"Erasing the Echo": The Usability of Data Deletion in Smart Personal Assistants

Cheng Cheng
University of Bristol
cheng.cheng@bristol.ac.uk

Kopo M. Ramokapane
University of Bristol
marvin.ramokapane@bristol.ac.uk

Abstract

Smart home personal assistants (SPAs) have gained widespread popularity in recent years. Despite their widespread adoption, existing research indicates that many users remain unaware of how to delete from these devices. While data deletion has been explored in other contexts, in the context of SPAs it is not well understood. This paper addresses this gap by investigating users' understanding of data deletion in the context of SPAs and evaluating the usability of existing deletion mechanisms. To address this, we conducted an interview study with 20 Amazon Alexa and Google Home users, during which we also observed their interactions with deletion processes. Our findings reveal that users hold diverse mental models regarding data deletion, recovery, and data storage, often leading to uncertainty and a lack of confidence in the deletion process. Moreover, we identified several usability challenges, particularly when users attempted to delete data using voice commands. Based on these insights, we discuss the implications for the design of more effective and transparent data deletion mechanisms in SPAs.

Keywords

Data Deletion, Smart Personal Assistants, Smart Speakers, Usability of Data Deletion, Privacy

1 Introduction

With the rise of Smart Home Personal Assistants (SPAs) like Amazon Alexa and Google Assistant, users have embraced the convenience of voice-activated technology in their everyday lives. Using voice commands, these assistants enable users to control smart home features, search the internet, and play music. However, as the functionality of these devices has grown, so have concerns regarding the privacy and security of the data they collect. SPAs constantly listen for voice commands, storing large amounts of personal data, including voice recordings, search histories, and interaction logs. This has led to increased scrutiny over how user data is handled and particularly how easily users can manage or delete their data.

While privacy concerns related to SPAs are well-documented, the usability of data deletion features remains largely unexplored. Existing research from other domains suggests that users actively delete data to protect their privacy [9, 23, 31, 37–39]. However, regarding SPAs, studies indicate that many users are either unaware of or misunderstand how data deletion works [1, 9, 25]. For instance, users often mistakenly assume their voice recordings are deleted automatically, while others believe deletion removes their data completely. However, when deletion options are available, users often engage with these mechanisms in various ways, sometimes encountering challenges [9]. Yet, prior studies have not examined the specific usability issues users face when interacting with deletion features. Though most SPAs offer data deletion options, the process is often opaque, cumbersome, or inconsistent across different devices and platforms. Users may struggle to protect their privacy through deletion due to a lack of understanding of how SPAs function or because they are unaware of the full scope of data retained by these devices [1, 44]. This raises important questions about the usability of data deletion in SPAs and the extent to which they empower users to manage their personal information effectively. Prior research on SPAs has largely overlooked the usability of deletion mechanisms and users' understanding of data deletion in the context of personal assistants. Malkin et al. [25] examined participants' awareness of data retention and the existence of deletion options but did not investigate their mental models or the usability of these mechanisms. Similarly, Abdi et al. [1] explored users' mental models of SPAs through scenario-based interviews, yet their study did not consider scenarios with data deletion.

To address this gap, this paper investigates the usability of data deletion mechanisms in SPAs. We particularly focus on three aspects: users' perceptions of data deletion in the context of SPAs, the usability of deletion mechanisms, users' perceptions of the mechanisms, and the challenges they encounter. We also seek to understand their expectations and desired improvements for data deletion features. To explore these issues, we conducted semi-structured interviews and a walk-through experiment, where participants attempted to delete data and reset SPA devices to their default settings. Through this approach, we analyse users' mental models of data deletion, identify usability barriers, and propose improvements to make data deletion in SPAs more transparent, intuitive, and accessible. In particular, we aim to answer the following research questions:

RQ1: How do users perceive data deletion in the context of SPAs?
RQ2: How do users perceive the usability of data deletion mechanisms in SPAs?
RQ3: What do users want with regard to data deletion in SPAs?

Overall, our main findings are summarized as follows: Participants defined data deletion based on various factors, including the type of data being removed, purpose, consequence, recoverability, and access. While SPAs input audio data, participants view input data in terms of content, what content they share, such as reminders and playlists. We identified several complete and incomplete mental models regarding data deletion in SPAs, including deletion as permanent and allowing no recovery, expectation of feedback and

Proceedings on Privacy Enhancing Technologies 2025(4), 76–93

https://doi.org/10.56553/popets-2025-0120

confirmation after deletion, believing data is recoverable, and assuming deleted data is only available to providers and developers. The variety of models contributed to their concerns. For instance, some questioned whether deletion really protects their privacy if their deleted private conversations are not completely deleted. In terms of the usability tasks, we found that while the deletion of voice recordings also removes their textual scripts, this is not always clear to some users. Moreover, since voice is the primary mode of interacting with the assistants, participants preferred to use voice commands to reset the devices, despite the likelihood of accidental deletion. Participants desire the improvement of the voice deletion feature and app interfaces to make deletion options better and accessible. They also desire to have data recovery, assurance, and deletion options as part of the onboarding or setup process of SPAs. These findings highlight the critical need for better-designed transparent deletion mechanisms and improved usability in SPAs to ensure users have agency to manage their data. In summary, we provide the following contributions:

- **Insights into users' perceptions of data deletion in SPAs.** We identify nine distinct mental models that describe how participants understand data deletion in SPAs, including their assumptions, expectations, and concerns.
- **Identification of usability challenges in SPA deletion mechanisms.** Using a think-aloud method, we explore how users attempt to delete voice recordings, textual transcripts, and reset assistants. We highlight the obstacles they encounter and the strategies they adopt when their approaches fail.
- **Design recommendations to improve data deletion in SPAs.** Drawing from the findings of RQ1 and RQ2 and participants' expressed desires, we propose actionable design improvements to enhance the usability and transparency of data deletion features in SPAs.

2 Background and Related work

2.1 Background

2.1.1 Overview. Smart Personal Assistants (SPAs) are voice-enabled systems designed to help users accomplish tasks through natural language commands. These systems are designed to interpret natural language and respond in a conversational manner, enabling users to ask questions, set reminders, and control other smart home devices, often without requiring direct physical interaction [4]. SPAs are embedded in various devices like smartphones and smart speakers and rely on cloud-based processing and context-aware intelligence for intent recognition and action execution [12]. For the purposes of this study, we provide background on smart-speaker-based assistants, Google Home, and Amazon Alexa.

2.1.2 SPA Architecture: A Multi-Layered Ecosystem. While SPAs may appear simple to users, they are built on complex, distributed architectures [3, 12]. These ecosystems typically consist of multiple devices, cloud-based processing capabilities, smart home infrastructures, and networked communications, which together enable various functions. A typical SPA architecture consists of the following components:

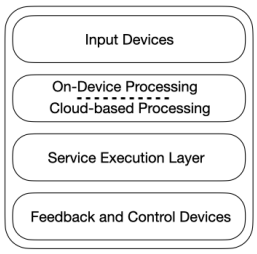


Figure 1: Typical SPA Multi-Layered Architecture.

- Input Devices: These include microphones embedded in smart speakers, mobile phones, or wearable devices that continuously listen for activation keywords (e.g., "Alexa").
- In-Device/Cloud-Based Processing: Once a command is detected, it is transmitted to the cloud, where speech recognition, natural language understanding, and intent classification are performed.
- Service Execution Layer: Based on the parsed intent, a request may be routed to internal services (e.g., calendar) or external third-party services (e.g., smart home APIs).
- Feedback and Control Devices: These are endpoints like smart light bulbs or thermostats that receive and execute commands.

These components work together to deliver a seamless user experience. For instance, a command to switch on smart lighting involves capturing the audio, cloud-based parsing, intent resolution, invocation of a third-party smart home service, and device-level execution [3]. This process highlights the complexity of data flow required to complete a task [1, 44]. The involvement of multiple devices, cloud-based processing, and networked communication adds layers of complexity to data management, particularly data deletion. Deleting data in these systems is not straightforward; when a user requests deletion, the request must propagate through various systems, logical and physical layers, and, in some cases, third-party services, making data deletion challenging to assure. This complexity may make it difficult for users to fully understand and control data deletion, thereby compromising their ability to safeguard their privacy.

2.1.3 Data Types and Deletion Mechanisms. While most users interact with SPAs via voice commands, these commands are converted into text for processing, something users may not be aware of when using the device or attempting to delete data. Moreover, output data (e.g., interaction history) generated by the assistant may not always be apparent to users. By default, Alexa stores voice recordings indefinitely, while Google Home does not. However, Google may still store textual transcripts of users' interactions by default as part of users' broader "Web & App Activity."

SPA providers usually offer users multiple interfaces to delete their data. Users can delete data through smartphone apps, voice commands, web interfaces, customer service, or directly interacting physically with the speaker. SPAs providers often offer an accompanying smartphone app that allows users to manage and delete data. The voice interface enables users to command the smart assistant to delete data, while web interfaces, for example, account management

Table 1: SPA data types, Control interfaces, and Deletion methods

SPA Platform	Data Type	Data Origin	Control Interface	Deletion Method
Amazon Alexa	Voice recordings	User generated	Alexa App, Voice, Website	Manual, Voice, Auto-delete
	Text transcripts	System generated	Alexa App, Website	Limited; Tied to the deletion of voice recordings
Google Home	Voice recordings	User generated	Google Assistant App, Voice, Website	Manual, Voice, Auto-delete via activity settings
	Text transcripts	System generated	Google Assistant App, Website	Can be viewed and deleted separately from recordings

portals, allow users to interact with and remove stored data. For Alexa, voice deletion is not enabled by default; users need to enable it under "Settings". However, voice deletion is enabled by default in Google Home assistants. Additionally, Google and Amazon also offer "auto deletion" capabilities; users can set their assistant to automatically delete voice recordings after a certain period of time. Some providers also allow users to delete their data through their customer service options. Beyond user interfaces, some SPAs offer users a full set of settings for viewing and deleting voice recordings, as well as setting up automatic deletion of recordings and pausing the storage of recordings [46]. Overall, the SPA ecosystem contains various features and complex processes that may affect how users perceive data deletion or delete data from their devices. Table 1 summarizes types of data collected and generated by Google and Amazon and interfaces available.

