understanding privacy implications of CCPA enforcement. For example, the consumer's device model is PI, and not PII, because many consumers share the same device type; however, a driver's license number is PII since it uniquely identifies an individual. As shown later, this requirement to provide PII for identity verification creates an unintended privacy paradox when exercising one's CCPA rights.

1.1. Research Questions

Motivated by curiosity about the current state of CCPA compliance and the practical challenges consumers face when exercising their privacy rights, this work seeks to answer the following three research questions:

RQ1. What challenges are encountered in the VCR submission process, especially in the identity verification step?

VCR submission is not a streamlined process. $\mathcal{DBR}$s can simplify it by adopting user-friendly privacy policies and employing reasonable VCR submission methods. We measured the time needed for the researcher to find each $\mathcal{DBR}$'s contact information on their privacy policy, as well as the time needed to submit a VCR, along with the amount of PII needed for the $\mathcal{DBR}$ to verify the researcher's identity.

RQ2. To what extent do $\mathcal{DBR}$s comply with CCPA after VCR submission?

The goal is to discover what fraction of registered $\mathcal{DBR}$s comply, fully or in part, with CCPA requirements described in Section 3. After measuring VCR response rates and response timelines, we consider the factors that correlate with CCPA compliance.

RQ3. What kind of data do $\mathcal{DBR}$s collect and how is it shared with the consumer following a VCR?

After receiving the PI from $\mathcal{DBR}$s, we classify both the type of PI collected and its accuracy. We consider whether this data is sent securely to consumers, and whether it is provided in a standard and usable format (e.g., TXT, CSV, PDF, or JSON) as required by CCPA, rather than in some proprietary, obscure, or otherwise unparsable format. That PI must also be provided "...in a format that is easily understandable for the average consumer" (1798.130 (a)(3)(B)(iii)) [2].

1.2. Study Overview

This paper reports on a comprehensive and systematic study of the $\mathcal{DBR}$ ecosystem. The study involved all 543 $\mathcal{DBR}$s duly registered in California, listed under the California Privacy Protection Agency. We examined six key aspects of the VCR submission process: (1) consumer burden of composing VCRs, (2) variability in identity verification, (3) response time fluctuations, (4) quality of responses (i.e., content), (5) what actual PI is being collected, and (6) privacy issues, if any, that arise in the process of requesting PI.

To begin the study, we submitted VCRs to all 543 $\mathcal{DBR}$s according to their individual guidelines, if those were available. We then measured their response timeline and analyzed common practices. We found that only 57% responded to VCRs, meaning that 43% are blatantly non-compliant.

We found that the majority of $\mathcal{DBR}$s did not have any PI about the researcher who submitted VCRs. However, we still had to send PII to every $\mathcal{DBR}$ as part of their identity verification process.

Although some prior work studied CCPA and/or GDPR compliance, its focus was on entities with which the consumer has a direct relationship [5], [6], [7], [8], [9], [10], [11], [12], [13]. Furthermore, previous research on $\mathcal{DBR}$s examined only a small subset of $\mathcal{DBR}$s: people search websites [14]. Other types of $\mathcal{DBR}$s have not been investigated. A previous study evaluated the usability of data access and removal of 20 such websites [14], offering a rather narrow view of the $\mathcal{DBR}$ ecosystem.

1.3. Organization

Section 2 presents an overview of prior work, and Section 3 summarizes relevant CCPA sections. The methodology is described in Section 4, including ethical considerations in 4.2. Section 5 follows with results. Finally, privacy issues and study outcomes are described in Section 6.

2. Related Work

Related work is summarized in Table 1, which specifies the number of entities to which VCRs could be submitted, followed by the total number of entities involved in each study, in parentheses. Note that, for the sake of simplicity and uniformity, we use the term **VCR** to refer to privacy rights requests, and the term **consumer** to refer to an individual whose PI is collected, regardless of the specific legal framework in the context of which they occur.

Most prior work on compliance with privacy laws (in terms of data access VCRs) focused on studying popular

websites or apps [5], [6], [7], [8], [9], [10], [11], [12], [13]. Some prior results [8], [11], [12], [13] ensured that targeted entities already had data about the specific consumer. To do so, they picked popular websites or applications, created accounts whenever possible, and used the services for a certain time. For example, [5], [6], [10], focused on selected entities that had already collected data about the researcher(s). Also, some prior work [5], [6] also examined vulnerabilities in the data request process and tried to access consumer PI through impersonation.

Borem et al. [15] investigated data export procedures, in response to data access requests, for 5 major service providers: Amazon, Facebook, Google, Spotify, and Uber. A total of 33 participants requested data; each participant had $\geq$ 2 years of use history with the relevant entity. Participants found that data exports were confusing and overwhelming, due to large amounts of information and a lack of explanation for the types of data collected and/or the reasons for collection.

Adhatarao et al. [7] reported on attempts to submit VCRs to websites they visited using only an IP address for identification, which was unsuccessful for all attempts.

Boniface et al. [9] and Urban et al. [16] evaluated submission of VCRs to well-known third-party tracking services. The former did not report on the results of VCRs, instead focusing on security aspects of the requester's identity verification and the ease of making a request. The latter used stored cookies created by tracking services for identification: only 58% of tracking services replied, whether positively or negatively.

Very recently, He et al. [17] analyzed pricing, coverage, and efficacy of 10 PII removal services, all of which claim to handle the VCR process with data brokers on behalf of consumers. Results showed that PII removal services differed greatly in terms of which data brokers they cover. They also generally failed to remove the majority of users' records from the data brokers.

Take et al. [14] is the most similar to our work. It reported on a study testing the compliance of 20 people-search websites, which is a type of $\mathcal{DBR}$, e.g., Intellus, Truthfinder, and Whitepages. VCRs were made for each such website under both GDPR and CCPA, depending on the researchers' residence. Response data was compared to that received after paying for user reports. However, results did not specify whether requests made under one law yielded better results than those made under another. Furthermore, the timeline of responses was not provided. Also, we believe that significant results or trends about the $\mathcal{DBR}$ landscape cannot be derived from the qualitative analysis of only 20 people-search websites.

In contrast, the study described in this paper systematically and comprehensively evaluates the behavior of **all** registered $\mathcal{DBR}$s after submission of CCPA data access VCRs.

TABLE 1: Comparison of this work to relevant prior results.

	# Entities	Law	Resp. rate	Year
$\mathcal{DBR}$s				
this	454 (543)	CCPA	57%	2024-2025
People Search Websites				
Take et al. [14]	20	CCPA, GDPR	82-100%	2024
Tracking services				
Boniface et al. [9]	25 (30)	GDPR	-	2018-2019
Urban et al. [16]	36 (39)	GDPR	58%	2018
Common websites				
Martino et al. [6]	40	GDPR	93%	2021
Adhatarao et al. [7]	109 (124)	GDPR	57%	2021
Bufalier et al. [8]	317 (341)	GDPR	71%	2020
Martino et al. [5]	55	GDPR	93%	2019
Boniface et al. [9]	27 (50)	GDPR	-	2018-2019
Wong et al. [10]	229 (230)	GDPR	75%	2018
Herrmann et al. [11]	119 (120)	GDPR	77%	2016
Smartphone apps				
Samarin et al. [13]	109 (160)	CCPA	81%	2023
Kroger at al. [12]	216-224 (225)	GDPR	77-83%	2015-2019
Herrmann et al. [11]	144 (150)	GDPR	55%	2016

3. Background: CCPA and Data Broker Registration

The General Data Protection Regulation (GDPR), enacted in 2016 by the European Union, established comprehensive privacy regulations and rights for European consumers [1]. It also inspired new privacy regulations internationally. In particular, the California Consumer Privacy Act (CCPA), voted on in 2018 and amended in 2020 by the California Privacy Rights Act (CPRA), grants California residents certain rights over the data collected about them by businesses [2]. These rights are commonly known as *the rights to know, correct, or delete PI collected by a business*.

Personal information. CCPA defines PI as information that:

> "identifies, relates to, describes, ..., or could reasonably be linked, directly or indirectly, with a particular consumer or household." (1798.140(v)) [2]

This includes identifiers (names or usernames) and other consumer data, such as addresses, biometric information, geolocation information, and network activity. It excludes publicly available information, i.e., public government records and consumer information made public by the consumer or another person, assuming no restrictions placed by the consumer.

In addition, CCPA distinguishes *sensitive* PI, which includes an individual's SSN, driver's license or passport number, genetic information, racial or ethnic information, and account credentials. (1798.140(ae)) [2].

$\mathcal{DBR}$s. Businesses must comply with CCPA if they satisfy one of the following (1798.140(d)) [2]:

1) Have a gross annual revenue over $25.625 million.
2) Handle over 100,000 California residents' PI.
3) Get more than 50% of their revenue from the sale of California residents' PI.

The Data Broker Registration law establishes further registration requirements for $\mathcal{DBR}$s [18]. A $\mathcal{DBR}$ is defined as:

"a business that knowingly collects and sells to third parties the personal information of a consumer with whom the business does not have a direct relationship" (1798.99.80(c)) [18].

