USENIX

THE ADVANCED COMPUTING
SYSTEMS ASSOCIATION

Addressing the Address Books' (Interdependent) Privacy Issues

Kavous Salehzadeh Niksirat, *University of Lausanne / Max Planck Institute for Security and Privacy;* Lev Velykoivanenko, *University of Lausanne;* Samuel Mätzler, *University of Zurich;* Stephan Mulders, *Maastricht University;* Aurelia Tamò-Larrieux, Marc-Olivier Boldi, Mathias Humbert, and Kévin Huguenin, *University of Lausanne*

https://www.usenix.org/conference/usenixsecurity25/presentation/niksirat

This paper is included in the Proceedings of the 34th USENIX Security Symposium.

August 13–15, 2025 • Seattle, WA, USA

978-1-939133-52-6

Open access to the Proceedings of the 34th USENIX Security Symposium is sponsored by USENIX.

Addressing the Address Books' (Interdependent) Privacy Issues

Kavous Salehzadeh Niksirat[*†], *Lev Velykoivanenko*[*], *Samuel Mätzler*[‡], *Stephan Mulders*[§], *Aurelia Tamò-Larrieux*[*], *Marc-Olivier Boldi*[*], *Mathias Humbert*[*], *and Kévin Huguenin*[*]

[*]*University of Lausanne* [†]*Max Planck Institute for Security and Privacy* [‡]*University of Zurich* [§]*Maastricht University*

Abstract

Interdependent privacy (IDP), which refers to situations where individuals affect the privacy of others, is a growing concern and has been studied in various contexts. Digital address books (DABs), where users store personal information about others on online services, are a compelling yet understudied case of IDP. In this paper, we present a multi-faceted analysis of DABs. In particular, we conducted two online survey studies with $N = 463$ and $N = 459$ DAB users to understand how they interact with their DABs, perceive and manage associated privacy risks, and support data protection rights. Our studies notably reveal that (i) the privacy leakage due to DABs is substantial, (ii) users are well aware of the privacy (incl. IDP) risks of DAB data but have only moderate privacy concerns and are quite comfortable granting access to this data, and (iii) users are relatively open to respecting the rights of data subjects. We conclude with concrete design recommendations for a privacy-aware DAB ecosystem.

1 Introduction

In recent years, the concept of privacy has evolved, revealing that it is no longer solely within the control of the individuals concerned. A particularly noteworthy development is the concept of interdependent privacy (IDP) [8, 40] (or bystander privacy [68] in the context of ubiquitous technologies) that refers to situations where an individual's privacy is compromised by others. This issue has been explored in various domains, including social networks [21, 76, 78], location-based services [66], genomics [39, 41], app permissions [57], voice assistants [1], smart homes [29], and augmented reality [64].

However, one significant domain that remains underexplored is that of digital address books (DABs), where individuals store contact information about *other* individuals. DABs can be stored online, typically as part of an e-mail service.[1] DABs represent a simple, yet fundamental, example of IDP: Users of online DAB services (inadvertently) share personal data about their contacts—often without their contacts' consent or even awareness. In some cases, DAB service providers (DAB-SPs) even actively encourage their users to input more personal information about their contacts; for instance, Google does so for birthdays (see Figure 7).

DAB data contains personal and identifying information, in a structured format, and thus easily exploitable without using any complex processing, unlike, for instance, location [65] and genomic data [39]. Such information includes first and last names, profile photo, e-mail addresses and phone numbers, birthday, and home address. Recently, pronouns were added to the vCard format, as well as to popular apps and services. Some of these so-called contact fields can be used for profiling (e.g., race from photo, socio-economic status from job title, age from birthday, gender identity from pronouns) [28, 89] or even for identity theft [74, 90]. Some fields, alone or grouped, constitute unique identifiers. Most importantly, having multiple such unique identifiers in a contact card reveals that they correspond to the *same* individual, thus enabling profiles associated with the different identifiers to be *merged*. For instance, a profile associated with an individual's phone number could be merged with a profile associated with their e-mail address, or the profiles associated with an individual's old and new phone numbers could be merged.

Despite extensive research on various aspects of IDP [1, 29, 41, 57, 60, 64] and of DABs [7, 20, 25, 42, 49] as independent topics, the IDP-related challenges within the context of DABs have received only little attention (mostly in the context of mobile permissions [20, 46]). In this work, we conduct an in-depth analysis of the IDP aspects of online DABs for both users and non-users; more specifically, we pose the following research questions (RQs):

RQ1. How do users interact with their DABs, what contact data do they store, and how complete are their DABs?

RQ2. How do users perceive the privacy risks associated with storing others' personal data, and what is their

[1] Some apps (e.g., instant messaging) handle their users' DABs internally. Also, some DABs are automatically populated based on the interactions users have with others. In the paper, we focus only on DAB data (contact cards with various fields), input and managed by users.

level of awareness and support for enforcing the data protection rights (e.g., access) of those individuals?

RQ3. What kinds of privacy-preserving remedies do users envision to mitigate these risks and support these rights?

To address our RQs, we present a multi-faceted investigation of the (interdependent) privacy implications of the use of DABs. In a brief preliminary legal analysis, centered on the GDPR, we study the legal roles of the different parties (i.e., DAB users and their contacts, DAB-SPs) and their associated rights and obligations. And through data-access requests sent to five popular DAB-SPs, we also study the DAB-SPs' interpretations regarding these roles. Our user-centered analysis is composed of two complementary large-scale online user surveys ($N = 463$ and $N = 459$). The first survey (i.e., Study #1) collects self-reported data on how users use their DABs and on how they perceive the associated privacy issues and legal rights, whereas the second (i.e., Study #2) collects actual user DAB data and behavioral data on users' decisions to share their DAB data.

One conclusion in our preliminary legal analysis is that, under some conditions (that, for instance, Google Contacts meets), the individuals whose information appears in users' DABs should be considered as data subjects hence have the associated rights, including the right of access (to their data) and right to object (to the processing of their data). Our results show that, in practice, these individuals cannot exert these rights—especially when they are not themselves users of the considered service. In contrast, our survey respondents are rather in favor of enabling these individuals to exert these rights. The results of our user-centered analysis reveal that respondents have modest concerns about the privacy implications of DABs; they are most concerned with photo and address data. The vast majority of our respondents realize that DAB-SPs can technically access DAB data from their users, and a substantial proportion of them are aware of the associated IDP aspects. However, a non-negligible proportion of the respondents are willing to share their DAB data for only a few dollars and a large majority of the respondents grant access to their contact data to at least one third-party mobile app. Our results also show that contact cards rarely contain information such as photo, birthday, and address but that, aggregated over all users, the chances that a DAB-SP has access to such information about an individual is quite high.

Our paper makes the following contributions: we provide empirical findings from a large-scale online survey capturing how users perceive the IDP risks of DABs, including how often such data is shared with third parties mobile apps, their awareness of and support for extending data protection rights to individuals whose data they store. We conducted a second large-scale user study combining actual DAB data with behavioral measures, revealing what types of personal information users typically store, and how users value their DAB data when making real sharing decisions. Based on these key findings, we discuss design implications.

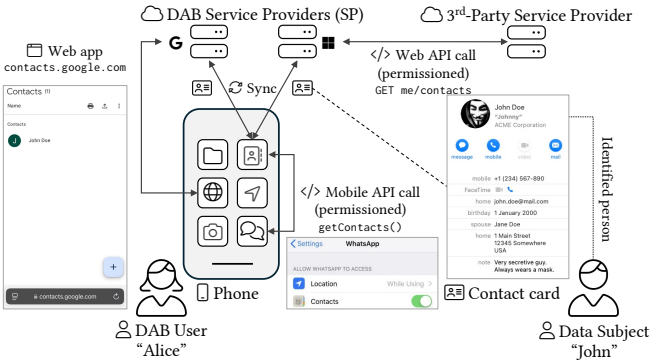


Figure 1: A typical digital address book (DAB) ecosystem.

2 Background and Related Work

Digital Address Books. Figure 1 gives an overview of a typical digital address book (DAB) ecosystem.

Ecosystem. People can manage their DABs—i.e., contact lists—on their electronic devices. DABs contain contact cards that include data fields such as first and last names, phone numbers and e-mail addresses (unique identifiers), company and job title (employment information), photo, important dates (e.g., birthday, anniversary), addresses, related persons, social profiles (e.g., LinkedIn and Instagram, unique identifiers as well), and personal notes. The vCard format [69], which is used as a data interchange format between apps, specifies such data fields and is regularly updated (e.g., gender and pronouns were introduced in v4).

Henceforth, we refer to an individual who manages their DAB using an app as a *(DAB) user* (who we name Alice), and an individual whose personal information is included in the DAB of a user as a *(data) subject* (who we name John). Note that we focus on actual DAB data (i.e., contact cards input and managed by users), not on DABs populated based on communication meta-data (i.e., who communicates with whom) or DAB meta-data (i.e., who has whom in their DABs).

Users can use online services to store their DABs. Doing so enables them to access their DAB through a web app or a desktop/mobile client app[2] synchronized with their online account.[3] Doing so also provides backup and between-device synchronization functionalities. Popular DAB services providers (DAB-SPs) include Apple, Google, and Microsoft. Proton [70] is an interesting example as it focuses on privacy. Open-source implementations of such services (i.e., DAB server apps[2]) exist, e.g., Radicale v3 [47] and Zimbra [83].

Third-party mobile applications (TPAs) and online services can request access to users' DABs through (permissioned) mobile (e.g., [2, 31]) and web (e.g., [32, 59]) APIs. For instance, the WhatsApp mobile app can access a user's DAB

[2]The entity that develops and provides the app is not relevant here as, *a priori*, it does not have access to the data processed in the app.

[3]CardDAV is a standard protocol to synchronize DABs [22].

stored/synced on their mobile phone and facebook.com can access a user's Google Contacts DAB on Google's servers.

When specific protections are not used (i.e., data is not encrypted or the DAB-SP has access to the decryption key), the DAB-SP can technically access "in the clear" the data in their users' DABs. A standard protection measure is end-to-end encryption (E2EE),[4] whereby the data is encrypted "at rest" on the DAB-SP's servers and only the user has the decryption key, typically on their trusted devices where they can decrypt and use their contact data. Most DAB-SPs, including Apple, Google, and Microsoft, do *not* offer such a feature. Apple offers E2EE ("advanced data protection") for photos, files, etc. but not for contacts [3]. Proton, however, partially does [71].

Threat Model. We consider the case where data about an individual (the subject) is communicated/leaked–via DAB entries (i.e., contact cards)—to a SP by one of its users (not necessarily with malicious intents), thus potentially resulting in a privacy violation. In short, the considered adversary is any SP that can access the contact cards input and managed by some users in their DABs . The considered SP can be the DAB-SP or a third party SP that accesses the users' DABs through a mobile or web API. The leaked data could be further used by the SP, for instance for profiling the subject (and monetize these profiles by selling them or through targeted advertisement), who might not be a user of the SP.

Specifics. Some specifics of DABs should be noted. First, some DAB data might be populated automatically by SPs (e.g., basic contact cards with names and e-mails of the individuals with whom the user exchanged e-mails). Second, if the subject is also a user of the DAB-SP, the data about the subject that other users put in their DABs might already be known to the DAB-SP. Third, while associated with personal data about the subject, parts of the data in a contact card about the subject is subjective data produced by users (e.g., notes).

General Data Protection Regulation. European data protection, codified in the GDPR, sets requirements for data controllers/processors processing personal data of data subjects. We define each term briefly. First, data controllers are entities who, alone or jointly with others, determine the *purpose* and the *means* of processing personal data (Art. 4(7) GDPR; European Data Protection Board (EDPB)-Guidelines 07/2020 [23]). A data controller can be a legal entity or a natural person; albeit a natural person processing data purely for a personal activity will not fall within the scope of the law (the so-called *household exemption*). Second, the processing of personal data, as the material scope of the law, is defined broadly [30, 72] to include any processing along the life-cycle of data (from collection to erasure). Case law and data-protection authorities (WP29 [4]) broadly define personal data to include not only names, birthdays, and addresses but also dynamic IP addresses (CJEU, Breyer-Case C-582/14). Third, personal data always relates to a data subject, specifically a natural person whose personal data is processed (Art. 4(1); CJEU, IAB Europe-Case C-604/22). Aside from data controllers, the GDPR also defines data processors, i.e., natural or legal persons that process personal data on behalf of the controller. The distinction between these legal roles—data controller and data processor—are relevant because only the controller is responsible for answering a data access request (DAR)[5] by a data subject; a data processor merely has to assist the data controller. The distinction is therefore important for data subjects who want to make use of their individual data-protection rights that are provided within the GDPR, e.g., the right to obtain access to the data that is being processed (Art. 15), the right to rectify data (Art. 16), the right to erasure (Art. 16), the right to restrict the processing (Art. 18), and the right to object to the processing (Art. 21).