2.2 Related Work

2.2.1 Users' security and privacy issues. Due to the popularity of SPAs, several studies (e.g., [10, 22]) have investigated their security and privacy implications. Key concerns include the 'always listening' nature of the speakers and the collection, storage, and potential misuse of recorded data [2, 12, 19, 45, 50]. While users acknowledge that data collection can improve personalization or device functionality, they worry about how their data is handled and shared [2, 49]. For instance, Cowan et al. [11] found that the storage of voice recordings is a significant factor driving privacy concerns among users. Nevertheless, there are studies showing that most users are unaware that these devices store and share recordings. For instance, Malkin et al. [25] surveyed 116 users and discovered that nearly half were unaware that their recordings were stored and available for review. Only a quarter had ever reviewed their interaction records, and very few had deleted them. Moreover, they are also unaware of the types of data being collected by speakers [45]. In this paper, we explore the usability of deletion mechanisms to understand why many users do not delete stored voice recordings. We also examine how users perceive and interpret the data collection and processing practices of SPAs.

Prior research [1, 19, 22, 36, 45] suggests that many users do not fully understand how SPAs function, particularly regarding data collection and device functionality. This limited awareness contributes to low risk perceptions and their ability to protect themselves. For instance, Zheng et al. [50] found that users often prioritize convenience and connectivity over privacy. Abdi et al. [1] also reported that users tend to focus on household use and device manufacturers, often overlooking third-party integrations, which results in incomplete threat models and gaps in their protective strategies. Our work builds on these works with the aim of understanding how users perceive deletion in such complex ecosystems. While previous studies have focused on the broader aspects of privacy and user preferences, they fall short in exploring the technical and procedural nuances of how data is actually deleted from SPAs. We argue that addressing this gap is critical for improving deletion usability and user trust. Through our study, we provide insights that can inform the design of more transparent and effective deletion mechanisms.

2.2.2 Users' data deletion practices and preferences. Data deletion in SPAs remains underexplored. However, prior studies [9, 25] indicate that users are often unaware that their recordings are stored and available for review or deletion. When users become aware of such recordings, they often express discomfort and a strong desire to delete them, but often lack knowledge on how to do so effectively [23, 25]. Cho et al. [9] indicated that providing users with an option to delete their voice history enhances trust in smart speakers. However, the effectiveness of these controls is unknown, as users' data is often distributed across the SPA ecosystem [12]. Our study addresses this gap by examining the usability of data deletion mechanisms in SPAs. We find that many users struggle to locate and use these controls effectively.

In a non-SPA context, several studies [18, 31, 34, 38, 43] have examined users' deletion behaviors across different digital platforms. These studies suggest that users delete data for various reasons, including protecting their privacy, correcting data (e.g., deleting incorrect or outdated information), and reducing clutter. In the context of social media, studies [28, 48, 51] found that users often delete posts they regretted sharing. Failure to delete such content can lead to unintended disclosures, clutter, regret, and even emotional distress. Incomplete mental models of deletion and how technology functions have been identified as a major reason why users fail to delete data. For instance, Ramokapane et al. [38] found that users hold various misconceptions about cloud storage deletion, leading them to ineffectively remove data from their accounts. Similarly, Murillo et al. [31] observed that users' understanding of deletion largely depended on the interface. Many users assumed that data was deleted simply because it was no longer visible to them. Ion et al. [20] highlighted that users often lack awareness of critical deletion concepts, such as timeliness and data retention policies. In instant messaging apps, Schnitzler et al. [43] found that ambiguous terminology in deletion options often caused users to misunderstand the consequences of their actions. Our study contributes insights around how users perceive data deletion in SPAs and how they interact with the deletion mechanisms provided.

Other studies [17, 18, 24] suggest that users fail to delete because they are unaware of deletion controls, or the interfaces and deletion options are not usable. For instance, Habib et al. [18] analyzed data deletion mechanisms across 150 websites and found that while 74% offered deletion controls, only 18% provided a direct link to a

deletion tool or request form. Regarding user expectations, users prefer flexible deletion options rather than a one-size-fits-all approach [37, 39]. Moreover, they also want clearer explanations and more transparent deletion processes that align with their expectations [17, 18, 24, 37]. Through our RQ3, we provide insights on what users want with regard to data deletion in SPAs.

3 Methods

To understand the usability of data deletion mechanisms in SPAs, we interviewed 20 users of Amazon Echo and Google Home assistants.

3.1 Ethical Considerations

Our study was approved by our IRB and followed trauma-informed practices [7]. The study material (i.e., the Participation Information Sheet and Consent Form) was sent to participants before the sessions to ensure they understood the goal of this study and their participation. Participants gave consent to voluntarily participate and be audio recorded. They were informed they could withdraw from the study at any time before data anonymization and analysis without giving any reason. Audio recordings were deleted immediately after transcription.

3.2 Study Design

Our study consisted of two parts: the first part sought to understand how users perceived data deletion from their SPAs, while the second part was practical, where participants were asked to show how they normally delete data from the assistants. We did not set out to compare deletion between the Amazon Echo and Google Home, nor compare users' performance when deleting from both assistants. Our focus was on the overall deletion experiences of users.

3.2.1 Perception and understanding of deletion. To explore how users perceive and understand data deletion, we developed an interview protocol as follows: The first aspect focused on building rapport and making participants comfortable. These questions explored how frequently they used devices and their primary use cases. We then examined users' knowledge of data processing and storage in SPAs. The third set of questions explored users' perceptions and expectations around data deletion. Participants were asked to describe data deletion, including where the deleted data goes and whether they believe data can be recovered. This helped us identify their expectations, mental models, and gaps between their perceptions and reality. The final set of questions examined users' data deletion practices, the challenges they face, and how they address those challenges.

3.2.2 Deletion tasks. In the second part of this study, we made two SPAs available: Google Nest Mini (2nd Gen) and Amazon Echo Plus (1st Gen). We then provided each participant with an SPA device which they were most familiar with. Each SPA was paired with a smartphone through the Google Home app (App Version 3.22.101) and the Amazon Alexa app (App Version 2024.17) respectively. Participants were then asked to complete three data deletion tasks:

- **Task 1:** Deleting voice recordings,
- **Task 2:** Deleting transcripts of voice recordings,
- **Task 3:** Resetting the SPA back to factory settings.

These tasks allowed us to observe how users interact with the various deletion mechanisms available for SPAs. We started with commonly executed tasks: Task 1, deleting voice recordings, and Task 2, deleting textual transcripts. Task 3 focused on resetting the assistant to its default factory settings, a mechanism that many users may not frequently use. While deleting voice recordings (i.e., *Task 1*) also removes the corresponding transcripts (i.e., *Task 2*), we designed these as separate tasks for two reasons. First, we aimed to assess whether participants understood that SPAs not only process voice commands but also generate and store textual transcripts. Second, we wanted to evaluate whether participants knew where to locate these transcripts and how to manage them independently within the settings. Distinguishing the tasks helped us better understand participants' mental models of what data is stored and how deletion mechanisms are perceived and used.

3.3 Pilot Study

We conducted two pilot interviews to test our protocol. After the first interview, we revised some of our questions to remove duplicates and add more prompts to encourage more detailed responses. Following the second pilot, we refined the tasks and added follow-up questions to ensure that we had a better understanding of users' practices. Since the changes were minimal, we therefore did not conduct another pilot interview. Data from these pilot studies were not included in the final results.

3.4 Recruitment and Screening

We targeted participants who owned or used an Amazon Alexa or Google Home smart assistant for at least a month. We recruited from our city through posters, word-of-mouth, and snowball sampling [32]. We also encouraged participants to share our study with others who might be interested. Interested individuals were asked to complete a screening survey via a QR code or short link, which included questions about device ownership, usage experience, and demographics. We did not formally pre-test the screening survey but internally reviewed it prior to deployment. On average, our screening survey took 2.5 minutes to complete and can be found in Appendix C. We received 34 responses in total, of which 24 met our eligibility criteria and were invited to our study. Finally, 20 participants showed up and completed our study. There were ten (10) participants for each assistant. Eleven (11) participants identified as primary users, having set up and managed the smart speaker themselves, while the remaining nine (9) were secondary users, meaning they used the speaker regularly but did not configure or manage it directly. Table 2 summarizes the demographic information (see Appendix A).

3.5 Study Procedure

All the sessions were conducted in person on the university premises. The lead researcher conducted all the interviews; twelve (12) were conducted in English and eight (8) in Mandarin Chinese. At the start of each session, she explained the study objectives and the withdrawal process, and obtained consent for participation and audio recording. Following the interview guide (see Appendix D), the researcher asked the questions in the first three parts and then

instructed the participant to complete the three tasks under observation. All participants completed usability tasks on lab-provided devices. These included an iPhone 11, Lenovo Legion Y700, Google Nest Mini, and Amazon Echo Plus. Prior to beginning the tasks, participants were asked which smartphone operating system and smart assistant they were most familiar with. Based on their responses, they were given the device they were most comfortable using to simulate a realistic usage scenario. During the tasks, participants were encouraged to describe their thought process as they were completing the tasks (i.e., think-aloud process [47]). This enabled us to observe and record how participants reasoned through each step. The session concluded with a discussion of what information about data deletion participants believed should be communicated to SPA users. On average, interviews lasted for 44 minutes. At the end, participants received a £10.00 voucher as compensation for their time. One participant was unable to complete the tasks due to an emergency. We only analyzed their data for the first part of the study.