As described in Section 1, unlike other businesses that collect data of their customers (i.e., consumers who previously used their services), $\mathcal{DBR}$s collect data from individuals who have never used their services. $\mathcal{DBR}$s specialize in collecting, analyzing, and monetizing consumer data by selling it to third parties.

The Data Broker Registration law requires $\mathcal{DBR}$s to register annually[1] with the California Privacy Protection Agency (1798.99.82) [18]. The registration fee is currently $6,600, with a 2.99% additional electronic payment fee (Tit. 11 Div. 6 §7600) [19]. $\mathcal{DBR}$s that fail to register may face a fine of $200 per day of non-compliance. The law also requires $\mathcal{DBR}$s to undergo regular third-party audits starting in 2028 (1798.99.86(c)) [18].

VCR processing. CCPA encourages verification of a consumer's identity using an existing password-protected account, although it prohibits businesses from requiring a consumer to create an account in order to submit a VCR (1798.130(a)(2)(A)) [2].

CCPA provides guidelines for identity verification of non-account holders (Tit. 11 Div. 6 §7062) [19], and requires a business to verify the identity of the consumer to a:

- "reasonable degree of certainty", when a user requests to know categories of PI currently possessed by the business, by matching at least two "reliable" data points provided by the consumer with data points possessed by the business.
- "reasonably high degree of certainty", when a user requests to know specific pieces of PI, by matching at least three "reliable" data points provided by the consumer with data points possessed by the business, along with a signed declaration under penalty of perjury.

CCPA states that a business may not use the PII provided during the identity verification process for any other purposes (1798.130(a)(7)) [2].

Once a business verifies the consumer's identity, it must: (1) confirm receipt of the request within 10 business days (Tit. 11 Div. 6 §7021(a)) [19], and (2) answer the request within 45 calendar days. If needed, a business may extend its response time by an additional 45 calendar days, as long as it notifies the consumer (1798.130(a)(2)(A)) [2]. CCPA requires businesses to deliver the information by mail or electronically, "in a readily usable format that allows the consumer to transmit this information from one entity to another entity without hindrance" (1798.130(a)(2)(A)) [2].

3.1. $\mathcal{DBR}$s Location

Almost all $\mathcal{DBR}$s are registered to an address within the United States, with only 28 $\mathcal{DBR}$s registered abroad. Table 2 specifies the geographical distribution of $\mathcal{DBR}$s to which VCRs could be submitted, by country.

TABLE 2: $\mathcal{DBR}$s by country of registration.

Country	$\mathcal{DBR}$ Count
United States	515
United Kingdom	10
Canada	5
Germany, Denmark	2 each
Malaysia, India, Poland, Sweden, Israel, New Zealand, Namibia, Italy, United Arab Emirates	1 each
Total	543

Within the United States, $\mathcal{DBR}$s are registered in 40 states. A large number are registered in three states: California, New York, and Florida, with 103, 74, and 46 registered $\mathcal{DBR}$s, respectively. Figure 1 shows a map of all $\mathcal{DBR}$s in the United States. Since no $\mathcal{DBR}$s are registered in Alaska or Hawaii, these states are not shown on the map.

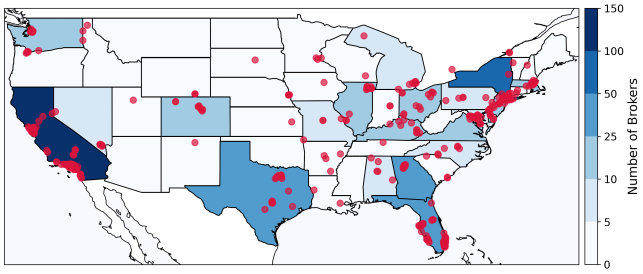


Figure 1: Map of US $\mathcal{DBR}$s.

4. Methodology

Recall that our goal is to assess $\mathcal{DBR}$s' CCPA compliance. To this end, we submitted VCRs[2] for PI collected by all registered $\mathcal{DBR}$s. We chose not to submit data deletion requests, since by doing so we would learn less information about $\mathcal{DBR}$s' data collection and response processes. In other words, a $\mathcal{DBR}$ could simply reply that relevant PI found in their database will be deleted, whether or not such data actually exists, with no proof of deletion. Submitting a VCR guarantees a database lookup by the (compliant) $\mathcal{DBR}$, which must then reply with collected PI or with a statement that it has none.

4.1. Study Design

We investigated all 543 registered $\mathcal{DBR}$s in the California Data Broker Registry during the 2024 year [20] and submitted VCRs to exercise our "right to know/access". A single co-author (denoted "researcher" hereafter), California resident since 2021, submitted all VCRs to ensure consistency and privacy.

1. The California $\mathcal{DBR}$ registry is available at https://cppa.ca.gov/data_broker_registry/

2. From this point forward, VCR refers specifically to data access requests.

Submitting VCRs. Each $\mathcal{DBR}$ in the registry provided a URL to their website's homepage. After finding a privacy policy, the researcher located each $\mathcal{DBR}$'s instructions for submitting VCRs. The researcher used each $\mathcal{DBR}$'s preferred VCR submission method to provide $\mathcal{DBR}$s with the best circumstances and maximize the likelihood of getting a response. In situations where the $\mathcal{DBR}$-provided forms were non-functional, the researcher used an alternative email contact method listed in the privacy policy. A standard (consistent) email template was used across all emailed VCRs; see in Appendix A.

In a few cases, when neither a form-based nor an email-based contact method worked, the only option was making a phone-call. Although CCPA requires $\mathcal{DBR}$s to provide at least a toll-free telephone number for submitting VCRs (1798.130(a)(1)(A)) [2], doing it by phone is not optimal since it leaves no proof or record of submission. Additionally, callbacks are easy to miss. Whenever a phone-call led to voicemail, the researcher left a message stating that they wanted to submit a VCR, along with their phone number for $\mathcal{DBR}$s to call back.

To accurately track all VCRs and responses, a dedicated email address was created solely for the VCR submission purpose. If $\mathcal{DBR}$s requested alternative email addresses (e.g., a business email), the researcher provided them.

The researcher duly provided all PII requested by $\mathcal{DBR}$s for identity verification purposes. If a copy of a government-issued ID was needed, the researcher sent a redacted version unless a full copy was specifically requested.

Metrics. We measured VCR submission time for each $\mathcal{DBR}$: the timer started when the researcher landed on the $\mathcal{DBR}$'s homepage and stopped when a form was submitted, an email was sent, or a phone-call ended. Thus, we measured the time needed to: (1) locate the privacy policy on the $\mathcal{DBR}$'s website, (2) identify the VCR submission method within that policy, and (3) perform the actual VCR submission. The data collection period lasted roughly 7 months total, from late 2024 to early 2025.

We recorded PII elements requested by each $\mathcal{DBR}$ for VCR submission and identity verification. To evaluate CCPA compliance, we tracked the response rate, timeline of responses, and the content of responses. We recorded the dates when: (1) the original request was sent, (2) the request receipt acknowledgment was received, and (3) the response (if any) was received. If a $\mathcal{DBR}$ requested additional PII, we recorded the time needed by the researcher to provide this information. In order to precisely and fairly assess response timeline, we exclude time taken by the researcher to submit the requested additional PII from our analysis. Thus, when a $\mathcal{DBR}$ requests additional information, the $\mathcal{DBR}$ is not penalized for the time taken by the researcher to reply to the $\mathcal{DBR}$.

Exclusions (Data Cleaning). 89 out of 543 registered $\mathcal{DBR}$s were excluded from the full study:

Duplicates. 50 corresponded to duplicate entries, i.e., different registry entries leading to the same website, email address, or form. We only submitted one request for each duplicate registration.

Noncompliant. 8 $\mathcal{DBR}$s' websites were completely inaccessible during the study period. 7 $\mathcal{DBR}$s lacked any privacy policy or contact information to submit VCRs. 11 $\mathcal{DBR}$s for which email-based VCRs could not be delivered. 2 $\mathcal{DBR}$s only provided phone contacts, and there was no way to leave a voicemail; in each case, the researcher hung up after 5 minutes of phone silence. Finally, 1 $\mathcal{DBR}$ required notarization, although CCPA regulations (Tit. 11 Div. 6 §7060(e)) [19] state that: "...a business may not require a consumer to provide a notarized affidavit to verify their identity unless the business compensates the consumer for the cost of notarization" [19].

Lack of Security. 2 $\mathcal{DBR}$s with non-HTTPS websites and forms were excluded for security and privacy reasons.

Unable to complete VCR. 4 $\mathcal{DBR}$s required a website cookie ID, which was not present on the researcher's device(s). 3 $\mathcal{DBR}$s requested information that the researcher did not have, e.g., organizational position, website URL, or medical professional status.

"No longer a $\mathcal{DBR}$" (but registered anyway). 1 $\mathcal{DBR}$ announced on its main web page that it had exited consumer data sales.

After accounting for all exclusions, the analysis focused on the remaining 454 $\mathcal{DBR}$s, representing approximately 84% of the total.