Interdependent Privacy. IDP risks are studied in various contexts [40]. Biczók and Chia [8] coined the term and studied the concept in the context of Facebook apps, showing that the outcomes may be inefficient and contrary to the best interest of users and/or the platform vendor. In genomics, Humbert et al. [39] quantified the detrimental implications of sharing one's genomic data on the privacy of close relatives. Humbert et al. [41] developed an online tool for raising awareness about these risks. Olteanu et al. [65] showed that jointly sharing location with co-location information may significantly impact the location privacy of others. Liu and Biczók [52] analyzed IDP risks with third-party apps (incl. access to DAB data) and proposed a solution for filtering data collected by users but implicating others as data subjects. Several works studied, with game or graph theory, the interplay between the (strategic) behaviors of individuals involved in IDP situations [8, 37, 38, 66]. Most recently, recognizing the limitations of established privacy threat modeling techniques such as LINDDUN in addressing IDP threats, Liu and Biczók [53] proposed IDPA, the first threat modeling methodology specifically designed for IDP.

Several studies focused on bystander privacy [68], a concept somewhat related to IDP. Saqib et al. [79] recently presented a systematic literature review of bystander concerns in smart homes and potential solutions to address them. Interestingly, Marky et al. [56] reported that smart home hosts care about bystander privacy but struggle to protect it. Barocas and Levy [6] presented a survey of the regulatory aspects of privacy (inter)dependencies. They described how these dependencies operate, the values they implicate, and the legal and technical interventions that can be brought to bear on them.

[4] Here, E2EE refers to the encryption of the users' DAB data (i.e., their DAB data is only decrypted on their devices but not on the DAB-SPs' servers), not to the encryption of the communications/messages between users.

[5] A DAR is a request issued by an individual to obtain a confirmation as to whether some categories of personal data about them are being processed, to what ends, and with whom those data are shared. DAR are typically made electronically and answered by data controllers by sending a copy of the data.

DAB Data and Privacy. Several studies highlighted the privacy risks associated with DABs. Bentley and Chen [7] conducted a user study by combining an Android app for collecting DAB metadata and communication logs with a follow-up survey. Among other results, they reported that participants could not recognize the names of 29% of their contacts. They further revealed a significant disparity between DAB size and its actual usage, thus raising concerns about the privacy issue of storing unused contact data, which goes against the data minimization principle. Bösch et al. [16] introduced the concept of privacy dark patterns (i.e., strategies designed to exploit user data without their informed consent) and highlighted a strategy used by apps to encourage users to share their DABs for contact discovery—thus enabling them to build shadow profiles of their contacts (possibly non-users of the app) without their consent [9, 10, 73].

Several works focused on DAB-related app permissions. Khatoon and Corcoran [46] studied the permissions of eight popular Android apps and showed that all of them had access to the contact lists of their users. Felt et al. [25] surveyed smartphone users to study their concerns about permissions and found that DAB-related risks ranked among the highest, emphasizing the need for careful management of DAB-related app permissions. Jindal et al. [42] studied privacy risks associated with the sharing of contact information from mobile devices and found that 69% of smartphone users have engaged, without the consent of data subjects, in privacy breaches associated with the sharing of DAB data. Tahaei et al. [84] explored app permissions from developer and end-user perspectives. The study highlighted key concerns related to DAB permissions, and revealed that some developers misused contact data they collected in their app. A few works proposed mechanisms for enhancing privacy for DABs. For example, Cha and Pak [19] proposed a system for protecting DAB data by applying policies to hide or replace real contact information with virtual data, when accessed by untrusted apps. Hagen et al. [35] examined privacy risks in contact discovery for mobile messengers, identifying vulnerabilities such as enumeration attacks and weak hashing protocols, and they proposed mitigation techniques such as cryptographic solutions. Similarly, Krahn et al. [48] highlighted the privacy risks associated with using DAB data for nearby-device discovery and proposed a cryptographic protocol to minimize the risks of disclosing DAB data.

Research Gaps. Despite extensive research on various aspects of IDP and of DABs privacy (independently), no prior work has specifically investigated IDP challenges related to DABs. More specifically, there is a lack of understanding regarding the legal and technical implications of DAB data, its usage, users' perceptions and concerns, user-centric privacy-enhancing solutions, and the monetary valuation of DAB data—all of which are addressed in this work.

3 Legal Analysis

We first conducted a legal analysis, with both a theoretical and an experimental component, of the interdependent data protection aspects of DABs. This enabled us to formalize the situation and capture the perspectives of the DAB-SPs.

3.1 Theoretical Analysis

Factual Setup. With respect to DABs, we typically have the following set-up, illustrated in Figure 1. A DAB user, Alice, determines which DAB-SPs she wants to use to store/sync the contact details of her friends, colleagues, *etc.* John is one of them. Both Alice and John are data subjects. Alice is a data subject for all the data in her DAB stored on the DAB-SP's servers; as a DAB user, she is identified through her username (typically her e-mail address). John is a data subject for the data stored in the contact card about him in Alice's DAB; he is identified through the unique identifiers[6] (e.g. , e-mail address, phone number) contained in the card. Alice will have an additional role with regards to data about John. As data subjects, they can each exert their rights. Next, we distinguish different scenarios that have different legal implications.

Scenario 1 (local DAB): The DAB user, Alice, stores her data only locally, on her device. She does not rely on an online DAB service; there is no DAB-SP. Alice is the only controller.

Scenario 2 (encrypted online DAB): Alice stores her data on an online DAB-SP in an encrypted format (i.e., using E2EE, meaning that only Alice has access to the decryption key) [75]. Alice is considered a controller, as long as she determines the purposes and means of the processing operation (e.g., by choosing the DAB-SP). The DAB-SP stores only the encrypted data and enables the synchronization of that data when Alice needs it; the DAB-SP can thus be considered a processor, as it processes personal data *for* Alice.

Scenario 3 (unencrypted online DAB; SP *not* involved): Alice stores her data online (as in Scenario 2) but unencrypted. Determining whether the DAB-SP is a processor or controller depends solely on whether the purposes and means of the data processing are determined by the DAB-SP. In this scenario, the DAB-SP only provides the service to store the contact details but does not determine the purposes (i.e., the reason the data is entered) nor the means. Although a step-by-step approach is needed for a final classification (see Scenario 4), based on the general EDPB guidance on the subject, this simplified Scenario 3 would consider the DAB-SP as a processor. However, it should be noted that, within online services, a market in which there is stark power asymmetry between consumers (users) and providers, and where a spectrum of different services can be offered, this legal classification has been challenged in more recent academic articles [26, 43, 44].

Legal implications for DARs in Scenarios 1-3. In these first three scenarios, the DAB user (Alice) is the sole data

[6] Or a combination of non-unique identifiers, such as address and birthday.

controller. Consequently, Alice is responsible for answering DARs. But, Alice is exempted from answering the request based on the household exemption (Art. 2(2)(c) GDPR). This exemption states that the GDPR does not apply to processing of personal data by a natural person, during a purely personal or household activity. Exceptions should be interpreted narrowly, but Recital 18 explicitly mentions the "holding of addresses" as an example of an activity that can be purely personal. A successful appeal to the household exemption means that the whole GDPR does not apply, regardless of the effect of the data processing. This means that John's DARs are left unanswered. From a data-protection law perspective, the legal implications are a trade-off between exempting natural persons from heavy compliance duties under the GDPR (through the household exemption) and the individual rights of other data subjects. It is important to keep in mind that data protection law is about protecting individuals' privacy and personality rights, hence data protection is not an absolute right but must be balanced against other rights (Recital 4).

Scenario 4 (unencrypted online DAB; SP involved): Here, the technical setup is the same as in Scenario 3, but the DAB-SP further processes the data in the users' DABs for its own purposes (e.g., sells the data to third parties or provides targeted advertisements based on it). Within this scenario, the attribution of controllership is that both the DAB user Alice and the DAB-SP jointly control the purposes and means of the data processing. Alice controls the data entry and the choice of DAB-SP, whereas the DAB-SP controls the processing of the data for its own purposes. This is a situation of joint controllership: Alice and the DAB-SP jointly determine the means and purposes of the data processing [23, 85].

Legal implications for DARs in Scenario 4. Here, the responsibility for answering John's request has to be shared among the joint controllers. Although Alice could invoke the household exemption, the DAB-SP is not a natural person and thus could not. Consequently, the DAB-SP would be bound to respond to John's DAR. Therefore, it will be critical to determine which parties (i.e., Alice and/or the DAB-SP) determine the purposes and means for each processing operation. This is important, as the duties under the GDPR will be shared among the joint controllers. However, the processor and controller roles are assigned based on an analysis of the circumstances of the case (CJEU, IAB Europe-Case C-604/22, para 61). As a result, we need to look at a real-world example. Google being the major e-mail (incl. contact) provider, we chose Google Contacts to conduct a case study.

A Case Study – Google Contacts. Policy documents, such as the Terms of Services and Privacy Policies, can be helpful tools to determine whether the DAB-SP is a processor or (joint) controller. However, keep in mind that these policies are written by the dominant party, the DAB-SP with their best interests at play. Indeed, it depends on how much influence the DAB-SP exerts in practice on the means and purposes of the data processing. Determining this, however, is challenging without proper insights into the data processing operations of the DAB-SP. From the Google Privacy Policy, we see that Google uses the content that users provide (e.g., entries in the DAB) to develop new Google services and to provide personalized advertisements. Google does not specify whether DAB entries are actually used for that. But, according to their privacy policy, they can use the DAB entries for purposes other than just keeping a DAB. Furthermore, they nudge individuals to add the birthdays of their contacts (see Figure 7).

Therefore, it is likely that Google qualifies as a sole controller for processing the DAB entries for the purpose of personalized advertisementand as joint controller for the purpose of DAB management. Arguably, most users enter data in the DAB for the sole purpose of keeping a DAB, not for receiving personalized ads. It is Google who decides that such data is to be used for other purposes and the way it is used. The DAB user does not have any control over these decisions, except by not using the service. As the DAB user exerts influence by its choice for a certain DAB-SP, the user could be considered a joint-controller. That is not relevant for our paper, however, because the user can invoke the household exemption. As long as the DAB-SP is the joint controller, the data subject (John) can request access to his data from the DAB-SP. This possibility of obtaining a response, however, requires John to know about Google Contacts in the first place. Although Google Contacts is a popular DAB-SP, those less popular but with similar policies are more difficult to identify.

Google Contacts corresponds to Scenario 4. Hence, John should be able to get a response from Google for DARs about DAB data. However, as the next section will show, obtaining access to such data has proven impossible in practice.

3.2 Experimental Analysis

We issued DARs as a preliminary, small-scale experiment to gain insight into DAB-SPs' perspectives regarding the legal roles and the associated rights and obligations of the different stakeholders: themselves, their users, and the individuals whose information appears in their users' DABs. Similar experimental approaches have been employed to analyze compliance with DARs [5, 12, 13, 34, 88].

We aimed for five e-mail service providers and identified the most popular ones from MailChimp,[7] as DABs are often tied to email services. We selected: Gmail (#1), Outlook (#2), and Proton (#3). We chose Yahoo! (#5) over AOL (#4) because AOL is part of Yahoo! Inc. We also included GMX (#8) as it is based in the EU and our legal study focuses on GDPR. In the end, two co-authors each created individual e-mail accounts with the five selected service providers from within Europe. Both researchers added the five e-mail addresses of the other researcher to their different DABs.

[7] https://mailchimp.com/resources/most-used-email-service-providers/.

Next, the researchers initiated DARs for each of the considered DAB-SP. For a first step, the right of access of the data subject was exercised via the automatic export feature offered by some services (Gmail, Outlook, Yahoo, GMX). These services allow personal data to be downloaded in an automated manner (e.g., Google Takeout). Note that this was possible only because the requesters (i.e., the researchers) happened to be, by design, also *users* of these providers. This might not always be the case: An individual who does not have a Google account might want to request the data Google has about them. This already reveals some information about these DAB-SPs' perspectives: **DAB-SPs do not seem to consider that they process the personal data of non-users**.

We analyzed the data returned by the DAB-SPs. Although we always found the data corresponding to the DABs of the requesters, we never found the data *about* the requesters, stored in the *other* researcher's DAB (i.e., the contact card that contains the e-mail address of the requester). This indicates that **these DAB-SPs do not consider the individuals whose information is stored in *other* users' DABs as the data subjects of the contact cards about them**. Users are considered to be the sole data subjects of the data stored in their DABs.