3.6 Thematic Analysis

All interviews were transcribed and, where applicable, translated by professional human transcribers and translators to ensure accuracy. After transcribing our recordings, we analyzed the scripts using the thematic analysis approach [5]. Two researchers independently coded two transcripts to develop a codebook, then met to discuss the identified codes and resolve any discrepancies to generate a unified codebook [27]. To ensure the validity of this codebook, the two researchers independently coded a third transcript, then discussed and refined their codebook. The primary researcher then proceeded to code the remaining transcripts, while the second researcher randomly selected and coded additional transcripts to maintain validity. We reached data saturation after 13 participants. However, because we initially had fewer than 10 participants per smart assistant group and the distribution was unbalanced, we continued recruitment until we had 10 participants for each assistant, resulting in a final sample of 20 participants. Since all discrepancies were resolved through discussions by two researchers, checking inter-coder reliability was deemed unnecessary [6, 8, 13, 15]. This is a qualitative study with a small sample size; therefore, we refrained from reporting the exact number of participants associated with a given theme to avoid overemphasizing the prevalence of a theme [16, 41]. Moreover, Braun and Clarke [5] argue that reporting counts can distort the nuanced, interpretive nature of qualitative analysis.

3.7 Limitations

Our study is not without limitations. First, the majority of our sample were students (13/20) and aged between 21 and 30 (14/20). This may have resulted in findings that reflect the views and experiences of a younger and more educated population. Prior studies (e.g., Kim and Choudhury [21]) also show that SPAs are also used by older adults as assistive devices. Future studies should aim to recruit more diverse populations. Second, part of our study relied on self-reported views. For example, some participants may have inaccurately reported whether they had previously deleted data or configured their assistants, which could have influenced their performance during usability tasks. This introduces uncertainty regarding the accuracy of some qualitative insights. Third, we did not explore participants' mental models or behaviors related to data deletion involving third-party apps or smart device integrations (e.g., deleting data associated with smart home devices connected via SPAs). This limits our understanding of deletion practices in more complex or integrated ecosystems. Fourth, participants used lab-provided smartphone devices. While some participants were accustomed to similar devices in their daily lives, others used newer or older models at home. These differences may have affected how intuitively participants navigated the deletion tasks. Lastly, as is common in qualitative research, our findings are not statistically generalizable. Some patterns observed in our study may not hold across broader populations. Future studies should validate these results with larger and more generalizable populations.

4 Findings

We present our findings in three parts: first, users' perceptions of data deletion in SPAs (RQ1), then usability tasks (RQ2), and finally, users' desired improvements (RQ3). We include a few illustrative counts to orient the reader. However, these figures are provided only for context and should not be interpreted as statistically representative of a broader population [26, 40, 42].

4.1 RQ1: How do users perceive data deletion in the context of SPAs?

4.1.1 Users' Mental Models of Data Deletion in the Context of SPAs. Our analysis revealed that participants' understanding of data deletion in the context of SPAs varied considerably. Participants characterized deletion in terms of the type of data being removed, its intended purpose, associated consequences, recoverability, and accessibility. We organize these perspectives according to the distinct mental models that emerged from our analysis.

Mental model 1: *"Input as Shared Intent, Not System Representation."* While all participants recognized that the assistant collected and processed audio, they consistently described input data in terms of the content of their commands and what they shared with the assistant, such as reminders, reading history, or playlists. This reflects a mental model in which "input" refers to the user's intended message, rather than the underlying data representations (e.g., audio files or transcripts). We also found that most participants (*n=11*) were unaware that their spoken inputs are transcribed and stored as text. For instance, P9 was alarmed: *"It has text transcripts too?"* This indicates a *partially complete mental model*, in which users correctly understand the high-level interaction but lack awareness of how the system processes, stores, and represents that data internally.

Mental model 2: *"Data Retention as Necessary for Functionality and Improvement."* Several participants (*n=4*) believed that retaining both voice recordings and textual transcripts was necessary to ensure the accuracy, personalization, and improvement of the assistant ecosystem. This model reflects a belief that storing multiple representations of input data supports technical needs such as speech recognition, debugging, and inclusivity. For instance, P20 noted that audio alone might not work well, hence the need for additional texts: *"The existence of text is necessary because the conversion from audio to text can sometimes differ."* P8 viewed voice data as essential for informing system improvements: *"[Voice recordings] help*

designers optimize and understand user needs, which then informs system improvements." Others, like P12, highlighted accessibility: *"Having both [voice and text] might cater to different needs... helping individuals with disabilities who can listen to voice recordings instead of reading text. They also accommodate those who prefer listening over reading, providing convenience and accessibility to various user groups."* While this mental model demonstrates a *more developed understanding* of how data supports assistant functionality, it often lacks a critical perspective on the long-term consequences of data retention. Participants tended to focus on the usefulness of data for improving services but gave little attention to how long this data is stored, who has access to it, or whether it can truly be deleted.

Mental model 3: *"Data Use Viewed Narrowly Through Immediate Functionality."* Some participants (*n=4*) agreed that both data types are useful. However, two believed that storing only voice recordings is sufficient for the assistant to work. Participants with this perspective often saw additional data formats, such as transcripts, as unnecessary or redundant. For instance, P13 said:*"I don't see a practical reason for keeping both formats. I believe the audio alone should be sufficient for understanding speech patterns."* This model reflects a *limited understanding* of the system's data processing pipeline, focusing narrowly on immediate speech recognition without accounting for how different data representations can be used. Moreover, this thinking also shows that some users do not know that textual scripts can be shared with third parties on their own without voice recordings (e.g., developers in the case of Amazon), which has implications for deletion.

Mental model 4: *"Voice and Textual Scripts as Equivalent."* Some participants (*n=5*) perceived no differences between the two data types; they explained that the scripts were simply transcribed recordings and served no other purpose than commanding the assistant to complete a task. This mental model treats transcripts as a byproduct of voice commands, assuming that both formats serve the same immediate functional purpose of issuing instructions to the assistant. P18 explained, *"I don't think there's much of a difference. They seem the same to me, I don't see a difference."* Similarly, P8 noted: *"they are linked together... Ideally, both should exist and be linked together"* This mental model is *incomplete*. Participants did not consider that voice and text may serve distinct backend roles. Because participants see voice and text as functionally equivalent and tightly coupled, they may assume that deleting one will automatically delete the other, or that both are removed together. In reality, SPAs often handle these formats separately, with different retention periods and technical affordances. This creates a mismatch between the perceived and actual scope of deletion.

Mental model 5: *"Voice recordings contain more information."* While some participants (*n=5*) did not perceive any difference between the two data types, others (*n=5*) highlighted that they were different; voice recordings contained more information about them and their surroundings, while textual scripts only represented the words of what they wanted the assistant to do. This mental model distinguishes between the two data types not just in format, but in informational richness and privacy sensitivity. P15 said *"Voice recording may have more personal information, while the text transcript is just words."* Other participants highlighted that the additional information in the voice recordings made them more privacy sensitive. For instance, P4 reported: *"Voice recordings contain more information, including my words, emotions, and voiceprint, making them more privacy sensitive. Text transcripts objectively reflect what I intend to do..."* Despite these differences, participants acknowledged that the average user might struggle to differentiate between the two data type or even understand their roles and implications. P19 explained:*"As an average user, it's hard to distinguish their specific roles or implications."* This is a *complete mental model* in terms of privacy sensitivity. Participants accurately identified the higher risk profile of voice data, which aligns with how such data is treated in SPAs. Voice recordings can be used for biometric authentication and profiling. Regarding deletion, users holding this model may expect or demand more robust deletion options for voice recordings or assume that deleting the transcript is insufficient to protect their privacy.

Mental model 6: *"Deletion Requires Confirmation and Transparency."* The majority of participants (*n=16*) stated that they expected additional information about their request, for instance, data retention or type of deletion. They also stressed that they had expected notifications about whether their request had been completed successfully. This model assumes that deletion is not complete unless the system explicitly communicates that the request has been processed. P3 explained,*"a notification, either through the mobile app or via email, confirming that all my data has been successfully deleted from the system."* This is a *partially complete* mental model. While this expectation aligns with good design principles, SPAs often lack clear, user-facing feedback on what data was deleted, how completely, and from where. Moreover, this model relies on visible confirmations as indicators for deletion and overlooks the complexities of how data is stored, replicated, or retained within various systems. As a result, participants may incorrectly assume that a notification assures full system-level deletion.

Mental model 7: *"Deletion as Permanent and Irrecoverable Removal."* Most participants (*n=12*) dismissed the concept of deleted data; they argued that when they request the assistant to delete data through voice or the app settings, they expect complete removal of data from the system, leaving no trace. They believed that data deletion from the assistants was complete. This mental model assumes a one-step, absolute process that once the data is deleted, it is irretrievable, with no residual copies. For instance, P18 explained that they had never deleted anything, and it came back:*"In my experience with this thing, I've never noticed anything that I've deleted coming back."* P7 argued that post deletion access invalidates deletion: *"I don't think so. If I can access it, that means they didn't delete it at all, or they have a backdoor to my deleted data."*

Others (*n=3*) associated the absence of a recovery mechanism like a recycle bin with complete deletion. As P6 explained: *"Unless there is a recycle bin, the data should not be recoverable. If a recycle bin or similar feature is in place, it would allow the user to restore deleted data if needed. But without such a system, once the data is deleted, it should be permanently removed with no option for recovery."* While the rationale behind this thinking varied, this model reflects a *simplified* and overly confident understanding of deletion. Participants with this understanding may overestimate their control over data and assume that issuing a delete command assures complete removal of data from all layers of the system. As a result, this mental model

can contribute to a false sense of privacy and misinform decisions around data sharing and deletion.

Mental model 8: *"Data recovery is possible in SPAs."* Some participants (*n=8*) viewed deletion as incomplete, holding the belief that data could still be recovered after a deletion request. They explained deletion is never complete; data is often hidden from users, companies need the data, and systems are designed to retain data to allow recovery. This model reflects an assumption that deletion primarily affects what is visible to the user, while underlying data remains stored in the system. P4 explained, *"data might only be removed at the user interface level, while the original file or its synced versions could still exist elsewhere ... so, even if the user chooses to delete, the original or synced versions of the data could still be lingering, making it [recoverable]."* While the majority had never tried to "test this theory," P12 shared an instance where data was recovered via customer service: *"I've personally tested this by contacting Amazon support regarding a 'deleted' music file, which they were able to recover for me."* This is a *partially complete* mental model. Participants correctly anticipated that deletion may not remove all copies of their data, this captures the complexities around deleting from cloud-based systems. However, it generalizes this behavior across different data types and lacks awareness around system constraints or data lifecycle policies in SPAs.