4.2. Ethical considerations

The authors' Institutional Review Board officially declared that this study constitutes non-human subject research. The lead researcher (using their own identity to compose VCRs) did so voluntarily, motivated primarily by the research goals. To preserve their privacy, only that researcher (who submitted all VCRs) had access to $\mathcal{DBR}$ responses. The researcher could choose to withdraw from the study at any time and not submit VCRs to certain websites if they felt uncomfortable. In particular, as mentioned above, the researcher chose not to make VCRs to $\mathcal{DBR}$s that had non-HTTPS websites out of concern for their privacy.

Also, recall that we observed that some $\mathcal{DBR}$s are registered more than once, i.e., multiple $\mathcal{DBR}$s' registrations led to the same VCR submission form or email. We avoided contacting them multiple times to be respectful of their resources. If we accidentally contacted a $\mathcal{DBR}$ twice, we only considered the first VCR for all further analysis.

Finally, besides the scientific goal of this study, the researcher had genuine and legitimate interest in learning what PI about them was collected by $\mathcal{DBR}$s.

5. Results

We now present quantitative results addressing **RQ1**, **RQ2**, and **RQ3**. Note that RQ1 is broken down into Sections 5.1 and 5.2.

5.1. RQ1: Submitting VCRs

As detailed in Section 4.1, we submitted VCRs by form, email, or phone-call, as outlined in the $\mathcal{DBR}$s' privacy policy.

VCR Submission Methods. Figure 2 depicts various submission methods encountered and their proportions. Since there is no CCPA-standardized method to submit VCRs, each $\mathcal{DBR}$ chooses what it prefers, which complicates VCR submissions for consumers. The researcher managed to submit most VCRs (92.5%) directly by email or by form.

Forms, denoted form-VCRs, were the most common method. While seemingly straightforward (just complete and submit), forms presented their own challenges mainly because they are not standardized, with each DBR using a different interface and requiring different information. The second common VCR method was email, denoted email-VCR.

We note that $\mathcal{DBR}$s can purchase CCPA compliance services from privacy-as-a-service companies. OneTrust [21] was the most popular service, used by 56 $\mathcal{DBR}$s: about 25% of form-VCRs were "powered by OneTrust". Curiously, despite using OneTrust, form fields varied quite a bit among the 56 $\mathcal{DBR}$s' form-VCRs.

For 7 $\mathcal{DBR}$s, VCRs had to be made by phone, denoted phone-VCRs. This corresponds to cases where $\mathcal{DBR}$s only posted their physical address and phone number, leaving no way to contact them electronically.

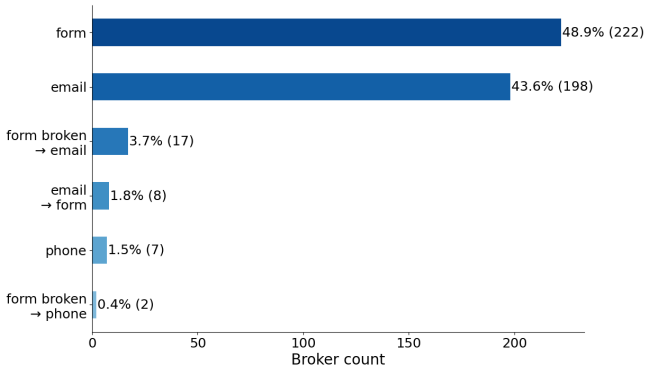


Figure 2: VCR submission methods.

27 submissions involved multiple steps. Upon submitting email-VCRs to 8 $\mathcal{DBR}$s, the researcher received links to resubmit through an online form. This occurred either because the $\mathcal{DBR}$ preferred email-initiated requests or because the form was not found in the $\mathcal{DBR}$'s privacy policy. Notably, certain $\mathcal{DBR}$s combine their "request to know" and "request to opt-out/delete" forms into one. This practice leads to confusion because the combined forms are frequently labeled only as "request to opt-out/delete". We believe that most consumers would not open a form titled "delete my data" when attempting to access their data.

19 forms could not be submitted, primarily due to broken links in the privacy policy. When such forms were encountered, the researcher emailed or called the $\mathcal{DBR}$. Some forms failed to submit due to an "invalid CAPTCHA", even though no CAPTCHA was displayed on the webpage. In another irritating case, the form required a frequent shopper ID (confirmed to be a string of digits by the $\mathcal{DBR}$), yet the field only accepted email address formats.

Finally, in a particularly time-consuming request case, the researcher had to call a $\mathcal{DBR}$ after being unable to submit the request form, since no email contact was specified in that $\mathcal{DBR}$'s CCPA privacy policy. After leaving their contact information by voicemail, the researcher received an email from the $\mathcal{DBR}$ asking to fill in a form (the link to which was broken) or to provide additional information by email. Despite replying with the required details and reporting on the form's malfunction, no further response was received.

VCR Submission Time. In total, 9 hours and 57 minutes were spent submitting VCRs to all 454 $\mathcal{DBR}$s. This corresponds to the average of 79 seconds to find the VCR method and to submit a VCR. Figure 3 shows the time to submit email-VCRs and form-VCRs. Phone and multi-step procedures are excluded from the Figure.

It took, on average, longer to submit form-VCRs than email-VCRs, with averages of 82 and 66 seconds, respectively. Shapiro-Wilk tests reject the null-hypotheses that email-VCR and form-VCR submission times are normally distributed (p-value < 0.001), and a permutation test suggests a statistically significant difference in submission times (p-value < 0.005). Note that, while forms were filled out manually, emails were copy-pasted from a template. Submitting OneTrust forms took roughly as long as other forms.

Submitting phone-VCRs was quite time-consuming. An average phone-call lasted 4 minutes and 37 seconds. It typically included talking with a company representative or leaving a voice message if no one was available. We were surprised by phone representatives being totally unaware of CCPA and confused by the nature of our requests. This was troubling since we made sure to use the phone number explicitly specified by $\mathcal{DBR}$s in their privacy policy.

KEY TAKEAWAYS - RQ1 PART 1. Submitting VCRs is a complex process. While email-VCRs are one of the most commonly used methods, their open-ended nature places a burden on consumers, who must properly word their emails to ensure that $\mathcal{DBR}$s do not simply dismiss their requests, and verify that the emailed VCR does not bounce. In comparison, form-VCRs are more straightforward, though more time-consuming. Multi-step processes incur an unjustified burden since consumers must keep track of (and complete) several steps, which are usually due to $\mathcal{DBR}$s' inadequate

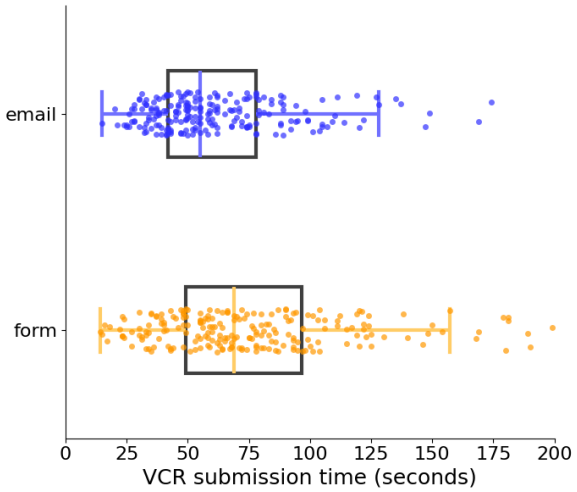


Figure 3: Time needed to submit VCRs.

privacy policies or outdated contact information. Finally, phone-VCRs are plain unhelpful. Even when the calls were answered, employees demonstrated an alarming lack of understanding about the nature of the inquiry.

5.2. RQ1: Identity Verification Process

Recall that $\mathcal{DBR}$s have to verify the consumer identity via VCRs before granting them access to their PI. To this end, they ask the consumer to provide a set of PII, which is presumably matched with previously collected PII.

Type of PII Requested We first categorize different types of PII requested by $\mathcal{DBR}$s to perform identity verification. Table 3 shows, for each PII category, the number of $\mathcal{DBR}$s that requested it.

Email address and name were requested by over 95% of $\mathcal{DBR}$s. Next, home address and phone number were requested by 42% and 23%, respectively. $\mathcal{DBR}$s likely use home address to confirm California residency and CCPA jurisdiction. Sometimes, only partial address information was required, e.g., city, zip code, or state. 21 $\mathcal{DBR}$s asked for professional information, e.g., the LinkedIn URL, and/or the name of the employer.

32 $\mathcal{DBR}$s asked for PII that is difficult to retrieve for the average consumer, such as cookie values, Mobile Advertising Identifier (MAID), or IP addresses. To obtain cookie information on Google Chrome, one needs to `Inspect` the webpage, click on `Application`, and find relevant cookie values under `Storage`. On Android devices, user MAID is accessible within the phone's settings. However, for iPhone users, Apple does not directly provide access to MAID. Instead, iOS users must download a third-party app to view their MAID, which is a bit disconcerting.