We further contacted all DAB-SP. We requested the DAB-SPs to indicate how many times the e-mail addresses associated with their account appear in the databases of the DAB-SPs (i.e., in the DABs of the *other* users of the DAB-SP). See supplementary material for for e-mail templates.[8] Note that unlike in the Princeton Privacy Study,[9] we did not ask SPs information they could not provide. Thus, beside the difference in scale (we contacted only five services and sent only a few e-mails), our study is also different from Princeton's in terms of approach. We discuss the ethical aspects in Section 7.

The responses varied widely. In several cases, the request for access to the data had to be further specified. Though none of the DAB-SPs provided us with the requested data, some did cater to our specific request. For instance, GMX argued that they are merely a provider of a product and are not to be regarded as a controller with regards to the data in the DABs of their users. They further stated that they have no influence on the use of their customers' DABs, and that they do not collect the e-mail addresses their customers store.

Other DAB-SPs responded mostly by saying that they do not have any information about e-mail addresses that are not from their system (Yahoo), or that e-mail addresses that are not connected to their services could not be verified and might adversely affect the rights and freedoms of others (Google). Although it diverges from our analysis, as Google was considered to be a joint controller, we agree that there is a trade-off between the subjects' right of access and the privacy of the users. One DAB-SP understood the request as our wanting to access personal data linked to multiple e-mail addresses but no longer replied to a follow-up e-mail (Outlook).

[8] See our OSF repository https://doi.org/10.17605/osf.io/x46aj.

[9] See Princeton Privacy Study: https://privacystudy.cs.princeton.edu/.

4 User-Centered Online Studies

We complemented our legal analysis with two user-centered online studies. This enabled us to capture the users' perspectives regarding the (interdependent) privacy issues related to the use of DABs and their perceptions regarding the data subjects' rights, including the right of access studied in our legal analysis. We were also able to collect actual DAB data, which enabled us to gain insights into DAB users' practices and perceived value of DAB data.

4.1 Methods

For recruiting survey respondents, we used the Prolific crowdsourcing platform that is considered reliable [67] and fair.[10] We conducted cognitive pre-tests with a few colleagues and soft launches with a dozen respondents to identify and fix potential issues in the design and implementation of our questionnaires.[11] Our studies were approved by our institutional review board; see Section 7 for a discussion about ethics.

4.1.1 Study Design

Study #1: User Attitudes and Perceptions. We conducted an online user study to collect data about users' (1) usage of DABs (RQ1—*self-reported*) and (2) perceptions regarding the IDP issues raised by the use of DABs (RQ2), as well as the potential solutions (RQ3). We targeted respondents in Germany and in the Netherlands as they are EU residents, and our legal analysis from Section 3 (which is based on EU GDPR) applies to them. Also, a high proportion of population in these countries are proficient in English [80]. To select respondents, we first deployed a screener survey (1 min, EUR 0.18), filtering those who *actively* use Apple iCloud Contacts or Google Contacts to *manually* store and manage their personal contacts. As we provided illustrations and instructions for these two (popular) services in the main survey, we chose to recruit respondents who were familiar with them.

In the main survey (20 min., EUR 3.50), respondents were asked between 20 and 23 questions (depending on the survey logic and on their responses). The questions were distributed across ten blocks, and the full transcript of the main survey is provided as a supplementary material.[8] As the phrasing of the questions could possibly affect the perspectives of the respondents, hence their responses, for some of the questions (i.e., S1.Q12/S1.Q13, S1.Q14/S1.Q15, S1.Q16/S1.Q17, and S1.Q18/S1.Q19), we created two versions of the question text: one phrased from the perspective of a 'subject' (i.e., a person whose information is stored in someone else's DAB) and one from the perspective of a 'user' (i.e., a person who

[10] See https://fair.work/en/fw/publications/work-in-the-planetary-labour-market-fairwork-cloudwork-ratings-2022/, visited: May 2025.

[11] Whenever we made any adjustments to the questionnaires after the soft launch, we discarded the collected data and did not use it in the final analysis.

uses a DAB). Following a *between-subject* design, at the beginning of the survey, each respondent was assigned with equal probabilities the 'subject' or 'user' role, thus enabling us to identify potential biases caused by the phrasing of the questions.

Next, we summarize the key blocks or questions. The survey began with a consent form (S1.B2) and screener questions (S1.B4). Next, we asked the respondents about their *practices* regarding the frequency of DAB usage and the completeness of contact information (S1.B5). For example, we asked how often their contact cards included various details, such as job title, using a five-point Likert scale from "Never" to "Always" (S1.Q7). S1.B6 investigated if respondents provide access to their DAB data to *third-party apps*. We asked whether they had ever granted access to their DAB data to any mobile app on their smartphone (S1.Q9), and further inquired about the number (and the names) of apps they had granted access to (S1.Q10, S1.Q11). S1.B7 explores respondents' *awareness* to understand whether they believe that DAB-SPs can access contact cards in a clear form. Responses were collected on a seven-point Likert scale from "Strongly disagree" to "Strongly agree." S1.B8 focused on *concerns*, aiming to understand how concerning respondents felt about the privacy of DAB data, using a five-point scale ranging from "Not at all concerning" to "Extremely concerning." 'Subjects' were asked to rate how concerning it is, from a privacy perspective, that specific personal information about them is stored in someone else's DAB (S1.Q14). 'Users' were asked to rate how concerning it was for them to store similar personal information of others in their own DAB (S1.Q15).

S1.B9 explored respondents' *preferences* regarding legal and technical rights to protect their privacy. For the rights, we focused on some of the individual rights specified within the GDPR, as it applies in the countries where the respondents were recruited. We asked respondents about their desired rights from the two perspectives of 'subjects' and 'users.' The questions covered whether DAB-SPs should be legally required to provide technical means to *prevent* the storage of personal information, *delete* stored data, *access* stored data, or *correct* inaccuracies (S1.Q16-S1.Q17). Next, respondents were asked to shift their perspective—'subjects' were asked to consider themselves as 'users' and vice versa—to explore whether they empathized with the opposite viewpoint regarding privacy rights (S1.Q18-S1.Q19). Additionally, respondents were asked to describe a hypothetical dashboard or interface that would provide access and control over personal information stored in other people's DABs (S1.Q20). The respondent could also optionally draw this dashboard (S1.Q21) for an additional bonus payment (EUR 1.2). Analyzing respondent drawings is a well-established method in security and privacy research (e.g., for understanding mental models [45, 50, 63, 91]).

Lastly, the survey concluded with demographic questions about the respondents' general privacy concerns about using the Internet Users' Information Privacy Concerns scale (IUIPC-8 [33, 54]; see S1.Q22), a standard scale regarding IDP [40] called Value of Other People's Privacy (VOPP [36]; see S1.Q23), and gender identity ([82]; see S1.Q24).

Study #2: User Behaviors. We conducted a second user study to collect statistics on actual DAB data (RQ1—*actual*). It also enabled us to collect preliminary insights on individuals' willingness to share their DAB data, as well as the financial value they attach to it–in the context of online surveys.

Although in Study #1 we targeted EU residents to ensure the relevance of the findings w.r.t the GDPR framework, for Study #2, our objective shifted toward users' DAB data and willingness to share it (GDPR was not relevant here), for which we prioritized access to a larger pool with native English proficiency, and thus targeted US residents. We acknowledge this sampling decision as a limitation (Section 4.1.3).

We first deployed a screener survey (1 min., USD 0.20) to select respondents who actively manage a DAB using Google Contacts for the main survey (2-5 min. expected duration; baseline payment of USD 1). The transcript of the main survey questionnaire is provided as a supplementary material.[8] Here, we summarize the key blocks of the survey. In the main survey, using the consent form (S2.B2), respondents were first informed that the goal of the survey was to collect *statistics* about (completeness of) their DAB data (e.g., "the proportion of their contact cards that include a birthday"). Then, they were given the choice to either (1) "Grant [us] access to [their] Google Contacts data", (2) "Manually respond to the survey questions about [their] Google Contacts data", or (3) "Withdraw from the survey" (see S2.Q3 for the exact and complete phrasing of the options). For each option, they were informed about the number of questions and the amount of time required to complete the survey. They were also informed about the additional bonus payment (i.e., incentive) they would receive if they chose to grant access to their Google Contacts data (i.e., Option 1), the data that we would collect (in a table; see [87][Fig. 9a]), and how we use it (i.e., academic research).

To evaluate how monetary compensation affects willingness to share data–in the context of online surveys–, we varied the financial incentive across survey batches. Thus, we conducted the study in *multiple* successive batches of approximately 120 respondents, each with different incentive amounts. After each batch, we decided whether to proceed with a new batch and, if so, on the new value of the incentive. To determine the new value, we considered the base value of the incentive, results from the literature [18, 24] and the respondents' input (S2.Q11).

Respondents who chose to grant access did so by using the DDS[12] platform (S2.B5). This enables researchers to person-

[12] See https://github.com/DataDrivenSurveys, visited: May 2025.

alize surveys with the data extracted from the respondents' online accounts, with their consent [87]. Details regarding data access, platform practices, and ethical considerations related to the use of DDS are provided in Section 7. Afterwards, they were shown the statistics that were extracted from their data (S2.B6).[13] Lastly, we asked respondents to briefly explain why they chose to grant access to their Google Contacts data rather than respond to the questions manually (S2.Q4).

Respondents who chose to respond manually were asked questions about their DAB data (see S2.B7; similar to those asked in Study #1, e.g., S1.B5). Then, they were asked to explain why they chose to manually respond to the questions instead of granting access to their data (S2.Q10). They were also asked how much they would need to be paid to be willing to grant access to their Google Contacts data (S2.Q11). For those who reported that they would never be willing to do so, they were asked what makes them feel comfortable sharing their contacts data with DAB-SPs such as Google but not with researchers (S2.Q12). All respondents were asked to provide their demographics (i.e., IUIPC-8 and gender).

Respondents who chose to withdraw were all liberated immediately. We invited them a few days later, via Prolific, to a brief, optional, follow-up survey where we asked only about their reasons for withdrawing (S2+.Q3; 2 min, USD 0.60). Respondents who just left the survey were not recontacted. See Section 7 for the ethical considerations of this procedure.

4.1.2 Data Analysis

For close-ended questions, we primarily used descriptive statistics to summarize the data. For some of them, we also used statistical tests, for comparison purposes. For open-ended questions, we used a reflexive thematic analysis [14] to analyze the responses. Using MAXQDA, one of the co-authors iteratively coded the responses inductively, thus refining the codes and organizing them into categories that were grouped into overarching themes. While the survey responses were generally concise, this approach allowed us to capture recurring ideas and perceptions relevant to our research focus. We acknowledge that the limited length and depth of responses constrain the richness of interpretation, which we discuss in the Limitations section. Also, consistent with Braun and Clarke [14, 15]'s approach, we did not employ multiple coders or calculate inter-coder reliability, as the method emphasizes the active role of the researcher in knowledge production rather than coding reliability.

All respondents answered the open-ended question about the proposed design (S1.Q20). However, only 199 respondents submitted an optional drawing (S1.Q21). Out of these, 26 were excluded due to low-quality submissions, irrelevant or unclear content, or clear use of generative AI. For the drawing task (S1.Q21), we first conducted a quality check to exclude low-quality and irrelevant submissions. We analyzed these drawings [55] alongside the open-ended responses to S1.Q20. After coding the open-ended responses (S1.Q20), we reviewed the associated drawings (S1.Q21) to identify any additional features or functionalities not captured in the text responses. If new features were found, they were added to the codebook. Ultimately, a unified codebook was developed for both S1.Q20 and S1.Q21. The codebook for open-ended questions and drawings is available on OSF.[8]

When reporting open-ended answers, we use the following determiner-to-percentage mapping for consistency: *very few* for $1-10\%$, *some* for $11-30\%$, *a substantial number* for $31-50\%$, *more than half* for $51-70\%$, *most* for $71-90\%$, and *almost all* for $91-100\%$ of respondents.

4.1.3 Limitations

Our studies are subject to a few limitations. One of them is the use of samples from different countries across Studies #1 and #2. However, we verified that key behavioral patterns, such as DAB completeness, remained consistent (see Figure 6 and Figure 10), suggesting that this limitation does not critically impact our findings. Also, our qualitative data in both studies were based on brief open-ended survey responses, limiting the depth and richness associated with reflexive thematic analysis. Next, although we used a commitment question and an evaluation of open-ended answers for quality insurance, we did not use attention checks; hence, our final dataset might still contain some responses from careless respondents [58].