Mental model 9: *"Deleted data access is limited to providers."* Several participants (*n=9*) believed deleted data is accessible to service providers and their developers. This mental model assumes that deletion affects only user-facing visibility but that service providers can still retrieve or view the data. For instance, P8 noted, *"I think employees of the service provider, especially those managing the backend operations, can access it. I think they might have the rights to..."* P19 added, *"...relevant management personnel in the backend can access it."* This belief reflects a common assumption that technical expertise or privileged roles override deletion. This mental model is *partially complete*. While it reflects legitimate concerns about backend access and the limitations of user-deletion interfaces in SPAs, it does not clearly differentiate between data marked for deletion and data actually retained, both of which service-level administrators may have access to. Consequently, users may be uncertain about who controls deleted data and whether this access is due to service provider policy, technical capability, or both.

4.1.2 Concerns Around Data Deletion. Regarding concerns around data deletion in SPAs, participants were concerned about deletion not being permanent, uncertainty of deletion, access to deleted data, and having no full control over deleted data.

Deletion not being complete and permanent. Eight participants raised concerns that data may not be obsolete, emphasizing that when they request data to be deleted, they expect its complete removal from the system. Some further explained that they were worried about how deleted data could be used, while others were concerned that undeleted data may not be sufficiently anonymised, leaving them at risk. For instance, P15 stated, *"I definitely want it cleaned as cleanly as possible without any trouble or risks...like what I said before, this might be private and really important for me, so I don't want them to use those methods that might cause any issues."* Moreover, others highlighted a lack of transparency in data retention practices, expressing concerns that companies may retain their data without providing explicit notification or obtaining user consent. *"Some companies might secretly retain my data without informing me, which is not good and can be upsetting"* (P20).

Complete deletion is uncertain. Some participants (*n=6*) expressed uncertainty about whether data deletion in SPAs is complete or effective, as they were unsure whether data is completely removed from their systems. P14 noted, *"I'm not sure if this will remove all the data from the speaker. I think it will, but I'm not certain."* While some participants attributed this uncertainty to the absence of deletion confirmation mechanisms, others cited the lack of recovery features, such as recycle bins or trash folders, which they normally use to know whether deleted data is still available in the system. In addition, several participants (*n=9*) reported that their inability to verify deletion contributed to their uncertainty. They explained that SPAs do not provide mechanisms to confirm whether data has been completely deleted. Some stated while they could review their deletion actions within the app, they cannot be sure that data has been permanently deleted. As P9 explained, *"I can double-check and ensure the correct data is deleted. The current process doesn't give me confidence that I've truly deleted anything."* Due to the absence of verification mechanisms, some participants reported feeling compelled to trust that service providers would honour their deletion requests. *"I don't think I can ensure it. It's difficult for me to investigate or verify, whether online or offline ... When it comes to this issue, I can only place my trust in the ethical values and integrity of these companies"* (P1).

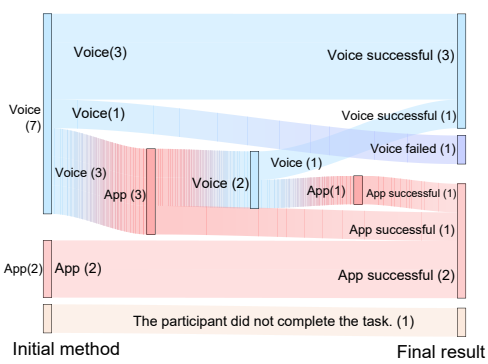
Access to deleted data. Other participants (*n=6*) were concerned about access to deleted data. Many participants reported discomfort that their deleted data can be accessed without their consent and knowledge. Some also highlighted that they do not have a list of who has access to such data. Participants said this was in violation of their privacy, especially that some delete to preserve privacy. P5 said *"I would be very offended if I found that deleted data can still be accessed by Google. Why should they have access to my private data, especially after it has been deleted."* Similarly, P9 noted, *"I would feel extremely uncomfortable, as if they were intruding into my private life. It would feel like an invasion of my personal space."*

Lack of control over deletion. Some participants (*n=4*) reported some concerns around lack of control over data deletion processes in SPAs. Many of them felt that deletion is out of their control as users but managed by service providers. Some said service providers had all the control, they could decide what is deleted completely, what can be recovered, retention without their input. P9 argued that the ability of service providers to allow recovery suggests that the deletion process is out of users' control. Other participants (*n=3*) explained that the SPA mechanisms are limited because when they delete, they lose control over their data. P2 highlighted, *"Because it's my data. I should have the right to control my own data. If they save [it] forever, I lose my control over my data."*

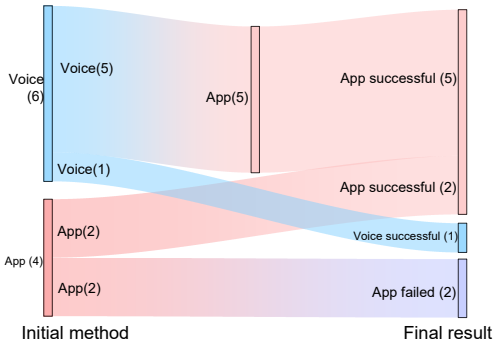
4.2 RQ2: How do users perceive the usability of data deletion mechanisms in SPAs?

Task 1 - Deleting voice recordings

Users' Approach. Participants approached the task using two primary methods: voice commands and the mobile app. The majority



(a) Task 1 results using Amazon Echo



(b) Task 1 results using Google Home

Figure 2: These are the results of Task 1 for both assistants.

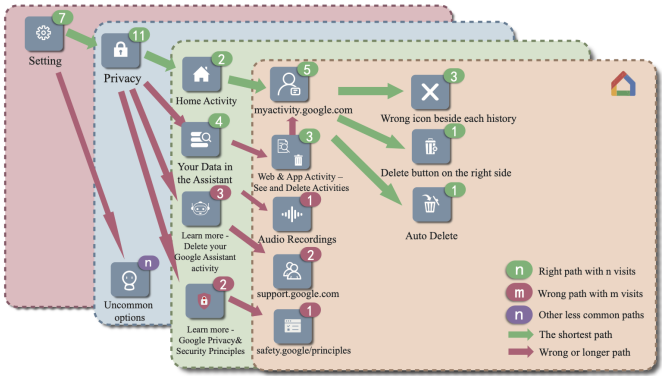


Figure 3: Users' paths when attempting to delete using the Google Home Mobile app. The green arrows shows the path that led to successful completion of the task, while the red ones shows unsuccessful paths. Different colored panes shows different screens users need to navigate to access different settings.

of participants (13/20) initially attempted to complete the task using voice commands. However, only four were successful on their first attempt, while eight failed and then switched to the mobile app. One participant in the Amazon group was unsuccessful and was unaware that the task could be completed using the app. Figure 2 summarizes these results. From those who used the app as their first approach, four successfully completed the task, while one did not. All participants who used both methods eventually succeeded in deleting the recordings. Figure 3 and Figure 4 show the success paths for participants using the Google Home mobile app and the Alexa app, respectively. Figure 4 is in Appendix B.

Participants who initially attempted to use voice commands cited prior experience, convenience, and the natural interaction with the assistant as their primary reasons. *"I think this way is more convenient and quicker because I don't have to do it myself if I just say it. If it can understand, it should execute my command immediately"* (P11).

However, two participants noted that voice commands were the only method they were familiar with, while another mentioned that their assistant was not easily accessible. P6 noted: *"The device is not physically accessible."* P16 and P17 were not aware that they could also delete from the mobile app: *"Yes, I prefer using voice commands. Is it possible to do it without voice commands?"* (P16) and *"I never knew another way to do that. And it's a smart speaker, so I think I can speak to it"* (P17).

Participants who opted for the mobile app cited prior experience, concerns that the assistant may inaccurately interpret their commands, habitual preference, and familiarity with the deletion option in the app. P8 explained that their choice was driven by habit and the perception that voice commands may not always execute correctly: *"It's a habit, and I feel that directly instructing the device might not clearly convey which specific data I want to delete. Manually selecting the entries gives me clearer control over what I am deleting, like choosing specific times or entries."* Similarly, P20 stated that because they initially set up the device using the app, they preferred using it for certain tasks: *"Because usually I set up Bluetooth devices through my phone, like headphones, speakers, and GoPro. I might try voice commands, but I actually prefer using app to operate"* (P20). Two participants, P12 and P14, stated that they used the app because the app assured that the voice recordings were deleted, but they were not convinced the voice command deleted the recordings: *"I prefer manual deletion [app] to ensure the records are indeed removed."* (P12) and P14 noted: *"This gives me more control and assurance over my data privacy."*

User Experience. Participants expressed mixed feelings about the usability of the deletion process. While some (*n=8*) found it straightforward, others (*n=7*) faced challenges related to unclear navigation, confusing voice assistant interactions, and a lack of confirmation. We did not find any notable differences between the two assistants.

Perceived Simplicity & Ease of Use: Some participants (*n=8*) perceived the deletion process as straightforward and easy to navigate, particularly those with prior experience and who accidentally discovered the audio deletion feature. P1 acknowledged that while the task took some time, they were willing to invest the effort because

deleting personal data was important. *It's simple, but it took me some time. However, once I decide to delete my data, it means that deletion is very important to me, so I'm willing to spend time on it."* Others, for example, P19 found the process intuitive: *"The logic behind the process is straightforward, and I can easily find what I'm looking for."* However, they further acknowledged that there is some learning that needs to happen before knowing how to delete: *"For new users, there might be a learning curve, but for me, it's pretty straightforward and easy to use."* Two participants (who used voice command) stated that the process was simple because they simply had to tell the assistant what to do without needing to go through the app: *"Just talking to it directly, and I don't need to do this through apps or other things"* (P15). However, others felt the app was simpler to use than the voice interface: *"The process is relatively simple... Through voice, it seems impossible to directly delete data"* (P20).