Five $\mathcal{DBR}$s requested unusual information: marital status, gender, a list of recently visited locations, and a grocery store frequent shopper ID. Another five $\mathcal{DBR}$s presented a multiple-choice questionnaire asking the consumer to select their PI from a list. A screenshot of a sample questionnaire is shown in Figure 4.

TABLE 3: Types of PII elements required for VCR submissions.

PII type	Count (out of 454)
Email	444
Name	431
Address	Partial: 70, Full: 122
Phone number	103
Date of birth	Partial: 8, Full: 33
MAID	26
Official ID	Redacted: 13, Full: 4
SSN	Last 4 digits: 9, Full: 7
Signature/Signed affidavit	15
LinkedIn URL	11
Employer	11
IP address	8
Multiple-Choice Questionnaire	5
Selfie	5
Cookie	5
Utility bill (redacted)	6
Previous addresses	4
Driver's license number	2
Recently visited locations	2
Grocery store frequent shopper ID	1
Gender	1
Social media usernames	1
Marital status	1

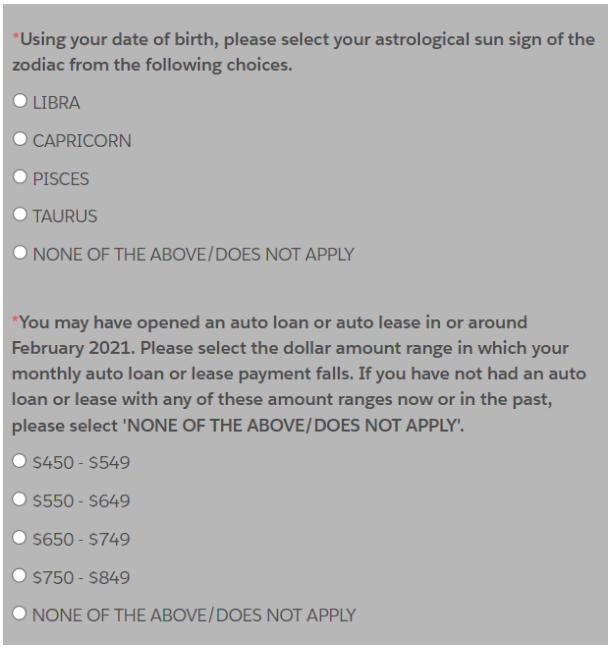


Figure 4: Questions from a questionnaire in a form-VCR.

Finally, 45 $\mathcal{DBR}$s asked for sensitive PII, as defined by CCPA. This included: full or partial SSN, full or partially redacted copy of a government ID, driver's license number, and biometrics, e.g., a signature or a selfie. This includes 4 $\mathcal{DBR}$s that used third-party systems requiring the consumer to take a live selfie along with a picture of a government ID.

Number of PII Elements in the VCR. The number of PII elements requested for identity verification varied depending on the VCR submission method. Since phone-VCRs were a small minority and mostly unsuccessful, a minimal amount of PII was shared through them, usually only phone number and name. Most email-VCRs require 2 PII elements, while about half of form-VCRs require 4 or more. A Mann-Whitney U test indicates that form-VCRs requested significantly more PII (p-value < 0.001) than email-VCRs. OneTrust forms did not ask for more PII elements than other $\mathcal{DBR}$ forms.

Submitting email-VCRs involved sharing the researcher's email address, from which the VCR was sent, and full name, signed at the end of the email. Seven $\mathcal{DBR}$s specified (in their privacy policy) additional PII elements to include in the email, and 23 requested additional information after the initial email, in order to continue with identity verification. In contrast, form-VCRs typically required all PII elements up front, at the initial submission time.

Additional Verification. Besides providing PII, a consumer needs to take additional steps to successfully submit form-VCRs. For instance, 101 $\mathcal{DBR}$s required confirming access to the supplied email address or phone number through a link or one-time code, or requested a copy of a recent utility bill to prove California residence. Many $\mathcal{DBR}$s with form-VCRs mandated ticking a box to acknowledge (under penalty of perjury) that the researcher is the same consumer for whom the request was being made. Nevertheless, 13 $\mathcal{DBR}$s asked for a separate *signed affidavit*.

Moreover, 53% of form-VCRs were CAPTCHA-protected. While CAPTCHAs are used to prevent malicious bot activity, CAPTCHA solving is considered difficult for humans. Furthermore, bots are known to be faster and better than humans at this task [22], [23]. Access to 2 $\mathcal{DBR}$s' form-VCRs resulted in multiple CAPTCHAs: reCAPTCHA and a text-based or math-based CAPTCHA. Table 4 shows the frequency of CAPTCHA types observed on forms.

OneTrust forms overwhelm consumers with additional verification: they all require CAPTCHA completion, using either reCAPTCHAs or text-based CAPTCHAs. 77% require confirming access to the supplied email address. In contrast, only 37% of *non*-OneTrust forms are CAPTCHA-protected. Compared to other form-VCRs, OneTrust forms require significantly more additional verifications before submission (Mann-Whitney U test p-value < 0.001). The researcher also needed to solve a text-based CAPTCHA every time they wished to access a OneTrust portal after VCR submission, e.g., to verify the request's status.

KEY TAKEAWAYS - RQ1 PART 2 The identity verification process for consumers is burdensome and intrusive. While most $\mathcal{DBR}$s require only a few PII elements, the type of these elements varies significantly. Some ask for PII that is inconvenient to obtain, such as mobile advertiser IDs or signed affidavits. The process can be highly privacy-invasive, with certain $\mathcal{DBR}$s requesting highly sensitive PII, e.g. SSNs or government IDs. Finally, $\mathcal{DBR}$s often implement additional verification steps (e.g., email confirmation, CAPTCHA solving) before allowing VCR submissions, primarily to prevent spam, yet further complicating the VCR submission process for consumers.

TABLE 4: CAPTCHA type frequency in CAPTCHA-protected form-VCRs.

CAPTCHA type	Percentage
reCAPTCHA	72%
text-based	20%
hCAPTCHA	8%
math-based	3%

5.3. RQ2: $\mathcal{DBR}$ responses

A substantial fraction of $\mathcal{DBR}$s (195 out of 454) **never responded** to VCR requests. Only 51.5% of the $\mathcal{DBR}$s responded on time, i.e., within 45 calendar days. Although no $\mathcal{DBR}$ asked for a permitted extension, 5.3% responded after the 45-day limit, thus failing to comply with the prescribed timeline.

We use the term *responding* $\mathcal{DBR}$s to denote those that replied to a VCR, whether on time or late. The rest, even if they initially acknowledged the VCR or corresponded to get more PII for identity verification, are called *non-responding* $\mathcal{DBR}$s.

Most responding $\mathcal{DBR}$s reported having no PI about the researcher. Only 22 $\mathcal{DBR}$s provided some PI, detailed in Section 5.4. Figure 5 shows various replies from responding $\mathcal{DBR}$s.

14 $\mathcal{DBR}$s could not verify the researcher's identity. In 7 cases, their response arrived almost instantly after submitting a VCR form. These $\mathcal{DBR}$s did not ask for additional information to verify the identity. 8 $\mathcal{DBR}$s answered with irrelevant information, such as stating that the researcher was now "opted out" of future sales or communication. Remarkably, one $\mathcal{DBR}$ refused to fulfill the request, falsely claiming that the researcher is not a California resident.

Response rate. Since a significant fraction of VCRs were unanswered, we examined factors associated with receiving a response, such as $\mathcal{DBR}$ location, VCR submission method, number of PII elements requested, and sensitivity of the PII supplied. Note that the analysis of number of PII elements and sensitivity of PII supplied (Figures 6 and 7) excludes multi-step VCRs.

Geographical location. First, we analyzed response rates depending on $\mathcal{DBR}$s' location. We optimistically expected that $\mathcal{DBR}$s located in the US, especially those in California, would have a better grasp of CCPA and would therefore be more likely to answer. 56.6% US-based $\mathcal{DBR}$s and 60% non-US $\mathcal{DBR}$s responded to VCR requests. However, the

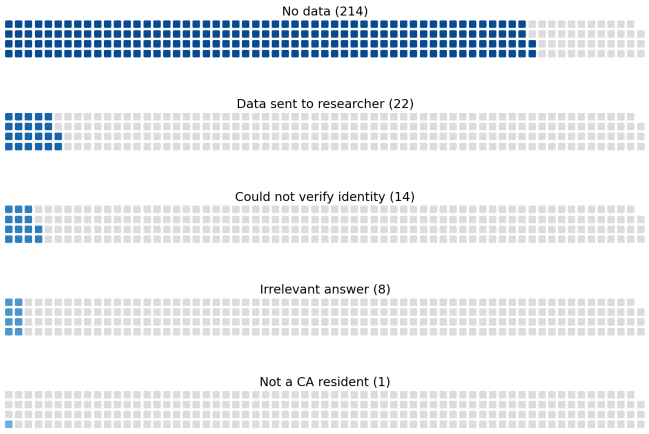


Figure 5: Distribution of answer categories received from $\mathcal{DBR}$s.

latter are geographically dispersed and few in number: 25 total.