Regarding Study #2, one limitation is the potential for priming, as the consent form explicitly mentioned privacy perceptions. This might have made some respondents more privacy-aware than they might be in everyday contexts. Lastly, respondents' willingness to share their DAB data may have been influenced by multiple overlapping incentives—namely, the bonus payment and the opportunity to complete the survey more quickly. This overlap is consistent with the privacy calculus framework [51], which acknowledges that users often disclose data in exchange for various perceived benefits—not only money, but also time savings or access to functionality. Also, the decision to share was made in a specific context: a user survey conducted by a university for academic research.

4.2 Results

Study #1. We deployed the screener to 1255 respondents, of which 626 were eligible (i.e., they actively manage a DAB using Google Contacts or Apple iCloud Contacts)[14], then we invited those eligible for the main survey. A total of 498

[13] If a respondent completed the process but their DAB data contained fewer than five contact cards, we contacted them and offered them the possibility to re-grant access with a different Google account, in case they connected with an account that they do not use for managing their contacts.

[14] Google and Apple were the most frequently-used services (39.0% and 33.9% resp.), thus justifying our choice for inclusion and illustrations.

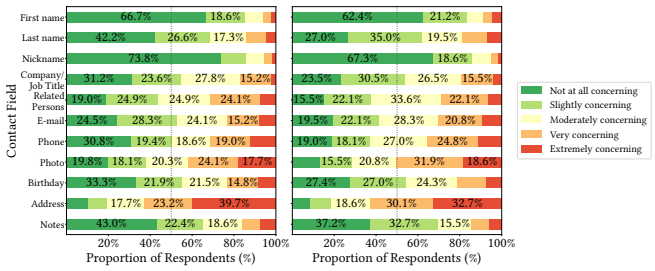
started the main survey, and we obtained a total of $N = 463$ valid responses (i.e., finished, consented, committed to providing thoughtful answers, confirmed their answers from the screener). The median completion time was 18 min and 38 s.

Demographics. The sample of respondents was diverse and relatively balanced in terms of age ($M = 31.9$, $SD = 9.1$, $Min = 18$, $Max = 73$) and gender (Man: 50.5%, Woman: 47.7%, Non-Binary: 0.9%, Self describe: 0.4%), and its distribution of IUIPC score ($M = 5.8$, $SD = 0.8$; on a scale from 1 to 7) and VOPP score ($M = 5.1$, $SD = 0.8$; on a scale from 1 to 7) were relatively high. As for country of origin, 70.0% of the (valid) respondents were from Germany and 30.0% from the Netherlands. For their roles, which determined the phrasing of some of the questions, 51.2% were assigned the 'subject' role and 48.8% the 'user' role.

Usage. We asked respondents about the platforms (dedicated phone/tablet/computer app and web browser) through which they access their DABs (S1.Q8). Almost all the respondents reported accessing their DABs from their phones (only 0.9% reported 'never'). Whereas, a large proportion of the respondents (41.7%) reported never accessing their DAB from a web browser. Also, a substantial proportion of the respondents (29.4%) reported accessing their DAB from a single platform (their phones). Although we collected self-reported data about the completeness of the respondents' DAB (S1.Q7), we do not report it here.[15] Instead, in the next section, we will report the (same) data collected from Study #2 (see Figure 6), as we can compare it to the data extracted from the actual DAB data of the respondents (from the same study) who granted us access to their Google Contacts account. We did not observe major differences between the self-reported data collected in Study #1 and Study #2.

We asked respondents how many mobile apps (i.e., third-party apps or TPAs) currently have access to their DAB on their smartphones (S1.Q9/S1.Q10). The vast majority of respondents (90.5%) reported having at least one such app installed and, among these, 33.4% they had ten or more apps installed on their smartphones. Regarding the types of such mobile apps, a total of 1742 apps were reported. After combining manual and automatic analyses to handle typos and discrepancies in app names between iOS and Android, we obtained a total 199 unique app names. These apps spanned a wide variety of categories (23 in total).[16] Communication apps (WhatsApp, Gmail, Telegram)[17] were the most frequently reported category, accounting for 48.4% of responses. Social networking apps (Instagram, Snapchat, Facebook) were the second most reported category, representing 13.9% of responses. The next three most reported categories were Productivity (Calendar, Google Drive, Microsoft Outlook) with 9.2%, Finance (PayPal, Revolut, N26) with 5.9%, and Travel & Local (Google Maps, Maps (iOS), Polarsteps) with 5.2%.

[15] We depict the data distribution in the appendix; see Figure 10 on page 19.
[16] We relied on the categories from the Google Play Store.
[17] We list the three most frequently reported apps in descending order.



(a) subject ($n = 237$) (b) user ($n = 226$)

Figure 2: Distribution of the respondents' level of privacy concerns for different contact fields (source: S1.Q14/S1.Q15). The median can be determined using the line drawn at 50%.

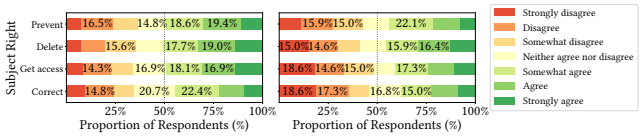
Understanding. When asked whether the DAB-SPs could technically access (in the clear) the data users store in their DABs (S1.Q12/S1.Q13), 80.8% expressed some agreement (i.e., at least *somewhat agree*). As DAB-SPs can access this data, as long as E2EE is not used, these results show a good level of understanding and awareness among the respondents.

Concerns. When asked about their privacy concerns regarding the different fields included in contact cards, respondents expressed overall modest, yet diverse, concerns. The results are depicted in Figure 2. The two fields that raised the highest levels of concern are 'Address' and 'Photo', with 62.9% and 46.0% of the respondents who expressed being very or extremely concerned, respectively.

We observed the same trend for the two different roles (i.e., 'subject' vs. 'user'). This tends to indicate that the phrasing of the questions (second person for 'subject' S1.Q14 vs. third person for 'user' S1.Q15) did not have a noticeable effect on the privacy perceptions of the respondents.

Opinions. When asked about the rights of subjects whose personal information is stored in users' DAB (i.e., prevent others from storing information, delete the information stored, obtain access to the stored information, correct the inaccurate stored information; see S1.Q16 and S1.Q17), the respondents were slightly more inclined to agree overall, with levels of agreement of 46.4%, 47.9%, 45.1%, and 41.3% respectively (i.e., at least *somewhat agree*). The results are depicted in Figure 3. The levels of agreement were somewhat comparable across the four different rights and, as expected, they were overall higher in the 'subject' role, even though the medians were the same (except for 'Delete'). When shifting the respondents' perspective through different phrasing ('subject' $\leftrightarrow$ 'user'), we also observed that respondents agreed slightly more when taking the perspective of a subject. This is unique to DABs as for typical online services, respondents are usually asked only about their opinions regarding their *own* rights; here, due to the IDP aspects of DABs, it is not the case.

Proposed Design. The respondents were asked to envision a dashboard that allows subjects to access and control



(a) subject ($n = 237$) (b) user ($n = 226$)

Figure 3: Distribution of the respondents' levels of agreement for the different data subject rights. (source: S1.Q16/S1.Q17).

their personal information stored in users' DABs (S1.Q20 and S1.Q21 [optional]). More than half the respondents suggested options for *managing the dashboard*. Some suggested that the government should manage this and emphasized trust in governments to oversee a centralized system supported by data protection laws. Some, however, preferred an independent, neutral organization that works in collaboration with governments, regulatory bodies, and DAB-SPs, to ensure cross-platform and cross-border compatibility. Some proposed that DAB-SPs or mobile OS developers (e.g., Apple for iOS and Google for Android) should manage this. As DAB management has little to do with the OS, except for the dedicated mobile permissions, and because both Apple and Google provide online DAB services, this last option is not ideal. Some respondents stressed the need for cross-platform compatibility (e.g., across DAB-SPs and countries), reliable identity verification to validate subjects and users, and subjects acting as the authoritative source by providing and validating their own accurate personal information (see F1 in Figure 11).

Several *core functionalities* were proposed. Regarding *information access and control*, more than half of the respondents emphasized the ability of subjects to view personal information about them stored by others and edit/delete it—including correcting inaccuracies (see F2 in Figure 11). Opinions diverged on whether edits/deletions should require users' consent or if they should be mandatory. Note that such a functionality creates a direct tension between the subject's right to access their information and the user's right to privacy. Indeed, some respondents expressed *ethical and privacy* concerns. This trade-off can be navigated by fine tuning the granularity of the information displayed, e.g., *"Two users included your birthday in their digital address books."* vs. *"Alice included your birthday as 'Jan. 1st, 2000' and Bob included it as 'Jan. 1st' in their digital address books."*. This is particularly acute for personal notes. A substantial number of the respondents suggested various *ways to access and visualize* stored information (see F3 in Figure 11), including displaying details about who stored which information and variations of stored data (e.g., name or phone number formats), summaries such as counts of mutual contacts or incorrect entries, with only a few advocating the anonymization.

Some respondents highlighted the importance of *managing access* (see F4 in Figure 11), including restricting, with the flexibility to change them over time, certain personal information for specific users and setting preferences for information sharing. This functionality resembles that of social networks. Very few also highlighted *distinctions between natural persons and legal entities* (i.e., individuals versus organizations storing the data) and a substantial number of them expressed a particular interest in identifying which DAB-SPs (e.g., Google Contacts) and TPAs (e.g., WhatsApp) have access to their accounts (see F5 in Figure 11).

A substantial number of the respondents called for *notifications* to inform subjects when their information is stored (including incorrect entries), updated, or breached and to notify users about subjects' actions on their stored information (see F6 in Figure 11).

Study #2. We deployed the screener to a total of 2700 respondents, of whom 889 were eligible (i.e., they actively manage a DAB using Google Contacts and have an Android or iOS smartphone). We contacted eligible respondents in batches, hoping that about 120 complete responses per batch. In total, 619 respondents began the survey and 574 of them agreed with the terms of the study (S2.Q2), thus reaching the question that asked them to choose how they want to share information about their Google Contacts (S2.Q3). We collected complete and valid data (i.e., finished the survey, and—in the case where they chose to grant access to their Google Contacts account—successfully granted access and had at least five contacts in their DAB) for $N = 459$ respondents. The median completion time was 5 min and 23 s.

Demographics. The sample of respondents was diverse and relatively balanced in terms of age ($M = 40.5$, $SD = 11.7$, $Min = 18$, $Max = 77$) and gender (Man: 47.4%, Woman: 50.7%, Non-Binary: 1.7%), and its distribution of IUIPC score ($M = 6.0$, $SD = 0.8$) was relatively high and comparable to that of Study #1.

Choice. We first look at the choices (i.e., 'grant access', 'manual', 'withdraw') made by *all* the respondents, *including* those who later left the survey before the end (i.e., 'not finished'). Figure 4 depicts, for each batch, the proportions of respondents who made each of the possible choices. For each choice, we further broke down the proportions, depending on whether the respondents finished the survey and, in the case where they granted access to their Google Contacts, whether they had at least five contacts. The results are split by batch. Each batch corresponds to a different value of the *additional* financial incentive associated with the respondents' choices of granting access to their Google Contacts data (i.e., bonus of 0, 1, 2, or 5 USD). We found that even without any *financial* incentive,[18] 19.9% of the respondents chose to grant access to their DAB data, but that *only* 10.6% actually completed the questionnaire and had five or more contacts in their DAB. There are several possible reasons for this disparity,

[18] There exist other incentives for this option, incl. a reduced response time.