Navigation challenges: Despite some participants finding the process straightforward, other participants (*n=4*) struggled with the interface. They described it as counterintuitive and challenging to navigate. Some explained that they had to go through multiple steps to access the deletion settings. P8 highlighted confusion about where data was stored, explaining that they initially searched under the "Activity" section, which seemed logical but did not contain recordings. They noted, *"The design isn't very intuitive. It was difficult to find because you have to go to the Privacy settings and then navigate to the official website to view the activity records."* Others expressed frustration with redundant or unclear sections, such as Alexa Privacy, Notifications, and Assistant, which made it difficult to determine where deletion options were located: *"It feels like everything is mixed together, and it's hard to see clearly"* (P5). The complex navigation process and inconsistent organization of options led many participants to struggle in locating the correct deletion function.

Voice interaction challenges: Several participants (*n=4*) found the voice assistant's responses unhelpful, inconsistent, or confusing when attempting to delete their recordings. Some reported that the assistants misunderstood their requests or provided unrelated responses. For instance, P13 said, *"I can't seem to get it to delete the recordings as I want. For example, I tried saying, 'Alexa, delete voice recordings,' but it responded with unrelated actions like playing music"*. Others, for instance, P9 stated that the response or instruction to delete recordings was not clear: *"The response wasn't very clear. It was too long and didn't provide a straightforward answer."*

Lack of clarity and confirmation: Some participants (*n=3*) were unsure whether their data had been successfully deleted, as the assistants did not provide sufficient confirmation. P9 noted, *"It didn't feel secure. There was no clear confirmation of deletion, and there wasn't enough feedback to know if the operation was successful. It felt like the feature was just a facade."* Other participants encountered conflicting responses from the assistant. P3: *"The good part is that Alexa at least told me where to go to find my voice data. The downside is that it also told me, 'I don't know.'."* These inconsistencies led some participants to doubt the effectiveness of using voice commands for deletion, forcing them to rely on the mobile app instead.

Suggested Improvements. In terms of improving the deletion of recordings, other than requesting for better voice interaction for deletion, both Google Home and Amazon Alexa owners suggested improvements in app navigation and deletion options. Focusing on the mobile app, participants suggested that deletion mechanisms should not be buried under any settings and that the labeling in the app should be simpler. P1 suggested that the deletion of recordings should be easy to find in the Alexa app, *"I don't think it should be buried within the privacy settings in the software, requiring multiple steps to enable."* Other participants, such as P8, a Google Home Mini owner, suggested improvements in the level of control users have when choosing what to delete. They wanted more options for managing the deletion of their recordings, *"Yes, it should have the option for batch deletion. There should be a way to select a period, like 18 months, for automatic deletion or something similar"* (P8).

Task 2 - Deleting transcripts or history of commands

Users' Approach. While deleting voice recordings also removes the accompanying transcripts as they are combined, we still asked participants to delete these text records to examine whether they would use the same method or adopt a different approach as to the first task. Overall, many had completed the task when completing task 1, but we found that the majority (*n=15*) were not aware that they deleted the scripts as a result of deleting the voice recordings. For those who attempted to complete the task, they still employed the two primary methods: voice commands and the mobile app. However, some were surprised to learn that the assistant stored text transcripts and questioned their purpose. *I wasn't aware that they also have text"* (P18).

While the majority of participants used voice commands, some attempted to specifically instruct the assistant to delete transcripts, which often led to misunderstandings or misinterpretations by the device. For instance, P9 asked: *"Can you delete the transcripts of my interaction history?"* which the Google Home assistant responded, "I couldn't find anything related to transcripts of my interaction history." This led to the user being confused as to why the assistant could not fulfill their request.

Similar to some who used the voice command, participants who used the app specifically searched for text transcripts, which often led to failing the right screen to complete the task. *"The interface does not clearly indicate the correct option for deletion"* (P16). They were aware that both voice recordings and transcripts were combined and could be found in the same location, enabling them to complete the task more easily. *"I'll go into Privacy, then review activity history. I think this includes the text records, and I can delete them there. It also has the voice records. In the previous interface, I didn't notice that it had both voice and text records together. I usually just clear everything without looking at each specific entry"* (P6).

One participant (P10) was unable to complete the task using either method, explaining that they typically delegate such complex processes to other members of their household: *"Because I usually don't know how to do these things. It's mostly my son who handles it"*. They tried to ask the assistant, *"'Alexa, can you delete the data?'...It doesn't seem to work. What does it mean? Is it saying I need to delete it on my phone?"* While the assistant gave him instructions to delete from the app, they could not do it: *"I didn't quite understand. I think it might be telling me to delete it from my phone, but I'm not sure. I don't use smartphones well, so I'm not sure how to delete it."*

User Experience. While some participants found the task easy to complete, others encountered several challenges related to both the voice and app interfaces. We did not observe any notable differences between users of different assistants. Some participants noted that the assistant struggled to understand their deletion requests: *"It should at least inform me if it can't find the information, or maybe it didn't understand the request. As a user, after trying a few times, I would assume that's the final answer"* (P9). Participants who used the app and could not complete the task reported difficulties with navigation, terminology, unclear logic, and rushed confirmation screens. Many expressed frustration over having to go through multiple steps to locate the correct option: *"The process should be more straightforward, with an option in the main interface to show user activity records directly. It shouldn't require navigating through settings and privacy options to find the delete option. These steps are too deeply buried and not user-friendly"* (P8). Moreover, some participants found the confirmation screen too brief, making it difficult to complete the task: *"The cancellation option was too rushed, appearing for just one or two seconds, which may not be enough time for new users to react"* (P8).

Suggested Improvements. Both Google Home and Amazon Alexa users suggested improvements in three key areas: app navigation, real-time feedback, and voice interaction during the deletion process. Participants focusing on voice interaction recommended that the assistant should guide users through the deletion process, confirming their choices or allowing them to select specific types of data for removal: *"The best solution would be integrating it into the voice module. For example, if I ask it to delete my data, it should ask me whether I want to delete voice recordings, text, or select from specific options or delete everything before I confirm my choice"* (P1).

Others emphasized the need for the interface to provide feedback on whether the deletion was successful, *"The UI didn't show the history that I deleted. I think some kind of confirmation should tell me that you have deleted those things"* (P15). P6 asked for clarifications during deletion: *"It should clarify whether the deletion is permanent or if there is a recovery option, like a recycle bin."* Lastly, participants called for improved navigation, arguing that data deletion is a common function and should be more accessible. P3 also suggested merging redundant screens to minimize confusion: *"There is redundancy in these two sections. Why do the two categories in Alexa Privacy (i.e., Review Voice History and Review Activity History) contain identical data? These sections could be merged into a single data record with different filtering options instead of repeating the same data across multiple categories. The delete button and menu logic should be unified to ensure consistency across different sections."*

Task 3 - Resetting the assistant to default settings

Users' Approach. When it came to resetting the assistant to default settings, participants adopted three methods: voice commands, the mobile app, and the physical buttons on the device. Those who attempted voice commands explained that this was their natural way of interacting with the assistant: *"I always use voice ... I think this is a very outdated process because, in my vision of the future, smart devices should have fewer physical buttons"* (P1). Others stated that they expected the voice command to work as it had worked for the first two tasks: *"Because that's how I did it for the first two tasks."* (P5) However, participants who attempted to reset the device using voice commands were unable to complete the task. Instead, the assistant provided instructions on how to reset it to default settings. In some cases, the voice command was misinterpreted as an attempt to de-register the device. Two participants (i.e., P2 & P3) stated that they preferred the manual method over voice commands, arguing that it helps prevent accidental resets: *"It's not a good idea to use voice. What if it's triggered by mistake?"* (P2) and P3 said: *"This design is quite effective in preventing accidental resets..."*

Participants who opted to use the mobile app did so either based on prior experience or personal preference. *"I prefer using the app..."* (P16) and P9 explained: *"Based on my previous experience, I would go to the app to reset to factory settings. It might send a command via Bluetooth or another method, and then the smart speaker would proceed with a factory reset."*

Finally, participants who chose to use the physical buttons stated that they had previously employed this method to reset their devices: *"The second method is on a physical level, where some devices have a physical button that can be long-pressed to reset the entire device."* (P2) Most participants who selected this approach were able to successfully reset the device (7/10). However, few did not complete the process correctly.

User Expectation. We asked participants what they expected regarding the outcome of the reset process. The majority (n=10) expected the device to erase all stored data, with some emphasizing that it should restore the device to its original, out-of-the-box state: *"I expect all user data to be deleted, including interaction records with the device, stored household information, and other device-related data. The device should be restored to a state as if it were new."* (P8) However, one participant (P19) expressed uncertainty regarding the scope of the reset. While they expected the deletion of data from the device itself, they were unsure whether this action would also affect their entire Amazon account and devices connected to the assistant: *"I wonder if resetting the speaker affects only the data associated with the speaker, not my entire account. It would be unreasonable if resetting the speaker deleted all personal data, including interactions with other devices. Resetting should only apply to the speaker, not the whole account. If I choose to delete my account data from the app, then all data might be removed, but resetting the speaker alone shouldn't affect the entire account"* (P19).

User Experience. Participants had differing experiences with the reset process. While some found it straightforward (*n=4*), others encountered challenges and described it as counterintuitive. Participants who found the task easy either had prior experience or followed instructions provided by the assistant, the app, or an online source. Some participants noted that though they were able to complete the process, they first had to learn how to do so, emphasizing that the steps were not necessarily intuitive for all users. *"I'm very familiar with this because I've done it many times before. For older models, you might need to use a pin to press the reset button, but for newer models like Alexa, you just hold down the power button to reset it. The task can be easily completed. You see, I just need to press it in like this"* (P19).