Table 5 provides $\mathcal{DBR}$s' response rate per state, in states where a VCR could be submitted to at least 10 $\mathcal{DBR}$s. Comparing three states with the highest $\mathcal{DBR}$ concentration (California, New York, and Florida), the response rate for California, at 58.1%, is only 5.6 and 1.3 percentage points higher than that of the other two, respectively. In other words, California $\mathcal{DBR}$s perform only marginally better than the US average of 56.6%.

TABLE 5: Response rates of $\mathcal{DBR}$s in US states.

State	$\mathcal{DBR}$ Count	Resp. rate
California	93	58.1%
New York	61	52.5%
Florida	37	56.8%
Texas	23	56.5%
Massachusetts	20	70.0%
Georgia	16	62.5%
Virginia	14	50.0%
Colorado	14	50.0%
New Jersey	14	71.4%
Illinois	12	58.3%
Washington	11	72.7%
Ohio	10	70.0%

VCR submission method. Beyond geographical factors, we hypothesized that $\mathcal{DBR}$s that used form-VCRs would be more likely to respond to a VCR. The rationale was that $\mathcal{DBR}$s with established VCR submission processes would be more inclined to allocate both time and personnel to address these requests. Figure 6 shows $\mathcal{DBR}$s' answers depending on the VCR submission method.

Aligning with our assumptions, form-VCRs yielded a higher response rate of 71.5%. Email-VCRs, the second most common VCR submission method, yielded a response rate of 42.4%. This suggests that the majority of these $\mathcal{DBR}$s *only* meet a minimum regulatory requirement by providing an email address, without actually responding to received requests. Finally, phone-VCRs performed worse: only 42.9% received a response. $\mathcal{DBR}$s consistently failed to return calls after voicemails were left. In one notable instance, upon the researcher's request to exercise their CCPA rights, a $\mathcal{DBR}$ representative abruptly transferred the call to voicemail. No follow-up communication was received.

Predictably, 19 $\mathcal{DBR}$s with broken forms that had to be contacted via email or phone exhibited low response rates: 41.2% responded to emails, and none responded to phone calls. 6 of 8 $\mathcal{DBR}$s that sent a form to complete after an initial email contact responded to the VCR, and 2 did not answer after they specifically pointed to a form.

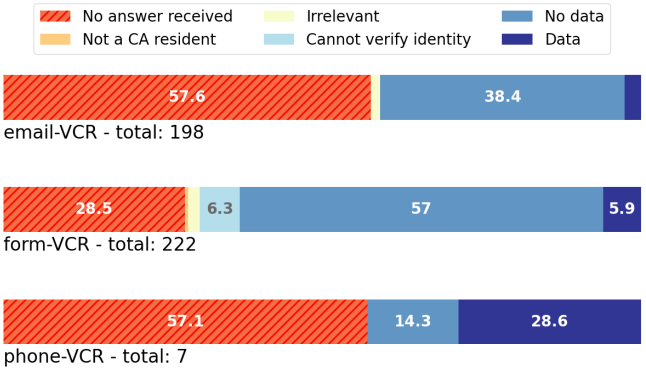


Figure 6: $\mathcal{DBR}$ responses for email, form, and phone VCR submissions (in %).

Number of PII elements requested. To verify the requester's identity, $\mathcal{DBR}$s ask for PII to match with their own records. As confirmed in Section 5.2, due to the lack of standards in CCPA legislation each $\mathcal{DBR}$ asks for different types of PII.

Figure 7 shows the response rate for email-VCRs and form-VCRs depending on the number of PII elements requested for identity verification. 168 email-VCRs were given 2 PII elements by the researcher: email address and name, and were completed in 36.3% of cases only. 30 email-VCRs that required more than 2 PII elements were answered at a higher rate (76.7%). The number of PII elements requested in form-VCRs had no effect on the response rate.

Sensitive PII requested. Finally, we checked whether $\mathcal{DBR}$s that asked for sensitive PII responded to VCRs. Following CCPA's definition of "sensitive personal information", Table 6 lists the seven elements of sensitive PII occasionally requested, and $\mathcal{DBR}$s' response rate when requiring these sensitive PII elements. Most $\mathcal{DBR}$s that required a full or redacted copy of a government ID completed the VCR. Meanwhile, only 2 of 6 $\mathcal{DBR}$s asking for a full SSN responded. This result is especially concerning:

> *Who, now, has access to this information? Does exercising one's CCPA rights in turn put a consumer at risk of identity theft?*

Two $\mathcal{DBR}$s that requested the last 4 digits of the researcher's SSN or that were sent a redacted copy of a government ID responded saying that they could not

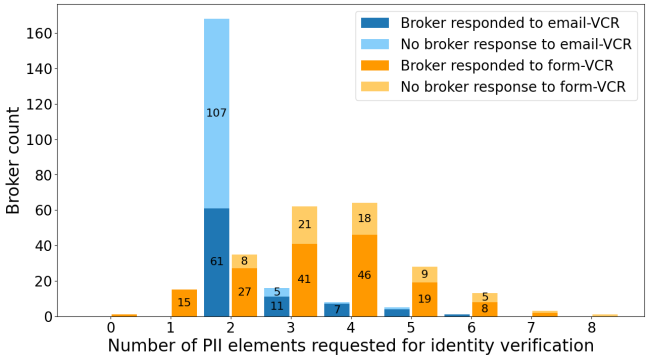


Figure 7: Number of responding $\mathcal{DBR}$s given increasing number of PII elements requested

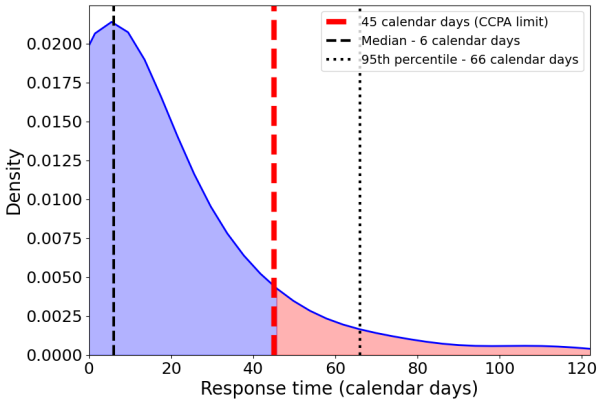
TABLE 6: $\mathcal{DBR}$s response rate when sensitive PII is asked.

Sensitive PII	Completed VCRs
Last 4 digits of SSN	5/9 (55%)
Full SSN	2/6 (33%)
Redacted ID	11/13 (85%)
Full ID	3/4 (75%)
Driver's License Number	1/2 (50%)
Signature/signed affidavit	10/15 (67%)
Selfie	4/5 (80%)
Total	29/45 (64%)

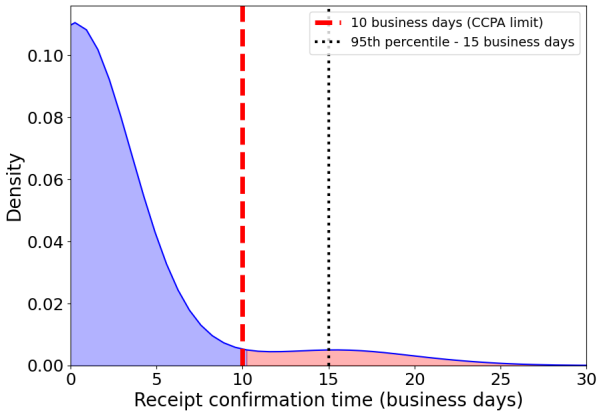
verify the researcher's identity: sharing sensitive PII was ineffective in these cases.

Response Timeline. Figure 8a shows the distribution of the response timeline, in calendar days, between VCR submission and $\mathcal{DBR}$ response, for the 259 responding $\mathcal{DBR}$s. Recall that CCPA requires $\mathcal{DBR}$s to respond within 45 calendar days. Half of the responding $\mathcal{DBR}$s answered the VCR within the first 6 days, and 90.7% responded within the allotted 45 days. $\mathcal{DBR}$s also have to confirm receipt of a VCR within 10 business days. We considered that every form-VCR was immediately confirmed as received, through the form submission confirmation message. For all other VCRs, we assume that the $\mathcal{DBR}$'s first email to the researcher was confirmation of VCR receipt. Figure 8b shows the distribution of time, in business days, between VCR submission and confirmation of receipt. 93.6% of responding $\mathcal{DBR}$s confirmed receipt of the VCR within the allotted 10 business days. In fact, 81.8% confirmed receipt of the VCR within one business day, typically via an automatic reply system. Of the 195 $\mathcal{DBR}$s that did *not* respond, 69 confirmed receipt of the VCR without following up.