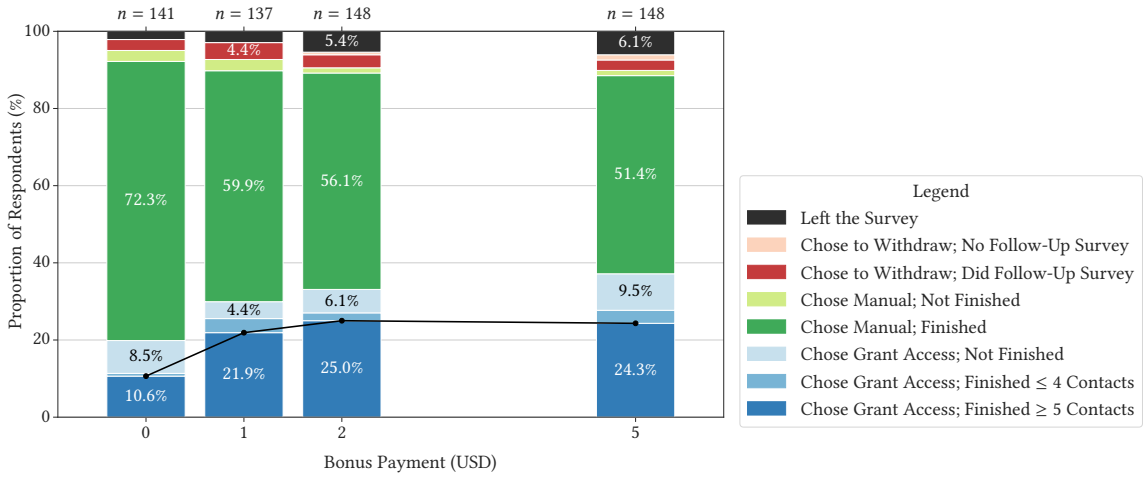


Figure 4: Distribution of the choices made by the respondents regarding how they wanted to share information about their Google Contacts with us in Study #2, depending on the value of the bonus payment (source: S2.Q3). We depict the breakdowns between the different choices (Grant Access, Manual, Withdraw, or simply Leave without making a choice) and complete/valid (darker) and incomplete/invalid (lighter) responses. We elaborate on the underlying reasons in the "Reasons behind Choice" sub-section.

including their having second thoughts, not *actually* having access to their Google credentials, not having a Google Contacts account, and their connecting to a secondary Google account with few contacts. Unsurprisingly, the proportion of respondents who make the choice to grant access to their Google Contacts data increases with the value of the financial incentive associated with it. However, the proportion of respondents who make this choice *and* finish the survey with valid data appears to plateau slightly below 25%, with only a marginal increase when increasing the incentive from 1 to 5 USD. These results provide an estimate of the value users attach to their DAB data (i.e., a few USD)—in the context of answering online surveys with DAB data— and, to some extent, to their privacy and that of their contacts.

DAB Data. We look at the size and completeness of the respondents' DAB data. We considered only the data, aggregated over all the batches, of the respondents who completed the survey and, for those who granted access to their Google Contacts, who had five or more contacts. For respondents who granted us access, we have *actual behavioral* data, whereas for the others, we have only *self-reported* data (S2.Q6). The median number of contacts is 69 ($M = 159.4$) for the respondents who granted us access and 78 ($M = 260.1$) for those who responded manually. The distributions are shown in Figure 5. Despite the large number of contacts, prior research shows that users often communicate with only a very small subset of their contacts [7]; the (interdependent) privacy costs come with almost no utility benefits.

Figure 6 depicts, for each of the main fields (first name, last name, photo, etc.), the distributions of the proportions of the respondents' contact cards for which a value is specified for this fields. The actual data is represented as violin box-plots (left-hand side). It can be observed that the median proportion of contact cards with a photo is 3.0%. This means that *half* of the respondents include a photo in *more than* 3.0% of their contact cards. Note that the mean proportion across respondents (not visible on the graph) is 8.0%. The numbers are lower for some fields such as birthday ($M = 3.6\%$, $Med = 0\%$) and address ($M = 7.7\%$, $Med = 1.7\%$). Unsurprisingly, contact cards almost always contain a first name and a phone number (which constitutes a unique identifier) and quite often a last name. This corresponds to the typical (basic) usage of a *phone* DAB [7]. Looking at the outliers (i.e., 95th percentile), it can also be observed that some respondents have very complete contact cards in their DABs, with more than 31.8% of their contacts cards including a photo, 20.7% a birthday, and 36.6% an address. The self-reported data is represented as stacked histograms (right-hand side).

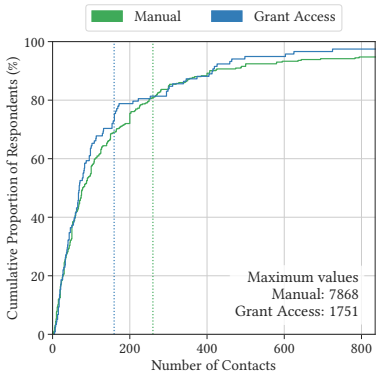
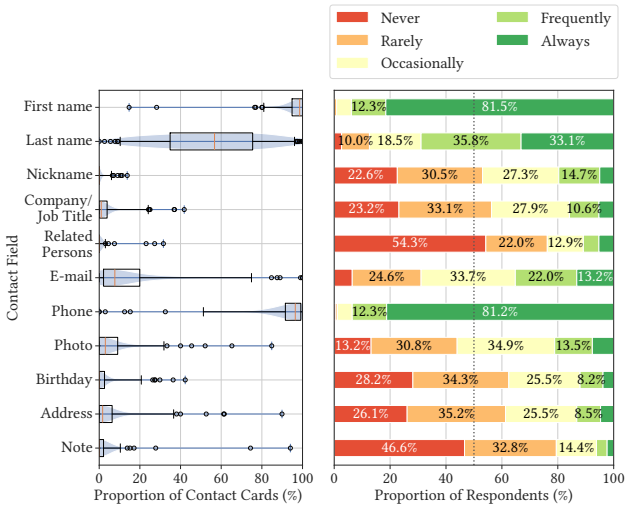


Figure 5: Distribution of the total number of contact cards in a DAB (CDF): extracted from the respondents' Google Contacts account or self-reported by the respondents.



(a) Actual data ($n = 118$) (b) Self-reported data ($n = 341$)

Figure 6: Distribution of the proportions of contact cards with a value specified for a given field (first name, last name, *etc.*): extracted from respondents' Google Contacts account (left) or self-reported by the respondents (source: S2.Q6) (right).

Surprisingly, the self-reported proportions are substantially higher than those computed from the actual data; yet, they show similar patterns. For instance, for birthdays, only 28.2% of the respondents reported never including them, whereas in the data extracted from actual DABs, this was the case for more than half of the respondents.

Reasons behind Choice. To understand users' perception of their DAB data, we look further at the reasons behind the choices made by the respondents (collected in S2.Q4, S2.Q10, and S2+.Q3 respectively). First, we looked at the differences, in terms of privacy scores (IUIPC [33, 54] and VOPP [36]) between the respondents who chose to grant access to their Google Contacts data and those who chose to respond manually. The distributions of the scores are depicted in Figures 8b and 9b in Appendix B. We observed a noticeable difference in the IUIPC score between the two groups ($M = 5.6$ for 'Grant Access' and $M = 6.2$ for 'Manual', but none for VOPP; this seems to indicate that only concerns about *their own* privacy are associated with the respondents' choices to grant access to their DAB data.

For the respondents who chose to *grant access* to their Google Contacts data (S2.Q4), the most common reason was *convenience*; most of the respondents noted that the procedure is easier, quicker, and more efficient (compared to manually navigating their contacts). Next, a substantial number of the respondents mentioned the *financial compensation* offered. Additionally, some respondents mentioned the reason they *did not refrain* from granting us access. They explained that they did not see any privacy issues with granting access to their contacts because (1) they trusted us (*"I know the data is safe in your hands."* [NB, 36 y.o.]) and/or (2) they thought the data was not sensitive (*"I wasn't really worried about it. I don't have any secret contacts."* [W, 66 y.o.]). This last point shows that some respondents have low awareness/concerns regarding (interdependent) privacy.

For respondents who chose to answer *manually* (S2.Q10), *privacy* emerged as the most common reason—for more than half of them: *"I like to keep my privacy."* [W, 43 y.o.]. Among the respondents with privacy concerns, some mentioned concerns about *interdependent* privacy. They stated they were uncomfortable sharing their DAB data, especially as it contains information about *others* who did not consent to the sharing: *"I don't want to reveal other people's private contact information without their permission."* [M, 45 y.o.]. Although this perspective is valid, it is worth noting that those individuals likely did not consent to sharing their information with Google either—and possibly with third-party apps, which is a prevalent practice as shown in Study #1 (i.e., S1.Q9).

Finally, among the 21 respondents who chose to *withdraw* from the survey, 19 completed the follow-up survey (S2+.Q3). Most of them mentioned *privacy concerns*, expressing discomfort with providing information about their contacts.

Expected Extra Compensation. For the respondents who chose to respond manually, we look at the value of the extra financial compensation that convinced them to choose the 'grant access' option instead of the 'manually respond' one, according to them (i.e., self-reported, S2.Q11). 32.3% of these respondents (36.3%, 32.9%, 34.6%, and 23.7% respectively, in the successive batches) indicated that they *would consider* granting access to their DAB data for an extra compensation.

The medians of the amounts reported by the respondents who would consider this, for the different batches, were 12.5, 10, 35, and 20 USD, respectively, in the successive batches. The overall median amount requested was 15 USD. These amounts were in line with the values reported in similar contexts (e.g., 25 USD for offline own PII such as age and address [18]). Also, similar to a previous work, the amount that individuals usually demand can vary widely: Some respondents report exaggeratedly high amounts, likely indicating *in fine* their unwillingness to share data ($Max = 500000$ USD in our study cf. $Max = 100000$ EUR in the context of IoT wearables [27]). However, it should be noted that these findings exclude the respondents who agreed to grant access to their data for only a little extra compensation (0, 1, 2, and 5 USD for the successive batches).

For those who responded they would never share their Google Contacts data, we look at the reason they feel comfortable sharing their contact data with, on the contrary, the DAB-SP, namely Google (S2.Q12). A substantial number mentioned Google's established reputation, as well as their trust in its privacy, security, and accountability policies. *"Google has a reputation of safeguarding user information and is a well-known and established company."* [W, 65 y.o.]. Among

them, very few respondents expressed a lack of trust towards researchers: *"Because Google is Google. You are a random researcher on the Internet."* [M, 44 y.o.]—which is understandable in the wake of the Facebook-Cambridge Analytica data scandal [17]—, whereas some acknowledged the apparent contradictory nature of the situation: *"I've never actually thought about it like that."* [W, 36 y.o.]. Some respondents shared their reluctant acceptance of Google's access to their contact data, as many expressed a sense of inevitability in sharing data with Google in general due to the prevalence of Google and the utility of DABs. *"Google already has all my information."* [W, 30 y.o.]; *"It is more or less a necessary evil to have functionality and ease of use."* [W, 37 y.o.].

5 Discussion

Legal Implications Determining the roles (i.e., data subject, controllers, and processors) of the different parties involves complex legal analyses. The limited availability of information makes this process time-intensive, even for legal experts, thus impractical for lay individuals. This asymmetry of power and information between individuals and SPs is well-documented [86]. A potential solution is standardized reporting by DAB-SPs to disclose their actual data-processing practices, including for non-users, thus enabling greater transparency. Also, in cases of joint controllership, the GDPR mandates shared responsibility between DAB users and the DAB-SPs, potentially complicating the assignment of responsibilities and the enforcement of data subject rights. Clearer guidelines are thus needed to delineate the responsibilities of each party. DAB-SPs must also meticulously review their data processing operations to facilitate DARs. DAB-SPs might re-purpose user-provided data, which does not align with users' expectations. This mismatch is not a new issue. A notable example is the creation of shadow profiles: profiles of non-users generated using data provided by users (e.g., Facebook) [9, 10, 73].

Reflections on the Results of the User Studies Our results from Study #1 showed that a substantial majority of respondents had granted to third-party mobile apps access to their DAB. These apps spanned a wide range of categories, with communication and social networking apps being the most common. While such widespread access raises clear privacy concerns, it is important to acknowledge that many of these apps have legitimate functional needs for DAB data. For example, instant messaging and social media apps use contact information to build friend lists, peer-to-peer payment apps rely on phone numbers to identify recipients, and travel applications may use stored addresses for route planning or location-based services. These utility-driven use cases help explain why users are relatively permissive in granting access, even when they are aware of the associated privacy risks.

However, the implications of such access are compounded by the interdependent nature of DAB data, as revealed in Study #2. The analysis of contact card completeness showed that many sensitive fields—such as photo, birthday, and address—are *sparsely* populated across individual users' DABs. For instance, only 3.6% of contact cards, on average, include a birthday. Yet, it should be noted that only *one* user needs to enter the birthday of an individual (a data subject) in their DAB for it to be known to the DAB-SP. Therefore, even if there is only a 3.6% chance that a user includes the birthday of a given individual in the corresponding contact cards in their DAB, if this individual appears in the DABs of 50 users, then the probability that the DAB-SP knows their birthday is $100\% - (100\% - 3.6\%)^{50} = 86\%$. This probabilistic accumulation illustrates how IDP breaches can arise not from any one user's behavior, but from the collective actions of many—posing a fundamental challenge for privacy protection.

The tension between perceived utility and collective privacy risk is further emphasized when financial incentives were introduced in Study #2. Among those who declined to grant access to their DAB data, only 32.3% reported that monetary incentives would convince them to grant access. This is in contrast with Study #1 findings, where 90.5% of the respondents reported granting access to their contact data to at least one third-party app. The discrepancy suggests that user motivations are highly context-dependent [62]. In real-world app context (Study #1), users may perceive tangible utility—which satisfies data sharing. In contrast, in the context of survey taking, when the only perceived benefits are financial (and marginally usability), users are more hesitant.