Participants also found the process challenging due to its lack of clarity, their unfamiliarity with the procedure, or difficulty locating the reset feature within the app. Some also noted that the

process lacked indicators confirming whether it was being carried out correctly or had been completed successfully. P19 argued that since data reset and deletion are not commonly used functions, they may not be immediately clear to all users: *"These features aren't commonly used or prominently displayed. They should be designed well enough that when I need them, I can find them and accomplish my goal with minimal effort. It's unrealistic to expect everything to be immediately obvious at first glance; that would be excessive."*

Others (*n=5*) reported difficulty finding the factory reset option within the app's interface: *"I found it a bit difficult to locate the factory reset option. I tried looking in the privacy and assistant settings, but there wasn't a clear option for factory reset. I could only find options for activity records or adding/removing users. There wasn't a dedicated factory reset option"* (P8). Some participants also highlighted that the instructions provided were not easy to follow: *"It only has images and text. It would be better if it also had a link to a YouTube video. When I was reading the instructions for the factory reset, I felt a bit confused. It wasn't very straightforward"* (P9).

Lastly, some participants (*n=3*) cited that the lack of feedback on the reset process led to uncertainty about whether it was completed successfully: *"This seems to be taking a long time. I'm not sure if it's working. I'll look it up online. Okay, let me check. ... Sorry, I got it wrong. It says I should hold the power button. But it didn't work earlier; I'm not sure why. I saw online that holding the power button should work. Is this right? Did I do it correctly?"* (P18).

Suggested Improvements. Participants proposed several ways to improve the reset process. Some suggested the use of voice command, while others suggested improvement of the app to make the default reset feature easier to locate. While some participants expressed concerns about accidental deletions through voice commands, others argued that voice commands could still be used if they required explicit user confirmation before executing the reset: *"I think voice commands [would] require repeated confirmation or specific confirmation phrases, such as, 'Alexa, I confirm the deletion of my data and understand this is an irreversible action' that would provide sufficient security"* (P1). Regarding the app, participants emphasized the importance of making the reset feature more visible and providing clearer navigational instructions: *"Clear delete options and explanatory text are very helpful. If I could directly find the delete option within the privacy settings in the app without needing to navigate to another webpage, it would be more convenient"* (P8).

Users' Strategies to Deleting in SPAs

In this section, we report the strategies participants employed when participants deleted outside the study environment, as well as the behaviors we observed in the laboratory setting. Participants reported searching for information online or consulting individuals who had originally set up the devices to learn how to delete data. During the study, we observed participants adopting various approaches to complete deletion tasks, particularly through voice commands and mobile app interfaces.

For Task 1 and Task 2 (Deleting voice recordings and scripts), most participants initially attempted to use voice commands. When unsuccessful, they transitioned to deleting data via the smartphone app or asked the assistant for guidance on how to complete the task. Only one participant indicated that they were unaware of any alternative deletion methods. Among those who initially attempted deletion through the mobile app, failure prompted them to search for solutions online; none of these participants asked the SPA for assistance.

For Task 3 (Resetting the assistant to default), when participants were unable to reset the assistant using either voice commands or the mobile app, the majority asked the assistant for instructions on how they could complete the task. However, some participants who failed to complete the task via the app opted to search online rather than consult the assistant. Similarly, participants who were unsuccessful using voice commands first switched to the app and, but when they failed again, they resorted to searching online.

4.3 RQ3: What do users want with regard to data deletion in SPAs?

Our participants suggested several ways to improve the usability and effectiveness of data deletion in SPAs.

4.3.1 Clear voice instructions and seamless communication. Several participants (*n=9*) suggested improvements to the voice interface, particularly regarding instructions around data deletion. They explained that during the tasks, the assistant either failed to understand their commands or provided responses that were unclear. Some participants also emphasized the need for the assistant to be more intelligent and for communication to be more seamless. For instance, P1 suggested: *"I think the guidance for voice commands should be clearer, such as where to navigate and what options will be available. I also hope Alexa can be smarter, as it often fails to understand my questions. It would be great if it could integrate more advanced AI models more quickly."*

4.3.2 Improvements to mobile applications and web interfaces. Most participants (*n=11*) expressed a desire for improved mobile app, particularly in terms of navigation. Others reported challenges related to the long explanation in smaller font size under each option, which made interaction with the app more challenges. They emphasized the need for a more user-friendly interface. P20 noted: *"Regarding control from the app, I think it could be more user-friendly. The labels provided could be clearer."* Similarly, one participant (P16) raised concerns about the web interface of their service provider, particularly regarding the length and complexity of the text presented: *"They provided a lot of information, but for non-native speakers, it's not very appealing to read. It doesn't really help much, and I didn't find it very engaging."*

4.3.3 Transparency around data retention and the need for data recovery mechanisms. A few participants (*n=4*) highlighted the importance of clear information about data retention. As reported earlier, several participants were unsure whether deleted data remained in the system or whether they could recover it after deletion. Some suggested the introduction of recovery mechanisms, such as a "recycle bin," to provide clarity on whether data was still accessible or permanently erased. For instance, P6 proposed: *"The confirmation message should include whether the deletion will result in the permanent loss of data or if the data will be moved to a recycle bin for potential recovery."*

4.3.4 Feedback, confirmation, and prompts. Several participants noted that the deletion process, particularly through the voice interface, was missing some interaction, and it required additional prompts and feedback. They suggested that the assistant should request confirmation before proceeding with deletion and provide feedback upon completing the task. Moreover, some participants (n=4) emphasized the need for more detailed feedback regarding the deletion process, including retention periods, data destinations, and whether deletion was complete. P2 explained: *"Detailed feedback. Like what kind of data I have deleted, how long, and how big it is. Just information about the deleted data. I'm expecting Alexa to give me some feedback on what she did."* Similarly, P9 emphasized the need for confirmation mechanisms to improve user understanding of the deletion process: *"Confirmation would help me understand the progress and extent of the deletion process. For example, it would be helpful to know if the data on the server is being deleted and if the deletion is immediate or has a delay."*

4.3.5 User control over deletion. Some participants (*n=2*) expressed a desire for greater control over SPAs deletion process. They called for mechanisms that would allow users to manage deleted data, including access and recovery options. They argued that such control is essential for transparency and privacy. P6 suggested: *"The user or account holder should have access to the data deletion process and any recovery options, such as a recycle bin. This ensures that only the person who owns the account or has the proper permissions can manage their data, whether it's deleting, restoring, or reviewing what's been removed. This kind of control is essential for privacy and transparency."*

4.3.6 Assurance and accountability from service providers. Some participants (*n=2*) emphasized the need for transparency, accountability, and assurances from service providers regarding data deletion. They expressed concerns that companies often mishandle or misuse user data without clear accountability. Some participants suggested that the lack of transparency about data retention may be intentional, as companies seek to avoid being held accountable. P9 highlighted that without proper monitoring, companies may prioritize profit over user privacy: *"Companies aren't clear on how to provide these functionalities. There's a very blurry line in between. For companies, if they're not held accountable, they might neglect this aspect altogether, focusing solely on profit and convenience."*

4.3.7 Starter pack - setup process and guides. Some participants (*n=3*) suggested increasing the visibility of data deletion options and usage of mechanisms during the setup process of SPAs. They recommended that key features related to data deletion be included in the product's packaging or the setup manual to ensure users are immediately aware of their options. For instance, P1 proposed: *"I think a simple instruction card should be included when purchasing the product, explaining data operation mechanisms or protection measures, so users can see it immediately."* P2 emphasized the need to move deletion-related information away from privacy policies and instead place it in more accessible locations where users can easily find it: *"It shouldn't be put in the long privacy policy. It should be clear and simple and put in a place where you see it when you start using it, like the guide page."*

5 Discussion

We discuss our results and lessons learned below.

5.1 Google Home vs Amazon Alexa

While we did not set out to compare deletion usability across platforms, we observed some meaningful patterns between Amazon Alexa and Google Home users. Among Amazon participants, many attempted using voice commands to delete data, yet none checked whether this functionality was enabled beforehand. Alexa users who used the app generally found the correct deletion path more easily than Google Home users, who explored more menu options before locating the relevant setting. This may be due to Google offering more granular privacy options, which, while detailed, appeared to increase navigation complexity. We observed no meaningful difference in success based on whether users had set up their assistant themselves, suggesting that technical familiarity may not strongly influence deletion task success. Of those who first attempted deletion via voice (Amazon: 7, Google: 6), three Amazon users were successful compared to just one Google user. When switching to the app, all Google users succeeded, while three Amazon users still failed, two of whom returned to using voice. Among those who used the app on their first attempt (Amazon: 2, Google: 4), both Amazon users succeeded, whereas two Google Home users succeeded and two failed. Overall, participants across both platforms often switched between interfaces, especially from voice to app when the initial deletion method failed, indicating that users adapt their strategies when deletion is not immediately successful.

5.2 Voice Deletion

Our findings extend and confirm prior research on SPAs, particularly regarding data deletion. Previous studies [9, 25] have shown that SPA users are often unaware of data deletion mechanisms or do not know they have the option to delete their data. Our findings suggest that even when users are aware, they may not always be able to delete their data due to their perceptions of data deletion and various challenges. While users find the voice interface intuitive, it does not always function as expected, leading them to seek alternative ways to delete data. Many attempted to use the app instead, but some face navigational issues. Figure 5 shows an example of navigational struggles of Amazon Alexa App users. These findings highlight the need to improve deletion mechanisms, particularly the voice interface, as many users find it intuitive and prefer to use it. We believe that improving the voice interface for data deletion will also benefit bystanders, as SPAs are typically connected to the app on the primary user's phone. Moreover, voice deletion is particularly important in cases where the primary user or administrator of the device can monitor how other household members interact with it, potentially violating their privacy. We argue that improving voice deletion mechanisms can help reduce power imbalances, particularly in situations of intimate partner violence.