Response contents. Figure 9 shows the response distribution of response times categorized by the type of answer provided. $\mathcal{DBR}$s that had PI about the researcher took longer to answer: this could be due to either a more thorough identity verification process, or the effort to find the collected PI. On the other hand, $\mathcal{DBR}$s that could not verify the researcher's identity or replied with



(a) $\mathcal{DBR}$ response time



(b) $\mathcal{DBR}$ VCR receipt confirmation time

Figure 8: Distribution of VCR completion and receipt confirmation time.

irrelevant data (about opt-out or deletion rights) answered more expeditiously. However, it is uncertain whether their answers can be trusted, e.g., did they reply just to quickly "complete" (i.e., get rid of) the VCR?

OneTrust. OneTrust form-VCRs yielded a response rate of 79%, notably higher than the 69% rate of non-OneTrust form-VCRs. Therefore, while OneTrust form-VCRs required additional verifications (CAPTCHA, email confirmations), they had a higher "success" rate. We conclude that the use of such an automated system, preferably standardized, might help $\mathcal{DBR}$s meet compliance requirements.

KEY TAKEAWAYS - RQ2. Measuring response rate revealed widespread non-compliance. A significant fraction of $\mathcal{DBR}$s failed to respond to VCRs, with only about half answering within the prescribed 45-day timeframe. Specifically, email and phone-VCRs were answered only in 42% of cases. Form-VCRs performed better, with a 71% response rate, potentially indicating that providing a form demonstrates a stronger commitment to CCPA compliance. Alarmingly, even $\mathcal{DBR}$s requesting PII exhibited low re-

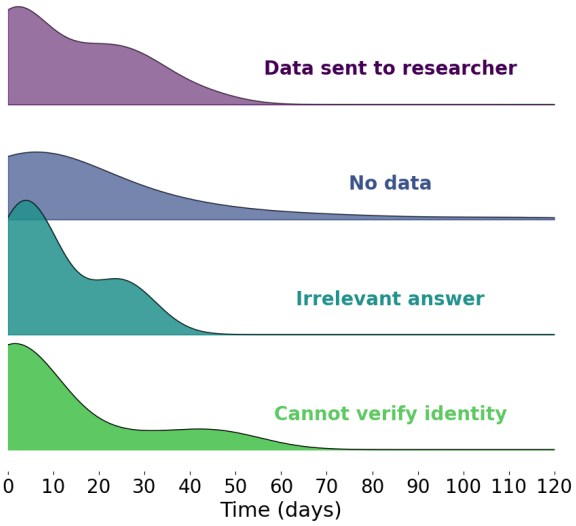


Figure 9: $\mathcal{DBR}$ response time depending on answer content.

sponse rates.

Among responding $\mathcal{DBR}$s, most adhered to CCPA-mandated 45-day limit. The pattern exposed a stark divide: $\mathcal{DBR}$s tend to either fully comply with or completely disregard CCPA obligations. Overall, these findings paint a rather disappointing picture of CCPA compliance.

5.4. RQ3: PI received from $\mathcal{DBR}$s

We analyze received PI to evaluate accessibility, accuracy, PI leakage, and tangentially, mapping PI requested in VCRs to the sensitivity of information $\mathcal{DBR}$s revealed. Recall that 22 $\mathcal{DBR}$s provided PI about the researcher. These $\mathcal{DBR}$s' registered addresses span 11 states within the US, and three countries (USA, Germany, Canada). Table 7 presents details regarding the VCR process and the PI received from these $\mathcal{DBR}$s. The following PI was obtained:

- *Confidential PI (2).* A credit reporting agency sent a credit report, and a major data analytics company sent past and current car insurance information: policy numbers, dates, coverage, and listed drivers. This information is very sensitive.
- *Behavior inference data (3).* These $\mathcal{DBR}$s provided such data for advertising purposes, e.g., `"Samba TV → Ad Exposure → Electronics"`.
- *Publicly available professional information (6).* These $\mathcal{DBR}$s provided a copy of the researcher's public LinkedIn profile.
- *Physical addresses and cookie values (2).* These $\mathcal{DBR}$s provided a limited amount of PI, only giving one or two home addresses and some website-specific cookie ID.
- *Data from the VCR (5).* These provided information that they likely just obtained from the VCR itself. These include only the researcher's name and email address, and VCR-provided IP address, cookie ID, and mobile advertiser ID.
- *Categories of collected PI (3).* These provided the categories, not the actual PI. Thus, we cannot verify whether these $\mathcal{DBR}$s actually collected the researcher's PI or if a generic answer was given.

PI Delivery. Two $\mathcal{DBR}$s that sent the most confidential PI did so by postal mail: the PI itself or a printed download link to the PI was sent to the researcher's home address. This measure ensured that only the consumer with the address matching that in the $\mathcal{DBR}$s' databases would receive the PI. It maintains data security as long as $\mathcal{DBR}$s are confident in the accuracy of their databases. Incidentally, the same two $\mathcal{DBR}$s required a higher-than-average number of PII elements and sensitive PII for identity verification.

Other $\mathcal{DBR}$s sent the PI via email or OneTrust portals. Three had PI directly accessible through their website, e.g., one prompted the researcher to press a "show me the data" button, after which they could download a CSV file containing behavior inference data. $\mathcal{DBR}$s usually provided the PI as PDF, CSV, or XLSX files. One $\mathcal{DBR}$ sent files as a series of nested HTML files with non-descriptive titles, making sense of which is likely to be unintuitive for the average consumer.

Accuracy of PI. Behavior inference data received from 4 $\mathcal{DBR}$s was vague, with multiple age ranges or contradictory interests. Some $\mathcal{DBR}$s sent thousands of data points, and others dozens. Certain inferences were inaccurate, e.g., two $\mathcal{DBR}$s provided data inferring that the researcher is Hispanic or Spanish-speaking, which is not the case.

PI Leakage. Alarmingly, the $\mathcal{DBR}$ from which the researcher received car insurance information included sensitive PI about the following consumers:

1) The researcher's family members, included as drivers in the researcher's policy: this might be acceptable, since the researcher entered this information themselves when starting the policy.
2) The researcher's housemate, on whose policy the researcher was listed as an excluded driver (an individual within the same household specifically listed as not covered under the car insurance policy). We consider this to be a privacy leak. The researcher was sent a third-person's PII, of which the researcher had no prior knowledge.
   Specifically, the researcher received the policy number, start and end dates, coverage limits, and **7 out of 8 characters** of the housemate's driver's license number. While the first half of the report redacted the first 5 characters of the 8-character-long California driver's license number (e.g. XXXXX678), the second half of the report redacted the last 4 characters (e.g. A123XXXX). Accordingly, the researcher obtained 7-out-of-8 characters (e.g. A123X678) and would only need to guess the 5th character, a decimal digit, to obtain their housemate's full driver's license number. We note that CCPA prohibits disclosing any consumer's "...driver's license number or other

TABLE 7: PI Received from $\mathcal{DBR}$s in Response to VCRs.

	Type of PI received	PI transmission method	File format	# PII for ID verif.	Ans. time (days)	Location
1	Credit report	Postal mail	N/A	6*	0	Georgia
2	Car insurance information	Link to PI sent via postal mail	pdf	7*	1	Georgia
3	Behavior inference for advertising	Email attachment	pdf	4	44	California
4		Expiring link	xlsx	4	29	Virginia
5		Directly website accessible	csv	0	0	New York
6	LinkedIn profile data	Email attachment	pdf	2	0	California
7		Email attachment	pdf	8	1	New York
8		Email attachment	pdf	3	5	Washington
9		Email attachment	json	2	17	Washington
10		Password-protected email attachment	xlsx	5	26	Delaware
11		Email content	N/A	4	1	California
12	Physical addresses and company-specific cookies	Expiring link	csv	7*	12	California
13	Website cookie ID	Directly website accessible	N/A	2	14	Canada
14	Email address	Expiring link	html	4	34	Virginia
15	Shortened first name	Directly website accessible	N/A	3	1	Colorado
16	IP address information	Email attachement	json	3	4	Nevada
17	Cookie ID and mobile advertiser ID	OneTrust Portal	N/A	5	32	Germany
18	Name, email address	OneTrust Portal	pdf	4*	22	New York
19	Categories of PI	Phone-call	N/A	2	0	Utah
20		OneTrust Portal	txt	4	0	Texas
21		Email attachment	xlsx, json	5	24	Massachusetts
22	History/details of services provided by the company	OneTrust Portal	pdf	5	20	Georgia

* indicates sensitive PII was requested

government issued identification number [in response to a request to know]" (Tit. 11 Div. 6 §7024(d)) [19].

KEY TAKEAWAYS - RQ3. While a few $\mathcal{DBR}$s shared confidential PI or advertising behavioral inference data, the majority of PI shared was not particularly meaningful. Most $\mathcal{DBR}$s simply returned the PII provided by the researcher, or at best, copied LinkedIn profile data.

$\mathcal{DBR}$s typically transmitted PI via email or OneTrust portal. Notably, two $\mathcal{DBR}$s that shared the most sensitive PII opted for postal mail delivery, demonstrating enhanced security measures. This suggests that $\mathcal{DBR}$s tailored their delivery methods to the sensitivity of the PII transmitted. File formats, with only one exception, were easily accessible and transmittable, following CCPA requirements. Surprisingly, one broker leaked sensitive PII about a third party.