Tensions between Users and Subjects GDPR provides exceptions to its application, such as activities of a natural person conducted in a "purely personal or household activity". This exemption, though relieving Alice of compliance burdens, limits the ability of John to exert his data-protection rights. As shown in Study #1, certain individuals appear to support the implementation of new techno-legal approaches that challenge the boundaries of this exemption. We deem such measures necessary, especially in situations where DAB is synced online, no E2EE is used, and DAB-SP uses the data (i.e., Scenario 4). These measures would enable John, (1) to prevent, for instance, Alice from inputting into DAB certain personal information about him and/or (2) to obtain some information regarding the data about him that users store in their DABs. However, implementing these measures would require deviating from the status quo regarding the balance of interests. Indeed, such measures could impose a burden on Alice and create potential tension between the rights of John and the autonomy/utility of Alice. For example, granting John the right to prevent Alice from storing her contact information could directly conflict with Alice's ability to efficiently manage her DAB. Additionally, providing John access

to information stored about him in Alice's DABs could raise privacy concerns for Alice: for instance, if John could access to the note composed by Alice in the contact card about him.

There are situations, however, where such measures would not be needed in the first place. One such situation is when DAB-SPs do not use the data for their own purposes (Scenario 3): This does not provide more control to John (e.g., right to object), but it does provide him with guarantees regarding the use of the data that concerns him. Essentially, only Alice can use the data, as she would do with a physical address book. An even better situation, in our opinion, is when DAB-SPs offer E2EE to Alice (Scenario 2), and that Alice activates it. This imposes only small constraints on Alice and it protects not only the privacy of John but also that of Alice.

Design Implications To illustrate how the current DAB ecosystem can be enhanced for greater subject privacy and to offer options for DAB users to (at least partially) preserve their autonomy/utility, we present a series of design ideas.

For instance, a large majority of respondents reported accessing their DABs primarily through mobile phones, with some of them relying solely on a single device. This has strong implications for E2EE, as usability concerns commonly tied to web-based decryption may not apply for those who never use web access, making E2EE feasible. Similarly, for users relying on a single device, storing the data on the DAB-SP's server is only useful for backups (as synchronization between devices and web access is not needed), hence E2EE would come at almost no cost. Additionally, when asked respondents to envision a dashboard that would allow them to access and control personal information stored about them in others' DABs, respondents proposed a range of solutions resembling decentralized data management systems like pods (e.g., [81]) or online social network profiles (e.g., Facebook, LinkedIn) where people compile personal information about themselves and grant others access to it. This model supports user agency and data accuracy but does not fully eliminate the problem, as users could still *add* new information to (or *edit* existing ones in) their local copies of their contacts' profiles. While some of these proposed mechanisms mirror solutions from related domains, such as IoT or mobile privacy, their application to DABs introduce unique challenges. This stems from the structured nature of DAB data (i.e., easily exploitable due to standardized fields), its duplicative storage (i.e., redundantly entered across users), its involuntary inclusion (i.e., data added without subjects' awareness or consent), and its inherently interdependent privacy dynamics (i.e., the privacy of one is affected by the action of others). Any viable design solution, must consider these unique aspects of DAB ecosystem.

Building on these insights, we now outline a set of steps John (data subject) would need to take to set privacy preferences and the corresponding effect this would have on Alice (DAB user). John begins by visiting a dedicated webpage (or app) managed by the DAB-SPs or by a centralized entity. He "authenticates" by proving ownership of one of his identifiers (e.g., e-mail addresses or phone numbers). Authentication is performed using a secure mechanism (e.g., one-time password (OTP) [61]) where a code is sent to John and then entered on the webpage to complete the verification process. Once authenticated, John can define the types of information that he does not want to be processed by the DAB-SP. For instance, John might opt out for his birthday but not for his job title.

Regarding the effect this would have on Alice, we envision two cases: (1) Once John has defined his preferences, the DAB-SP enforces these settings during synchronization. For example, if Alice adds John's birthday and job title to her contact card, the birthday will be excluded from synchronization, meaning Alice must manually re-enter the birthday on other devices if she wants to access it everywhere. Similarly, the birthday will not appear in the web app. (2) The DAB-SP employs E2EE (e.g., Advanced Data Protection for iCloud). In this case, the excluded fields are still synchronized but are encrypted. This ensures that this information is unavailable to the DAB-SP, and consequently—most probably—in the web app or any platforms that lack decryption capabilities.

6 Conclusion

Our work highlights the IDP challenges of DABs, thus addressing a critical research gap in IDP/DABs' legal, technical, and user-centric implications. We contribute legal insights, showing that individuals whose data appears in other's DAB qualify as data subjects but face significant obstacles in exercising their rights under GDPR. Despite recognizing IDP risks and expressing concerns about certain data types, many users still allow third-party apps access and are willing to share DAB data for minimal compensation. We propose actionable steps for DAB-SPs to enhance privacy compliance, and we outlined privacy-enhancing designs. Yet future research should focus on participatory design to co-create solutions and evaluate them. By raising awareness and offering design ideas, we enable a future where individuals gain greater control over their personal information in the DAB ecosystem.

Acknowledgments

The authors are grateful to Prof. Valérie Junod and Evanne Anthoine-Milhomme Phan for their preliminary investigation of the problem from a legal perspective and to Holly Cogliati for editing the paper. The authors also thank Lahari Goswami, Yesle Kim, and Oleksandr Velykoivanenko for participating in the cognitive pretests. The work was partially funded with grant #2024-09-24-173 from the Hasler Foundation.

7 Ethical Considerations

Our research received IRB approval. All respondents provided informed consent detailing the study's purpose, data handling, withdrawal procedure, and compensation. We followed Prolific's guidelines, compensating above the minimum recommended rate of GBP 9 per hour, consistent with ethical norms for fair compensation.[19] There are two primary stakeholders: (1) respondents, who were asked to respond to the survey and, in some cases, share DAB summaries; and (2) DAB-SPs, contacted via DARs. We assessed risks across these groups. We minimized respondent risk through transparent communication and secure data collection. While DAB-SPs faced no legal or privacy risks, we address ethical concerns around our DARs. Next, we discuss various ethical considerations.

Contacting DAB-SPs. We contacted five large DAB-SPs with GDPR-based DARs to explore how they assess interdependent data protection situations in practice—particularly how their corporate legal compliance teams view and treat IDP situations. For one DAB-SP (Proton), even though it uses E2EE, there was uncertainty (when sending DARs) around its legal roles due to technical unknowns (e.g., whether contact cards were E2E-encrypted and whether all users enabled E2EE). Since email addresses, as we later learned, are not encrypted, Proton could have answered our query ("How many contact cards in your database contain my email address?"). Similar ambiguity existed with other DAB-SPs, whose public documentation did not clarify whether they act as controllers, processors, or joint controllers. Thus, our DARs were carefully formulated to avoid requesting access to sensitive or encrypted user content. Also, even if a DAB-SP concluded it was not a controller, our limited inquiries—only 2–3 e-mail exchanges—posed no harm. Our intent was not to coerce responses but to explore real-world inconsistencies and responses within a gray area of legal interpretation. Unlike large-scale DAR campaigns,[9] our outreach was narrow, transparent, and focused, with minimal burden on recipients. We view this as an ethically appropriate legal research practice that helps clarify underexplored regulatory boundaries.

DDS Platform. In Study #2, respondents could choose to share structured summaries of their DABs using DDS [87]. The consent form specified that only aggregate statistics (i.e., the proportion of contact cards with certain fields) would be collected—no raw data would be stored. DDS uses OAuth-based, granular consent: respondents authorize access to specific data categories, preview what will be accessed (i.e., see [87][p. 15, Figure 9a]) and computed (i.e., see S2.B6), and may opt-out at any point (i.e., see S2.B2). All raw data is processed temporarily in memory to compute predefined variables (e.g., number of contacts with e-mail addresses), which are then uploaded to the survey platform (Qualtrics); researchers access only these variables. DDS aligns with data donation principles [11][20] while emphasizing user control and transparency and adding safeguards via API scopes and selective access. To reduce re-identification risk, we excluded DABs with fewer than five contacts. DDS does not otherwise differentiate based on the dataset size, but we explicitly evaluated the sensitivity of small summaries and found minimal risk when dissociated from respondent identity.

Treatment of Withdrawn Respondents. Those who withdrew from the main survey had their survey session immediately terminated, and no data from these sessions was retained or analyzed (except for the count of withdrawals). Only their Prolific IDs, completion statuses (i.e., "withdrew"), and timings remained accessible on Prolific for standard respondents tracking. Respondents also received distinct completion codes reflecting their paths through the survey. For example, participants who granted access to their Google Contacts data received a different code than those who preferred to withdraw. While this deviates slightly from Prolific's typical use of a single, uniform completion code per study, it enabled us to accurately manage the study. Using these completion codes, we were able to later send an invitation for a short, optional, follow-up study focused only on their decision to withdraw. This follow-up was conducted as a separate study with independent consent. While this procedure complies with Prolific's standard recruitment model and involved no reuse of personal data without re-consent, we acknowledge that our original consent form could have more clearly communicated the possibility of such follow-up contact.

Positionality Statement. Our team is composed of computer scientists, covering both technical and user-centric aspects in the field of security and (interdependent) privacy, and of legal scholars, specialized in digital law and data protection. One of the legal researchers is a practicing attorney focused on data protection and IT law. The team is based in Europe.

8 Open Science

In compliance with the research transparency criteria [77], we include the following materials in the OSF repository:[8] the script of e-mails sent for DARs by the researchers, the survey transcripts for Study #1 and Study #2, a redacted and de-identified version of the survey data, the codebook for open-ended responses, and example visualizations of respondents' drawings. Lastly, regarding the responses to our DARs, for those service providers who agreed, we shared their responses on OSF. After acceptance, we sent them a follow-up request with the camera-ready version of the paper attached.

[19] See https://researcher-help.prolific.com/en/article/2273bd.

[20] See https://datadonation.eu/data-donation/.

References

[1] A. Alshehri, J. Spielman, A. Prasad, and C. Yue. Exploring the Privacy Concerns of Bystanders in Smart Homes from the Perspectives of Both Owners and Bystanders. *Proceedings on Privacy Enhancing Technologies (PoPETs)*, 2022. doi: 10.2478/popets-2022-0064.

[2] Apple. Accessing the contact store, 2025. URL https://developer.apple.com/documentation/contacts/accessing-the-contact-store. Last visited: Jan. 2025.

[3] Apple. iCloud data security overview - Advanced Data Protection for iCloud, 2025. URL https://support.apple.com/en-us/102651#advanced. Last visited: Jan. 2025.

[4] Article 29 Data Protection Working Party. Guidelines, 1996. URL https://www.edpb.europa.eu/our-work-tools/general-guidance/endorsed-wp29-guidelines_en.

[5] J. Ausloos and P. Dewitte. Shattering one-way mirrors – data subject access rights in practice. *International Data Privacy Law*, 2018. doi: 10.1093/idpl/ipy001.

[6] S. Barocas and K. Levy. Privacy Dependencies. *Washington Law Review*, 2020.

[7] F. R. Bentley and Y.-Y. Chen. The Composition and Use of Modern Mobile Phonebooks. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*, 2015. doi: 10.1145/2702123.2702182.

[8] G. Biczók and P. H. Chia. Interdependent Privacy: Let Me Share Your Data. In *Proc. of the Int'l Conf. on Financial Cryptography and Data Security (FC)*. 2013. doi: 10.1007/978-3-642-39884-1_29.

[9] V. Blue. Anger mounts after Facebook's 'shadow profiles' leak in bug, 2013. URL https://www.zdnet.com/article/anger-mounts-after-facebooks-shadow-profiles-leak-in-bug/. Last visited: Jan. 2025.

[10] V. Blue. Firm: Facebook's shadow profiles are 'frightening' dossiers on everyone, 2013. URL https://www.zdnet.com/article/firm-facebooks-shadow-profiles-are-frightening-dossiers-on-everyone/. Last visited: Jan. 2025.

[11] L. Boeschoten, J. Ausloos, J. Moeller, T. Araujo, and D. L. Oberski. Digital trace data collection through data donation, 2020. URL http://arxiv.org/abs/2011.09851. arXiv:2011.09851 [cs].

[12] 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 *Proc. of the Annual Privacy Forum (APF)*. 2019. doi: 10.1007/978-3-030-21752-5_12.

[13] 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 *Proc. of the USENIX Security Symposium (USENIX Security)*. 2024.