5.3 Transparency Around Data Deletion

Many of our participants lacked clarity on whether deleted data was permanently removed or retained by service providers. This finding aligns with prior studies (e.g., [31, 38, 43]) indicating that users usually assume that deletion means complete removal of data

from the system, when in most cases, data may still be retained for operational or legal purposes. This highlights the need for greater transparency regarding data deletion and retention policies. To improve user awareness around deletion, service providers must design clear and accessible deletion policies that support users' mental models of data deletion. As suggested by participants, this information should not be restricted to privacy policies but should instead be integrated into easily accessible locations, such as system settings, app dashboards, or voice-guided instructions. We believe that increased transparency in data deletion will not only reduce user confusion but also increase trust in SPAs. Future research should explore how data deletion transparency influences users' trust in the adoption and continued use of SPAs.

5.4 Deletion Controls

Our results indicate that SPA users hold varied and often incomplete mental models of data deletion. This confirms findings by Abdi et al. [1], which demonstrated that most SPA users lack a good understanding of the SPA ecosystem, including how data is stored and deleted. Our study extends this by showing that misaligned mental models directly affect users' expectations around deletion of data, often leading to confusion about what happens after they request deletion. We also discovered several mental models that are common in other platforms like cloud storage or computer operating systems to SPAs. This suggests a need for service providers to introduce deletion controls that align with user expectations, such as a recycle bin feature. Implementing such recovery mechanisms would provide users with greater autonomy over their data while supporting their understanding of the deletion process. The alignment of SPA deletion mechanisms with existing user mental models could reduce errors and align their expectations, thus improving user experience.

5.5 Feedback and Confirmation Mechanisms

Our participants wanted more explicit feedback and deletion confirmation messages, including whether data deletion was immediate, delayed, or reversible. We believe these expectations are also transferred from other platforms where these features are common. To enhance trust and assure deletion, this calls for real-time feedback mechanisms that will give users status or information about deletion. This info could include deletion timelines, retention policies, and potential recovery options as suggested in other prior studies (e.g., [31, 39]). It is also important that this is integrated into the voice interface since voice is the primary mode of interacting with SPAs. Future research should focus on improving the voice deletion mechanisms.

5.6 Integrating Data Deletion Information into Setup Processes

Some participants suggested that data deletion options should be made more visible during the onboarding and setup process, rather than being buried within app settings or privacy policies. We also believe that making deletion settings part of the initial setup could serve as an early awareness mechanism, ensuring users understand their deletion options from the outset. Prior studies have shown that the majority of users are not even aware that they could delete from their assistants. We believe that auto-delete features, both available in Amazon Alexa and Google Home, should be integrated in the setup process to also raise awareness. Also, this could be enabled by default and allow users to turn off or opt out. The stater guides or onboarding process can also allow users to customize retention periods. Future studies should leverage the Contextual Integrity Framework [35] to explore users' preferred default deletion settings and how these preferences vary based on contextual factors.

5.7 Deletion Privacy

Consistent with prior studies, our study indicates that users delete data from SPAs primarily to preserve their privacy. Our participants expected deleted data to either be completely removed or made inaccessible. Since deleted data may contain personal or sensitive conversations, it is critical that service providers implement effective mechanisms to protect against unauthorized access to deleted data. We recommend that service providers adopt secure deletion practices, including:

- Ensuring deleted data is permanently removed after the retention period.
- Preventing unauthorized internal access to deleted data.
- Providing transparency regarding whether deleted data is anonymized, retained, or repurposed for other uses.

We also believe that if deleted data is not effectively protected, it may expose users to risks, as reported previously, including data breaches, third-party access, or misuse [14, 28–30, 33]. Companies must implement policies to ensure that when retention periods are over, data is securely and irreversibly destroyed to protect user privacy.

6 Conclusion

In this paper, we investigated the usability of data deletion mechanisms in Smart Personal Assistants and examined users' perceptions of the deletion process through a qualitative study with 20 participants. Our findings reveal that users hold diverse and often incomplete mental models regarding data deletion, recovery, and the types of data collected and stored by SPAs. Data deletion is not always a straightforward process, and while the voice interface is the most preferred method, it does not consistently yield the intended results. We also discussed critical concerns surrounding transparency, accountability, and user control over deleted data, emphasizing the responsibility of service providers to address these issues. Finally, our work contributes to ongoing efforts to address the asymmetry between service providers and users regarding data management. Future research should explore how improving usability, transparency, and control influence user trust, engagement, and long-term adoption of voice-assisted technologies.

Acknowledgments

We thank the anonymous reviewers for their helpful feedback. We are also deeply grateful to our participants. Lastly, we would also like to thank Dr Rui Huan for her feedback on the paper. Cheng Cheng is a PhD student supported by EPSRC CDT Funding EP/Y035313/1. Dr Kopo Marvin Ramokapane was supported by REPHRAIN EPSRC EP/Y036611/1.

References

[1] Noura Abdi, Kopo M. Ramokapane, and Jose M. Such. 2019. More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, 451–466.

[2] Noura Abdi, Xiao Zhan, Kopo M Ramokapane, and Jose Such. 2021. Privacy Norms for Smart Home Personal Assistants. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–14.

[3] Akshay Awasthi, Huw OL Read, Konstantinos Xynos, and Iain Sutherland. 2018. Welcome Pwn: Almond Smart Home Hub Forensics. *Digital Investigation* 26 (2018), S38–S46.

[4] Frank Bentley, Chris Luvogt, Max Silverman, Rushani Wirasinghe, Brooke White, and Danielle Lottridge. 2018. Understanding the long-term use of smart speaker assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–24.

[5] Virginia Braun and Victoria Clarke. 2012. Thematic Analysis. In *APA Handbook of Research Methods in Psychology, Vol.2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological*, H. T. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, and K. J. Sher (Eds.). American Psychological Association, Washington, DC, USA, 57–71.

[6] M. A. Cascio, Eunlye Lee, Nicole Vaudrin, and Darcy A. Freedman. 2019. A Team-based Approach to Open Coding: Considerations for Creating Intercoder Consensus. *Field Methods* 31 (2019), 116 – 130. https://doi.org/10.1177/1525822X19838237

[7] Janet X Chen, Allison McDonald, Yixin Zou, Emily Tseng, Kevin A Roundy, Acar Tamersoy, Florian Schaub, Thomas Ristenpart, and Nicola Dell. 2022. Trauma-informed computing: Towards safer technology experiences for all. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–20.

[8] Kason Ka Ching Cheung and Kevin W. H. Tai. 2021. The use of intercoder reliability in qualitative interview data analysis in science education. *Research in Science & Technological Education* 41 (2021), 1155 – 1175. https://doi.org/10.1080/02635143.2021.1993179

[9] Eugene Cho, S Shyam Sundar, Saeed Abdullah, and Nasim Motalebi. 2020. Will deleting history make alexa more trustworthy? effects of privacy and content customization on user experience of smart speakers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[10] Hyunji Chung, Michaela Iorga, Jeffrey Voas, and Sangjin Lee. 2017. Alexa, Can I Trust You? *Computer* 50, 9 (2017), 100–104.

[11] Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What Can I Help You With?" Infrequent Users' Experiences of Intelligent Personal Assistants. In *Proceedings of the 19th international conference on human-computer interaction with mobile devices and services*. Association for Computing Machinery, New York, NY, USA, 1–12.

[12] Jide S Edu, Jose M Such, and Guillermo Suarez-Tangil. 2020. Smart home personal assistants: a security and privacy review. *ACM Computing Surveys (CSUR)* 53, 6 (2020), 1–36.

[13] Victoria Elliott. 2018. Thinking about the coding process in qualitative data analysis. *The Qualitative Report* 23, 11 (24 Nov. 2018), 2850–2861. https://doi.org/10.46743/2160-3715/2018.3560

[14] Federal Trade Commission. 2016. Ashley Madison Settles with FTC and State Regulators for Failing to Protect Users' Data and Misleading Consumers. https://www.ftc.gov/news-events/news/press-releases/2016/12/ashley-madison-settles-ftc-state-regulators-failing-protect-users-data-misleading-consumers. Accessed: 2025-04-25.

[15] D. Garrison, M. Cleveland-Innes, Marguerite Koole, and James Kappelman. 2006. Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *Internet High. Educ.* 9 (2006), 1–8. https://doi.org/10.1016/J.IHEDUC.2005.11.001

[16] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods* 18, 1 (2006), 59–82.

[17] Hana Habib, Sarah Pearman, Jiamin Wang, Yixin Zou, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. 2020. "It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12.

[18] Hana Habib, Yixin Zou, Aditi Jannu, Neha Sridhar, Chelse Swoopes, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. 2019. An Empirical Analysis of Data Deletion and {Opt-Out} Choices on 150 Websites. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, USA, 387–406.

[19] Yue Huang, Borke Obada-Obieh, and Konstantin Beznosov. 2020. Amazon vs. my brother: How users of shared smart speakers perceive and cope with privacy risks. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[20] Iulia Ion, Niharika Sachdeva, Ponnurangam Kumaraguru, and Srdjan Čapkun. 2011. Home is safer than the cloud! Privacy Concerns for Consumer Cloud Storage. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 1–20.

[21] Sunyoung Kim and Abhishek Choudhury. 2021. Exploring Older Adults' Perception and Use of Smart Speaker-based Voice Assistants: A Longitudinal Study. *Computers in Human Behavior* 124 (2021), 106914.

[22] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–31.

[23] Sophie-Charlotte Lemmer. 2019. *Alexa, Are You Friends With My Kid? Smart Speakers and Children's Privacy Under the GDPR*. Graduate Student Research Paper 2018/9_6. King's College London, The Dickson Poon School of Law, Somerset House East Wing, Strand, London WC2R 2LS, United Kingdom. Available via SSRN; DOI: 10.2139/ssrn.3627478.

[24] Yijing Liu, Yan Jia, Qingyin Tan, Zheli Liu, and Luyi Xing. 2022. How Are Your Zombie Accounts? Understanding Users' Practices and Expectations on Mobile App Account Deletion. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 863–880.