6. Discussion

Given the study results, we now consider privacy tradeoffs in the VCR process, unexpected and noteworthy incidents, and finally recommendations stemming from these, for consumers, data brokers, and policymakers.

6.1. The Privacy Paradox

The goal of privacy laws is to give consumers greater control over their PI. Ironically, exercising one's CCPA-given privacy rights introduce new privacy risks and vulnerabilities. The researcher sent a large amount of PII to hundreds of $\mathcal{DBR}$s, 95% of which either did not respond or did not previously collect any PI about the researcher. In one email informing that no PI was collected about the researcher, a $\mathcal{DBR}$ even stated, about the email address and name given to submit the VCR:

"In fact, we never had this information until receiving it in your email below."

This highlights current privacy issues stemming from exercising one's CCPA rights. It is especially concerning when $\mathcal{DBR}$s request sensitive PII in VCRs. As noted in Section 5.3, 6 $\mathcal{DBR}$s asked for the researcher's full SSN, and 4 of them did not respond to the VCR. This raises serious concerns about identity theft risks created by the VCR process, which is intended to protect consumer privacy. Not only does the consumer submitting the request become vulnerable, but privacy threats to other individuals also emerge, i.e., as mentioned earlier, the researcher received sensitive PI about another consumer. On a related note, prior work shows that an impersonator could easily receive another consumer's PI, since the PII required for identity verification is often public or easily obtainable [5], [8], [24].

This leaves two imperfect choices:

1) $\mathcal{DBR}$s should require a copious amount of (potentially sensitive) PII to more accurately identify the consumer, which is detrimental to consumer privacy.

   OR

2) $\mathcal{DBR}$s should require less PII, thus favoring consumer privacy. However, they would be more prone to data breaches due to ease of impersonation resulting from a simplified identity verification process.

6.2. Unexpected and unintended outcomes

Even though the researcher only asked to exercise their PI access rights, many $\mathcal{DBR}$s, along with responding to VCRs, announced that they were putting the researcher's information on their opt-out list. They either assume that the researcher cares about their privacy and anticipate an opt-out/deletion request, or they hope to not receive any more VCRs from the researcher by proactively adding them to their opt-out list.

Several factors may explain why such a small percentage of $\mathcal{DBR}$s shared collected PI with the researcher. Recall that only 22 $\mathcal{DBR}$s, 4.6% of all contacted $\mathcal{DBR}$s, and 8.1% of

all responding $\mathcal{DBR}$s, shared PI which they collected about the researcher. Beyond non-compliance, there are several potential reasons. CCPA only requires $\mathcal{DBR}$s to share PI that was collected within the past 12 months (1798.130(a)(2)(B)) [2]. If PI was collected earlier, $\mathcal{DBR}$s might not have to disclose it. In addition, if the collected PI is considered "public" in any way, it is not considered PI under CCPA, and $\mathcal{DBR}$s do not have to disclose it (1798.140(v)(2)) [2].

Finally, we had no way of verifying the veracity of responding $\mathcal{DBR}$s that claimed they had no PI.

6.3. Noteworthy incidents

The study revealed several noteworthy incidents:

Noncompliance. Four $\mathcal{DBR}$s added the researcher's email address to their marketing/newsletter list, which clearly violates CCPA (1798.130(a)(7)) [2]. Two of three major US credit reporting agencies (Experian, Equifax, TransUnion) did not respond to our request. One $\mathcal{DBR}$ that definitely had the researcher's exact current address (it was one of the multiple-choice questionnaire options) did not answer the VCR.

Inaccessible VCRs. One $\mathcal{DBR}$ required the researcher to provide cookie values from a specified website. The website was down. Although we communicated this to the $\mathcal{DBR}$, no reply was received. One $\mathcal{DBR}$ required email access confirmation by clicking a link in an email, even though no link was included in that email.

Oddities. A couple of $\mathcal{DBR}$s used email obfuscation in their privacy policy, making it harder for both bots and humans to send emails. We also observed that employees of some $\mathcal{DBR}$s visited the researcher's LinkedIn profile, probably for identity verification purposes.

6.4. Recommendations

We now make a few improvement recommendations for all three stakeholders involved: policymakers, $\mathcal{DBR}$s, and consumers.

Policymakers. Policymakers need to mandate standardization of both the VCR submission process and the maximal set of information that $\mathcal{DBR}$s should be allowed to request for identity verification purposes. More importantly, increased enforcement is needed through random audits and legal actions against non-compliant $\mathcal{DBR}$s. While CPPA already fined some $\mathcal{DBR}$s for failing to register [25], [26], [27], we are unaware of any $\mathcal{DBR}$s facing legal repercussions for failing to process consumer requests.

Plus, there needs to be a standardized process for consumers to lodge grievances against unresponsive or otherwise non-compliant $\mathcal{DBR}$s. Finally, consumers need to be informed about a history of such grievances against each $\mathcal{DBR}$ (e.g., why bother submitting yet one more VCR to a $\mathcal{DBR}$ that has numerous recent complaints).

$\mathcal{DBR}$s. Of course, we strongly recommend that $\mathcal{DBR}$s must comply with CCPA. We also suggest that $\mathcal{DBR}$s provide direct links to relevant VCR forms or email contacts as part of their CPPA registration, instead of providing a generic link to their privacy policy. This may help $\mathcal{DBR}$s to not confuse different rights: $\mathcal{DBR}$s should not opt-out consumers from the sale of their data if such a request was not made.

Furthermore, $\mathcal{DBR}$s must make the VCR submission process easy for consumers. If asking for online identifiers is essential (e.g., via cookie ID or MAID), obtaining such information should be trivial. Finally, $\mathcal{DBR}$s must train designated employees to properly handle responding to phone-VCRs and email-VCRs.

Consumers. Consumers should exercise caution when submitting VCRs and, if problems are encountered, file official grievances using the CPPA complaint form [28].

6.5. Limitations

Throughout this study, one researcher sent VCRs to $\mathcal{DBR}$s. We acknowledge that having multiple individuals submit VCRs would yield more data. This is the subject of future work. In a similar vein, studying CCPA compliance for non-existent (or deceased) individuals might yield a better overall picture. However, that would also trigger more ethical issues. Comparing $\mathcal{DBR}$s' compliance with other privacy laws, e.g., GDPR, would also be interesting. However, GDPR unfortunately does not mandate $\mathcal{DBR}$ registration.

Our findings likely underestimate the full extent of non-compliance in the $\mathcal{DBR}$ ecosystem, since our methodology only assessed entities that have already fulfilled the basic legal CCPA obligation of $\mathcal{DBR}$ registration. Unregistered data brokers (which by definition are already failing to comply with CCPA) were beyond the scope of this study.

Our results, in particular, the number of $\mathcal{DBR}$s that had data about the researcher and their responses, might have been influenced by the specific researcher's characteristics and online behavior. The researcher in question is under 30 years of age, privacy-conscious, and not very active on social media platforms. This naturally limited the amount of data collected by $\mathcal{DBR}$s. It is likely that $\mathcal{DBR}$s would have more data on older individuals, and/or those with richer financial histories (e.g., loans, debts, real estate, etc.), as well as ones who are more active on social media.

7. Conclusion

This comprehensive study of 543 California-registered data brokers reveals concerning patterns of CCPA (non-) compliance. A substantial portion of data brokers (43%) completely ignored VCRs. This widespread non-compliance undermines CCPA's efficacy and highlights significant enforcement gaps in current privacy regulation. Our findings

expose a troubling paradox: exercising privacy rights under CCPA introduces new privacy vulnerabilities. Consumers must provide personal information – sometimes including sensitive PII, such as SSNs, government IDs, and biometric data – to data brokers who may never respond. This creates a privacy catch-22 where consumers must risk exposing additional personal data to potentially untrustworthy entities to learn what information these brokers already (do not) possess. The lack of standardization in verification procedures, places the burden on consumers to navigate each broker's unique process, providing various levels of personally identifiable information and creating substantial barriers to accessing their own data.

Acknowledgments

We thank IEEE S&P reviewers for their constructive feedback. This work was supported in part by funding from NSF Award SATC 1956393, and a 2024 Guggenheim Fellowship.

References

[1] European Parliament and Council, "Regulation (eu) 2016/679, general data protection regulation," 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/

[2] California Legislature, "Title 1.81.5. california consumer privacy act of 2018," 2018. [Online]. Available: https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

[3] Governo do Brasil, "LGPD - Lei Geral de Proteção de Dados," Ministério do Esporte, 2023, accessed: [5/28/2025]. [Online]. Available: https://www.gov.br/esporte/pt-br/acesso-a-informacao/lgpd

[4] California Office of the Attorney General, "Data broker registry," 2025, accessed June 4, 2025. [Online]. Available: https://oag.ca.gov/data-brokers

[5] M. D. Martino, P. Robyns, W. Weyts, P. Quax, W. Lamotte, and K. Andries, "Personal information leakage by abusing the GDPR 'right of access'," in *Fifteenth Symposium on Usable Privacy and Security, SOUPS 2019, Santa Clara, CA, USA, August 11-13, 2019*, H. R. Lipford, Ed. USENIX Association, 2019. [Online]. Available: https://www.usenix.org/conference/soups2019/presentation/dimartino

[6] M. D. Martino, I. Meers, P. Quax, K. Andries, and W. Lamotte, "Revisiting identification issues in GDPR 'right of access' policies: A technical and longitudinal analysis," *Proc. Priv. Enhancing Technol.*, vol. 2022, no. 2, pp. 95–113, 2022. [Online]. Available: https://doi.org/10.2478/popets-2022-0037

[7] S. Adhatarao, C. Lauradoux, and C. Santos, "Why ip-based subject access requests are denied?" *arXiv preprint arXiv:2103.01019*, 2021.