[14] V. Braun and V. Clarke. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 2019. doi: 10.1080/2159676X.2019.1628806.

[15] V. Braun and V. Clarke. *Thematic Analysis: A Practical Guide*. 2021. URL https://us.sagepub.com/en-us/nam/thematic-analysis/book248481.

[16] C. Bösch, B. Erb, F. Kargl, H. Kopp, and S. Pfattheicher. Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies*, 2016. doi: 10.1515/popets-2016-0038.

[17] C. Cadwalladr and E. Graham-Harrison. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach, 2018. URL https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election. Last visited: Jan. 2025.

[18] J. P. Carrascal, C. Riederer, V. Erramilli, M. Cherubini, and R. de Oliveira. Your browsing behavior for a big mac: economics of personal information online. In *Proc. of the ACM Int'l Conf. on the World Wide Web (WWW)*. 2013. doi: 10.1145/2488388.2488406.

[19] Y. Cha and W. Pak. Protecting contacts against privacy leaks in smartphones. *PLOS ONE*, 2018. doi: 10.1371/journal.pone.0191502. Publisher: Public Library of Science.

[20] Y. Cheng, L. Ying, S. Jiao, P. Su, and D. Feng. Bind your phone number with caution: automated user profiling through address book matching on smartphone. In *Proc. of th ACM Symp. on Information, computer and communications security (AsiaCCS)*. 2013. doi: 10.1145/2484313.2484356.

[21] M. Cherubini, K. Salehzadeh Niksirat, M.-O. Boldi, H. Keopraseuth, J. M. Such, and K. Huguenin. When Forcing Collaboration is the Most Sensible Choice: Desirability of Precautionary and Dissuasive Mechanisms to Manage Multiparty Privacy Conflicts. *Proc. ACM Hum.-Comput. Interact.*, 2021. doi: 10.1145/3449127. tex.ids=cherubini_when_2021-1.

[22] C. Daboo. CardDAV: vCard extensions to web distributed authoring and versioning (WebDAV), 2011. URL https://www.rfc-editor.org/info/rfc6352. Number: 6352.

[23] EDBP. Guidelines 07/2020 on the concepts of controller and processor in the GDPR, 2021. URL https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_en.

[24] L. Fan, S. Zhang, Y. Kong, X. Yi, Y. Wang, X. O. Xu, C. Yu, H. Li, and Y. Shi. Evaluating the Privacy Valuation of Personal Data on Smartphones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2024. doi: 10.1145/3678509.

[25] A. P. Felt, S. Egelman, and D. Wagner. I've got 99 problems, but vibration ain't one: a survey of smartphone users' concerns. In *Proc. of the ACM Workshop on Security and privacy in smartphones and mobile devices (SPSM)*, 2012. doi: 10.1145/2381934.2381943.

[26] C. Fischer. Re-thinking the allocation of roles under the GDPR in the context of cloud computing. *International Data Privacy Law*, 2024. doi: 10.1093/idpl/ipad023.

[27] M. Furini, S. Mirri, M. Montangero, and C. Prandi. Can IoT Wearable Devices Feed Frugal Innovation? In *Proc. of the Workshop on Experiences with the Design and Implementation of Frugal Smart Objects(FRUGALTHINGS)*, 2020. doi: 10.1145/3410670.3410861.

[28] P. Galopoulos, C. Iakovidou, V. Gkatziaki, S. Papadopoulos, and Y. Kompatsiaris. Towards a Privacy Respecting Image-based User Profiling Component. In *2021 International Conference on Content-Based Multimedia Indexing (CBMI)*, 2021. doi: 10.1109/CBMI50038.2021.9461886. ISSN: 1949-3991.

[29] C. Geeng and F. Roesner. Who's In Control? Interactions In Multi-User Smart Homes. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*, 2019.

[30] D. George, K. Reutimann, and A. Tamò-Larrieux. GDPR bypass by design? Transient processing of data under the GDPR. *International Data Privacy Law*, 2019. doi: 10.1093/idpl/ipz017.

[31] Google. Contacts Provider, 2025. URL https://developer.android.com/identity/providers/contacts-provider. Last visited: Jan. 2025.

[32] Google. People API, 2025. URL https://developers.google.com/people/. Last visited: Jan. 2025.

[33] T. Groß. Validity and Reliability of the Scale Internet Users' Information Privacy Concerns (IUIPC). *Proceedings on Privacy Enhancing Technologies*, 2021. doi: 10.2478/popets-2021-0026.

[34] A. A. Habu and T. Henderson. Data subject rights as a research methodology: A systematic literature review. *Journal of Responsible Technology*, 2023. doi: 10.1016/j.jrt.2023.100070.

[35] C. Hagen, C. Weinert, C. Sendner, A. Dmitrienko, and T. Schneider. Contact Discovery in Mobile Messengers: Low-cost Attacks, Quantitative Analyses, and Efficient Mitigations. *ACM Trans. Priv. Secur.*, 2022. doi: 10.1145/3546191.

[36] R. Hasan, R. Weil, R. Siegel, and K. Krombholz. A Psychometric Scale to Measure Individuals' Value of Other People's Privacy (VOPP). In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*, 2023. doi: 10.1145/3544548.3581496.

[37] F. Houssiau, P. Sapieżyński, L. Radaelli, E. Shmueli, and Y.-A. De Montjoye. Detrimental network effects in privacy: A graph-theoretic model for node-based intrusions. *Patterns*, 2023. doi: 10.1016/j.patter.2022.100662.

[38] M. Humbert, E. Ayday, J.-P. Hubaux, and A. Telenti. On non-cooperative genomic privacy. In *Proc. of the Int'l Conf. on Financial*

Cryptography and Data Security (FC). 2015. doi: 10.1007/978-3-662-47854-7_24.

[39] M. Humbert, E. Ayday, J.-P. Hubaux, and A. Telenti. Quantifying Interdependent Risks in Genomic Privacy. *ACM Transactions on Privacy and Security*, 2017. doi: 10.1145/3035538.

[40] M. Humbert, B. Trubert, and K. Huguenin. A Survey on Interdependent Privacy. *ACM Computing Surveys*, 2020. doi: 10.1145/3360498.

[41] M. Humbert, D. Dupertuis, M. Cherubini, and K. Huguenin. KGP Meter: Communicating Kin Genomic Privacy to the Masses. In *Proc. of the IEEE European Symp. on Security and Privacy (EuroS&P)*, 2022. doi: 10.1109/EuroSP53844.2022.00033.

[42] A. K. Jindal, V. Banahatti, and S. Lodha. People to People (P2P) Privacy in Mobile Devices. In *Proc. of the Int'l Conf. on COMmunication Systems & NETworkS (COMSNETS)*, 2022. doi: 10.1109/COMSNETS53615.2022.9668478. ISSN: 2155-2509.

[43] D. Kamarinou, C. Millard, and F. Turton. Protection of Personal Data in Clouds and Rights of Individuals. In C. Millard, *Cloud Computing Law*. Oxford University Press, 2021. doi: 10.1093/oso/9780198716662.003.0008.

[44] D. Kamarinou, C. Millard, and F. Turton. Responsibilities of Controllers and Processors of Personal Data in Clouds. In C. Millard, *Cloud Computing Law*. Oxford University Press, 2021. doi: 10.1093/oso/9780198716662.003.0009.

[45] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler. "My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security. In *Proc. of the Symp. on Usable Privacy and Security (SOUPS)*. 2015.

[46] A. Khatoon and P. Corcoran. Android permission system and user privacy — A review of concept and approaches. In *Proc. of the IEEE Int'l Conf. on Consumer Electronics (ICCE)*, 2017. doi: 10.1109/ICCE-Berlin.2017.8210616.

[47] Kozea. Radicale v3: Free and Open-Source CalDAV and CardDAV Server, 2025. URL https://radicale.org/v3.html. Last visited: Jan. 2025.

[48] D. Krahn, D. McCloskey, and K. Yeo. Offline, Privacy-Preserving, Contact-Based Nearby Device Discovery. *Defensive Publications Series*, 2024.

[49] F. Kreuter, G.-C. Haas, F. Keusch, S. Bähr, and M. Trappmann. Collecting Survey and Smartphone Sensor Data With an App: Opportunities and Challenges Around Privacy and Informed Consent. *Social Science Computer Review*, 2020. doi: 10.1177/0894439318816389.

[50] K. Krombholz, K. Busse, K. Pfeffer, M. Smith, and E. von Zezschwitz. "If HTTPS Were Secure, I Wouldn't Need 2FA" - End User and Administrator Mental Models of HTTPS. In *Proc. of the IEEE Symp. on Security and Privacy (S&P)*. 2019. doi: 10.1109/SP.2019.00060.

[51] R. S. Laufer and M. Wolfe. Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory. *Journal of Social Issues*, 1977. doi: 10.1111/j.1540-4560.1977.tb01880.x.

[52] S. Liu and G. Biczók. IDPFilter: Mitigating interdependent privacy issues in third-party apps. *Computers & Security*, 2025. doi: 10.1016/j.cose.2025.104321.

[53] S. Liu and G. Biczók. Modeling interdependent privacy threats, 2025. URL http://arxiv.org/abs/2505.18386. arXiv:2505.18386 [cs].

[54] N. K. Malhotra, S. S. Kim, and J. Agarwal. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information System Research*, 2004. doi: 10.1287/isre.1040.0032.

[55] K. I. Manktelow and M. C. Chung. *Psychology of reasoning: theoretical and historical perspectives*. 2004. URL https://www.taylorfrancis.com/books/9780203506936.

[56] K. Marky, N. Gerber, M. G. Pelzer, M. Khamis, and M. Mühlhäuser. "You offer privacy like you offer tea": Investigating Mechanisms for Improving Guest Privacy in IoT-Equipped Households. *Proceedings on Privacy Enhancing Technologies*, 2022.

[57] M. Marsch, J. Grossklags, and S. Patil. Won't You Think of Others?: Interdependent Privacy in Smartphone App Permissions. *Proc. ACM Hum.-Comput. Interact.*, 2021. doi: 10.1145/3479581.

[58] T. Matsuura, A. A. Hasegawa, M. Akiyama, and T. Mori. Careless Participants Are Essential for Our Phishing Study: Understanding the Impact of Screening Methods. In *Proceedings of the 2021 European Symposium on Usable Security*, EuroUSEC '21, 2021. doi: 10.1145/3481357.3481515.

[59] Microsoft. Outlook personal contacts API overview, 2025. URL https://learn.microsoft.com/en-us/graph/outlook-contacts-concept-overview. Last visited: Jan. 2025.

[60] A. Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel. You are who you know: inferring user profiles in online social networks. In *Proc. of the ACM Int'l Conf. on Web Search and Data Mining (WSDM)*. 2010. doi: 10.1145/1718487.1718519.

[61] Mozilla. WebOTP API, 2025. URL https://developer.mozilla.org/en-US/docs/Web/API/WebOTP_API.

[62] H. Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. 2009.

[63] M. Oates, Y. Ahmadullah, A. Marsh, C. Swoopes, S. Zhang, R. Balebako, and L. F. Cranor. Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. *Proceedings on Privacy Enhancing Technologies*, 2018. doi: 10.1515/popets-2018-0029.

[64] J. O'Hagan, P. Saeghe, J. Gugenheimer, D. Medeiros, K. Marky, M. Khamis, and M. McGill. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2022. doi: 10.1145/3569501.

[65] A.-M. Olteanu, K. Huguenin, R. Shokri, M. Humbert, and J.-P. Hubaux. Quantifying Interdependent Privacy Risks with Location Data. *IEEE Transactions on Mobile Computing*, 2017. doi: 10.1109/TMC.2016.2561281.

[66] A.-M. Olteanu, M. Humbert, K. Huguenin, and J.-P. Hubaux. The (Co-)Location Sharing Game. *Proceedings on Privacy Enhancing Technologies*, 2019. doi: 10.2478/popets-2019-0017.

[67] S. Palan and C. Schitter. Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 2018. doi: 10.1016/j.jbef.2017.12.004.

[68] A. J. Perez, S. Zeadally, and S. Griffith. Bystanders' Privacy. *IT Professional*, 2017. doi: 10.1109/MITP.2017.42.

[69] S. Perreault. vCard format specification, 2011. URL https://www.rfc-editor.org/info/rfc6350. Number: 6350.

[70] Proton AG. Proton: Privacy by Design, 2025. URL https://proton.me. Last visited: Jan. 2025.

[71] Proton Team. Introducing Proton Mail Contacts — the world's first encrypted contacts manager, 2022. URL https://proton.me/blog/encrypted-contacts-manager. Last visited: Jan. 2025.