[8] L. Bufalieri, M. L. Morgia, A. Mei, and J. Stefa, "GDPR: when the right to access personal data becomes a threat," in *2020 IEEE International Conference on Web Services, ICWS 2020, Beijing, China, October 19-23, 2020*. IEEE, 2020, pp. 75–83. [Online]. Available: https://doi.org/10.1109/ICWS49710.2020.00017

[9] C. Boniface, I. Fouad, N. Bielova, C. Lauradoux, and C. Santos, "Security analysis of subject access request procedures - how to authenticate data subjects safely when they request for their data," in *Privacy Technologies and Policy - 7th Annual Privacy Forum, APF 2019, Rome, Italy, June 13-14, 2019, Proceedings*, ser. Lecture Notes in Computer Science, M. Naldi, G. F. Italiano, K. Rannenberg, M. Medina, and A. Bourka, Eds., vol. 11498. Springer, 2019, pp. 182–209. [Online]. Available: https://doi.org/10.1007/978-3-030-21752-5_12

[10] J. Wong and T. Henderson, "The right to data portability in practice: exploring the implications of the technologically neutral gdpr," *International Data Privacy Law*, vol. 9, no. 3, pp. 173–191, 2019.

[11] D. Herrmann and J. Lindemann, "Obtaining personal data and asking for erasure: do app vendors and website owners honour your privacy rights?" in *Sicherheit 2016: Sicherheit, Schutz und Zuverlässigkeit, Beiträge der 8. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI), 5.-7. April 2016, Bonn*, ser. LNI, M. Meier, D. Reinhardt, and S. Wendzel, Eds., vol. P-256. GI, 2016, pp. 149–160. [Online]. Available: https://dl.gi.de/handle/20.500.12116/863

[12] J. L. Kröger, J. Lindemann, and D. Herrmann, "How do app vendors respond to subject access requests?: a longitudinal privacy study on ios and android apps," in *ARES 2020: The 15th International Conference on Availability, Reliability and Security, Virtual Event, Ireland, August 25-28, 2020*, M. Volkamer and C. Wressnegger, Eds. ACM, 2020, pp. 10:1–10:10. [Online]. Available: https://doi.org/10.1145/3407023.3407057

[13] N. Samarin, S. Kothari, Z. Siyed, O. Bjorkman, R. Yuan, P. Wijesekera, N. Alomar, J. Fischer, C. J. Hoofnagle, and S. Egelman, "Lessons in VCR repair: Compliance of android app developers with the california consumer privacy act (CCPA)," *Proc. Priv. Enhancing Technol.*, vol. 2023, no. 3, pp. 103–121, 2023. [Online]. Available: https://doi.org/10.56553/popets-2023-0072

[14] K. Take, J. Young, R. Bhalerao, K. Gallagher, A. Forte, D. McCoy, and R. Greenstadt, "What to expect when you're accessing: An exploration of user privacy rights in people search websites," *Proc. Priv. Enhancing Technol.*, vol. 2024, no. 4, pp. 311–326, 2024. [Online]. Available: https://doi.org/10.56553/popets-2024-0118

[15] A. Borem, E. Pan, O. Obielodan, A. Roubinowitz, L. Dovichi, M. L. Mazurek, and B. Ur, "Data subjects' reactions to exercising their right of access," in *33rd USENIX Security Symposium (USENIX Security 24)*, 2024, pp. 2865–2882.

[16] T. Urban, D. Tatang, M. Degeling, T. Holz, and N. Pohlmann, "A study on subject data access in online advertising after the GDPR," in *Data Privacy Management, Cryptocurrencies and Blockchain Technology - ESORICS 2019 International Workshops, DPM 2019 and CBT 2019, Luxembourg, September 26-27, 2019, Proceedings*, ser. Lecture Notes in Computer Science, C. Pérez-Solà, G. Navarro-Arribas, A. Biryukov, and J. García-Alfaro, Eds., vol. 11737. Springer, 2019, pp. 61–79. [Online]. Available: https://doi.org/10.1007/978-3-030-31500-9_5

[17] J. He, P. Snyder, H. Haddadi, F. E. Bustamante, and G. Tyson, "Measuring the accuracy and effectiveness of PII removal services," *Proc. Priv. Enhancing Technol.*, vol. 2025, no. 4, pp. 166–182, 2025. [Online]. Available: https://doi.org/10.56553/popets-2025-0125

[18] California Legislature, "Title 1.81.48. data broker registration," 2019. [Online]. Available: https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=CIV&division=3.&title=1.81.48.&part=4.&chapter=&article=

[19] C. Legislature, "California code of regulations." [Online]. Available: https://govt.westlaw.com/calregs/Index?transitionType=Default&contextData=\%28sc.Default\%29

[20] California Privacy Protection Agency, "Data Broker Registry." [Online]. Available: https://cppa.ca.gov/data_broker_registry/

[21] "OneTrust," https://www.onetrust.com/, OneTrust LLC, 2023, accessed: [05/28/2025].

[22] A. Searles, Y. Nakatsuka, E. Ozturk, A. Paverd, G. Tsudik, and A. Enkoji, "An empirical study & evaluation of modern {CAPTCHAs}," in *32nd usenix security symposium (usenix security 23)*, 2023, pp. 3081–3097.

[23] E. Bursztein, S. Bethard, C. Fabry, J. C. Mitchell, and D. Jurafsky, "How good are humans at solving captchas? a large scale evaluation," in *2010 IEEE symposium on security and privacy*. IEEE, 2010, pp. 399–413.

[24] J. Pavur and C. Knerr, "Gdparrrrr: Using privacy laws to steal identities," *arXiv preprint arXiv:1912.00731*, 2019.

[25] CPPA, "Data broker promoting ability to dig up "scary" amounts of information agrees to shut down," Feb 2025. [Online]. Available: https://cppa.ca.gov/announcements/2025/20250227.html

[26] ——, "Cppa orders florida data broker to pay fine," May 2025. [Online]. Available: https://cppa.ca.gov/announcements/2025/20250508.html

[27] ——, "Washington data broker agrees to pay fine for failing to register," Jul 2025. [Online]. Available: https://cppa.ca.gov/announcements/2025/20250729.html

[28] California Privacy Protection Agency, "California privacy protection agency complaint form," 2025, accessed: 2025-06-05. [Online]. Available: https://cppa.ca.gov/webapplications/complaint

Appendix A. Email template

Subject: Request for Information Under CCPA "Right to Know/Access"

To whom it may concern,

I am writing to you in order to exercise my rights under the California Consumer Privacy Act (CCPA) "right to know".

Please kindly provide the following information pertaining to my personal data:

- The categories of personal information collected **and** a copy of all specific pieces of information collected about me.
- The categories of sources of my personal information.
- The purposes for collecting my personal information.
- The categories of third parties with whom my personal information is shared or sold.
- The categories of my personal information that was sold or shared to third parties.

Please let me know if you need any further information from me.

Thank you for your prompt reply.

Sincerely,

<Name>

Appendix B. Meta-Review

B.1. Summary

This paper studies the compliance of data brokers with the California Comsumer Privacy Act (CCPA). One of the researchers manually submitted VCRs (Verifiable Consumer Requests) to all 543 data brokers that were registered with the California data broker registry in 2024. The paper analyzes the process of submitting these VCRs and the responses of these data brokers and aims to identify any potential non-compliance regarding the CCPA. It finds that most data brokers are not compliant, and with a few engaging in explicitly illegal behavior.

B.2. Scientific Contributions

- Independent Confirmation of Important Results with Limited Prior Research
- Provides a Valuable Step Forward in an Established Field

B.3. Reasons for Acceptance

1) The paper provides independent confirmation of important results with limited prior research. The paper conducts a large-scale study of data brokers and their compliance; this has been understudied in prior work. In addition to presenting the results of its requests, the paper provides a detailed overview of the challenges consumers might encounter while attempting to submit a VCR.
2) The paper provides a valuable step forward in an established field. While there has been prior work on data brokers, this paper takes an exhaustive approach towards evaluating data brokers' compliance with existing regulations and surfaces important findings.

B.4. Noteworthy Concerns

The paper is about one researcher's experience with submitting VCRs and the results (i.e., the data that was shared or not shared) are limited by the data that specific companies have about the researcher. The findings would be stronger if multiple researchers submitted requests or if the authors relied on data donations from participants.