[72] N. Purtova. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, 2018. doi: 10.1080/17579961.2018.1452176.

[73] A. Quodling. Shadow profiles - Facebook knows about you, even if you're not on Facebook, 2018. URL http://theconversation.com/shadow-profiles-facebook-knows-about-you-even-if-youre-not-on-facebook-94804. Last visited: Jan. 2025.

[74] W. Richter. The scary reality of identity theft. *XRDS*, 2013. doi: 10.1145/2542650.

[75] D. Rosenthal. Das neue Datenschutzgesetz. *Jusletter*, 2020.

[76] K. Salehzadeh Niksirat, E. Anthoine-Milhomme, S. Randin, K. Huguenin, and M. Cherubini. "I thought you were okay": Participatory Design with Young Adults to Fight Multiparty Privacy Conflicts in Online Social Networks. In *Proc. of the ACM Designing Interactive Systems Conf. (DIS)*, DIS '21, 2021. doi: 10.1145/3461778.3462040. tex.ids= salehzadeh_niksirat_i_2021-1.

[77] K. Salehzadeh Niksirat, L. Goswami, P. S. B. Rao, J. Tyler, A. Silacci, S. Aliyu, A. Aebli, C. Wacharamanotham, and M. Cherubini. Changes in Research Ethics, Openness, and Transparency in Empirical Studies between CHI 2017 and CHI 2022. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, 2023. doi: 10.1145/3544548.3580848.

[78] K. Salehzadeh Niksirat, D. Korka, H. Harkous, K. Huguenin, and M. Cherubini. On the Potential of Mediation Chatbots for Mitigating Multiparty Privacy Conflicts - A Wizard-of-Oz Study. *Proc.*

ACM Hum.-Comput. Interact., 2023. doi: 10.1145/3579618. tex.ids= salehzadeh_niksirat_potential_2023-1.

[79] E. Saqib, S. He, J. Choy, R. Abu-Salma, J. Such, J. Bernd, and M. Javed. Bystander Privacy in Smart Homes: A Systematic Review of Concerns and Solutions. *ACM Trans. Comput.-Hum. Interact.*, 2025. doi: 10.1145/3731755. Just Accepted.

[80] E. G. Site. EF English Proficiency Index: A Ranking of 113 Countries and Regions by English Skills, 2023. URL https://www.ef.com/wwen/epi/.

[81] Solid. Solid: Your data, your choice., 2025. URL https://solidproject.org. Last visited: Jan. 2025.

[82] K. Spiel, O. L. Haimson, and D. Lottridge. How to do better with gender on surveys: a guide for HCI researchers. *Interactions*, 2019. doi: 10.1145/3338283.

[83] Synacor. Zimbra: Email & Collaboration best productivity tools, 2025. URL https://www.zimbra.com. Last visited: Jan. 2025.

[84] M. Tahaei, R. Abu-Salma, and A. Rashid. Stuck in the Permissions With You: Developer & End-User Perspectives on App Permissions & Their Privacy Ramifications. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*. 2023. doi: 10.1145/3544548.3581060.

[85] B. Van Alsenoy. *Data Protection Law in the EU: Roles, Responsibilities and Liability:*. 2019. doi: 10.1017/9781780688459.

[86] P. J. van de Waerdt. Information asymmetries: recognizing the limits of the GDPR on the data-driven market. *Computer Law & Security Review*, 2020. doi: 10.1016/j.clsr.2020.105436.

[87] L. Velykoivanenko, K. Salehzadeh Niksirat, S. Teofanovic, B. Chapuis, M. L. Mazurek, and K. Huguenin. Designing a Data-Driven Survey System: Leveraging Participants' Online Data to Personalize Surveys. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*, 2024. doi: 10.1145/3613904.3642572.

[88] S. Veys, D. Serrano, M. Stamos, M. Herman, N. Reitinger, M. L. Mazurek, and B. Ur. Pursuing Usable and Useful Data Downloads Under GDPR/CCPA Access Rights via Co-Design. In *Proc. of the Symp. on Usable Privacy and Security (SOUPS)*. 2021.

[89] S. Yan, T. Zhao, and J. Deng. Interaction-aware Hypergraph Neural Networks for User Profiling. In *Proc. of the IEEE Int'l Conf. on Data Science and Advanced Analytics (DSAA)*, 2022. doi: 10.1109/DSAA54385.2022.10032374.

[90] Y. Zou, K. Roundy, A. Tamersoy, S. Shintre, J. Roturier, and F. Schaub. Examining the Adoption and Abandonment of Security, Privacy, and Identity Theft Protection Practices. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*. 2020. doi: 10.1145/3313831.3376570.

[91] N. Zufferey, K. Salehzadeh Niksirat, M. Humbert, and K. Huguenin. "Revoked just now!" Users' Behaviors Toward Fitness-Data Sharing with Third-Party Applications. *Proceedings on Privacy Enhancing Technologies*, 2023. doi: 10.56553/popets-2023-0004.

A Google Contact Case Study: Resources

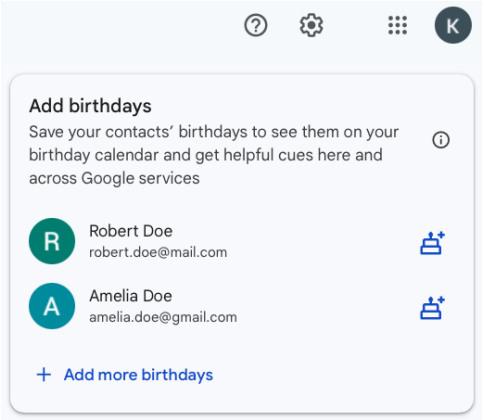
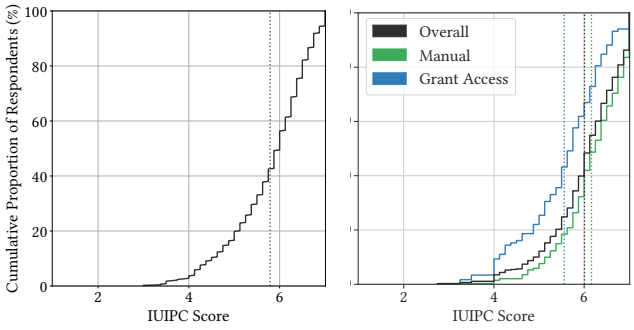


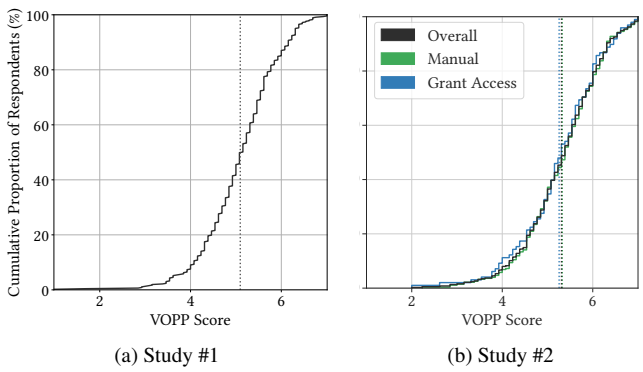
Figure 7: Screenshot of Google's suggestion to provide contacts' birthdays (source: https://contacts.google.com, visited: May 2025), edited/redacted for anonymization purposes. Note that the list of suggested contacts includes contacts that have not been saved by the DAB user but automatically saved by Google because the user interacted with them (e.g., sent them an e-mail).

B Additional Results from User Studies



(a) Study #1 (b) Study #2

Figure 8: Distribution of the respondents' IUIPC scores [54].



(a) Study #1 (b) Study #2

Figure 9: Distribution of the respondents' VOPP scores [36].

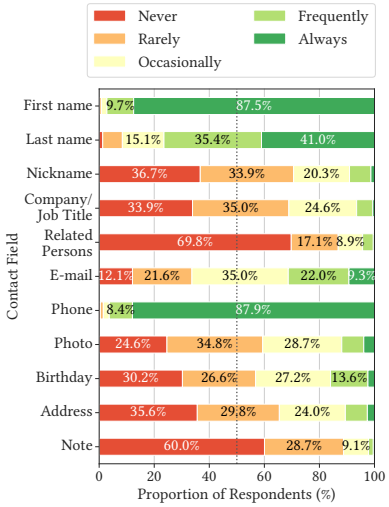


Figure 10: Distribution of the proportions of contact cards with a value specified for a given field (first name, last name, *etc.*), self-reported by the respondents (source: S1.Q7).

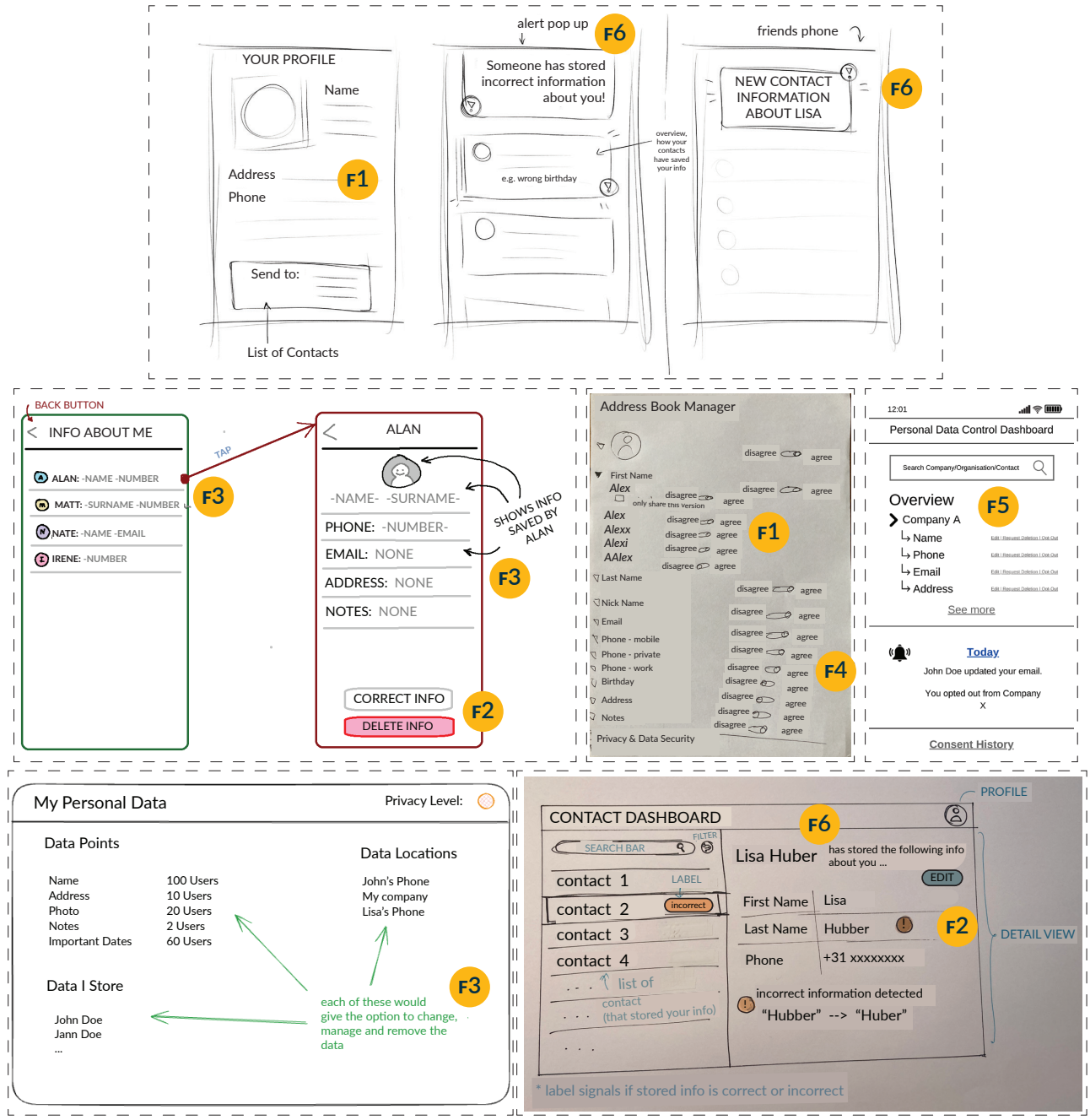


Figure 11: Example visualizations from S1.Q21 illustrating imagined features (F) of a dashboard for subjects to access and control their personal information stored in DABs. These features include: (F1) establishing accurate information, (F2) editing/deleting information, (F3) managing access, (F4) visualizing or quantifying data, (F5) distinguishing legal entities from natural users, and (F6) receiving notifications. To protect respondent anonymity, handwritten elements in sketches were replaced with text rendered in machine fonts.