[25] Nathan Malkin, Joe Deatrick, Allen Tong, Primal Wijesekera, Serge Egelman, and David Wagner. 2019. Privacy attitudes of smart speaker users. *Proceedings on Privacy Enhancing Technologies* 2019, 4 (2019), 250–271.

[26] Joseph A Maxwell. 2010. Using Numbers in Qualitative Research. *Qualitative inquiry* 16, 6 (2010), 475–482.

[27] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (Nov. 2019), 23 pages. https://doi.org/10.1145/3359174

[28] Mohsen Minaei, Mainack Mondal, and Aniket Kate. 2022. Empirical Understanding of Deletion Privacy: Experiences, Expectations, and Measures. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 3415–3432.

[29] Mohsen Minaei, Mainack Mondal, Patrick Loiseau, Krishna P Gummadi, and Aniket Kate. 2019. Forgetting the Forgotten With Lethe: Conceal Content Deletion from Persistent Observers. In *PETS 2019-19th Privacy Enhancing Technologies Symposium*. Sciendo (De Gruyter), Stockholm, Sweden, 1–21.

[30] Mohsen Minaei, S. Chandra Mouli, Mainack Mondal, Bruno Ribeiro, and Aniket Kate. 2021. Deceptive Deletions for Protecting Withdrawn Posts on Social Media Platforms. In *Proceedings of the Network and Distributed System Security Symposium (NDSS 2021)*. Internet Society, 11710 Plaza America Drive, Suite 400, Reston, VA 20190, USA, 462–479. https://doi.org/10.14722/ndss.2021.23139

[31] Ambar Murillo, Andreas Kramm, Sebastian Schnorf, and Alexander De Luca. 2018. "If I Press Delete, It's Gone": User Understanding of Online Data Deletion and Expiration. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 329–339.

[32] Mahin Naderifar, Hamideh Goli, and Fereshteh Ghaljaie. 2017. Snowball Sampling: A Purposeful Method of Sampling in Qualitative Research. *Strides in Development of Medical Education* 14, 3 (2017), –. https://doi.org/10.5812/sdme.67670

[33] BBC News. 2015. Ashley Madison infidelity site's customer data 'leaked'. https://www.bbc.com/news/business-33984017 Accessed on August 19, 2015.

[34] Kavous Salehzadeh Niksirat, Diana Korka, Quentin Jacquemin, Céline Vanini, Mathias Humbert, Mauro Cherubini, Sylvain Métille, and Kévin Huguenin. 2024. Security and Privacy with Second-hand Storage Devices: A User-Centric Perspective from Switzerland. *Proceedings on Privacy Enhancing Technologies (PoPETs)* 2024, 2 (2024), 22.

[35] Helen Nissenbaum. 2004. Privacy as Contextual Integrity. *Wash. L. Rev.* 79 (2004), 119.

[36] Kentrell Owens, Johanna Gunawan, David Choffnes, Pardis Emami-Naeini, Tadayoshi Kohno, and Franziska Roesner. 2022. Exploring Deceptive Design Patterns in Voice Interfaces. In *Proceedings of the 2022 European Symposium on Usable Security*. Association for Computing Machinery (ACM), New York, NY, USA, 64–78.

[37] Kopo Marvin Ramokapane and Awais Rashid. 2023. ExD: Explainable Deletion. In *Proceedings of the 2023 New Security Paradigms Workshop*. Association for Computing Machinery (ACM), New York, NY, USA, 34–47.

[38] Kopo Marvin Ramokapane, Awais Rashid, and Jose Miguel Such. 2017. "I feel stupid I can't delete...": A Study of Users' Cloud Deletion Practices and Coping Strategies. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association, Santa Clara, CA, 241–256.

[39] Kopo Marvin Ramokapane, Jose Such, and Awais Rashid. 2022. What Users Want From Cloud Deletion and the Information They Need: A participatory Action Study. *ACM Transactions on Privacy and Security* 26, 1 (2022), 1–34.

[40] Jane Ritchie, Liz Spencer, and William O'Connor. 2003. Carrying out Qualitative Analysis. In *Qualitative Research Practice: A Guide for Social Science Students and*

Researchers, Jane Ritchie and Jane Lewis (Eds.). SAGE Publications Ltd, 6 Bonhill Street, London EC2A 4PU, UK, 219–262.
[41] Margarete Sandelowski. 1995. Sample size in qualitative research. *Research in nursing & health* 18, 2 (1995), 179–183.
[42] Margarete Sandelowski. 2001. Real Qualitative Researchers do not Count: The use of Numbers in Qualitative Research. *Research in nursing & health* 24, 3 (2001), 230–240.
[43] Theodor Schnitzler, Christine Utz, Florian M Farke, Christina Pöpper, and Markus Dürmuth. 2020. Exploring User Perceptions of Deletion in Mobile Instant Messaging Applications. *Journal of Cybersecurity* 6, 1 (2020), tyz016.
[44] William Seymour, Noura Abdi, Kopo M Ramokapane, Jide Edu, Guillermo Suarez-Tangil, and Jose Such. 2024. Voice App Developer Experiences with Alexa and Google assistant: Juggling Risks, Liability, and Security. In *33rd USENIX Security Symposium (USENIX Security 24)*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 5035–5052.
[45] Madiha Tabassum, Tomasz Kosinski, and Heather Richter Lipford. 2019. " I don't own the data": End User Perceptions of Smart Home Device Data Practices and Risks. In *Fifteenth symposium on usable privacy and security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, USA, 435–450.
[46] Cayetano Valero, Jaime Pérez, Sonia Solera-Cotanilla, Mario Vega-Barbas, Guillermo Suarez-Tangil, Manuel Alvarez-Campana, and Gregorio López. 2023. Analysis of Security and Data Control in Smart Personal Assistants From the User's Perspective. *Future Generation Computer Systems* 144 (2023), 12–23. https://doi.org/10.1016/j.future.2023.02.009
[47] Maarten W. van Someren, Yvonne F. Barnard, and Jacobijn A.C. Sandberg. 1994. *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*. Academic Press, London; San Diego. 208 pages.
[48] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. 2011. "I Regretted the Minute I Pressed Share" A Qualitative Study of Regrets on Facebook. In *Proceedings of the seventh symposium on usable privacy and security*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 1–16.
[49] Eric Zeng, Shrirang Mare, and Franziska Roesner. 2017. End user security and privacy concerns with smart homes. In *thirteenth symposium on usable privacy and security (SOUPS 2017)*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, USA, 65–80.
[50] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–20.
[51] Lu Zhou, Wenbo Wang, and Keke Chen. 2016. Tweet Properly: Analyzing Deleted Tweets to Understand and Identify Regrettable Ones. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 603–612.

A Participants' Information

Table 2: Interview Participant Demographics

	No. of participants
Gender	
Male	11
Female	9
Age	
18 – 20	2
21 – 25	9
26 – 30	5
31 – 45	3
45 +	1
Educational Background	
High school/College course	3
Bachelors	9
Masters	7
PhD	1
Employment Status	
Full-time employed	7
Part-time employed	2
Student	13
Device Type	
Amazon Echo	10
Google Home	10
Period of Usage	
1 - 6 months	5
6 - 12 months	5
1 - 2 years	4
2+ years	6

Table 3: Participant usage and setup information

Participant	Device	Setup	Deleted before
P1	Amazon	Self	No
P2	Amazon	Self	No
P3	Amazon	Self	No
P4	Google	others	No
P5	Google	Self	Yes
P6	Google	others	Yes
P7	Amazon	Self	No
P8	Google	Self	No
P9	Google	others	Yes
P10	Amazon	others	No
P11	Google	others	No
P12	Google	Self	Yes
P13	Amazon	Self	Yes
P14	Google	Self	Yes
P15	Amazon	others	Yes
P16	Google	others	No
P17	Google	others	No
P18	Amazon	others	Yes
P19	Amazon	Self	Yes
P20	Amazon	Self	No

B Users' Navigation in the Assistants' Mobile Apps

Figure 4 shows the user paths attempted for deletion using the Alexa app to complete Task 1 - deleting voice recordings, while Figure 5 shows how participants navigated the Amazon App attempting to reset the assistant.

In Figure 5, since the only way to reset an Amazon assistant is by using the physical button on the device, and the app does not offer a reset option, participants' attempts to find it in the app were in vain. However, it is noticeable that users frequently visited "Settings and Device Settings," suggesting they "expected" to find the reset function in these sections.

C Screening Questions

(1) Do you agree to start the survey?
  - I agree, begin the survey
  - I do not agree to participate

(2) Which brand of smart speaker do you own?
  - Amazon Echo
  - Google Nest
  - Apple Homepod
  - Huawei Sound / AI speaker
  - Xiaomi Smart Speaker
  - Other (please specify):

(3) Which voice assistant do you use?
  - Amazon Alexa
  - Google Assistant
  - Siri
  - Xiaoyi
  - Xiaoai Tongxue
  - Other (please specify):

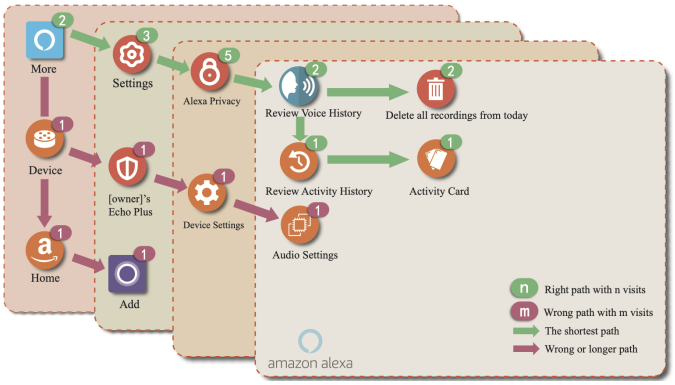


Figure 4: Users' paths when attempting to delete the using the Alexa mobile app. The green arrows shows the path that led to successful completion of the task, while the red ones shows unsuccessful paths. Different colored panes shows different screens users need to navigate to access different settings.

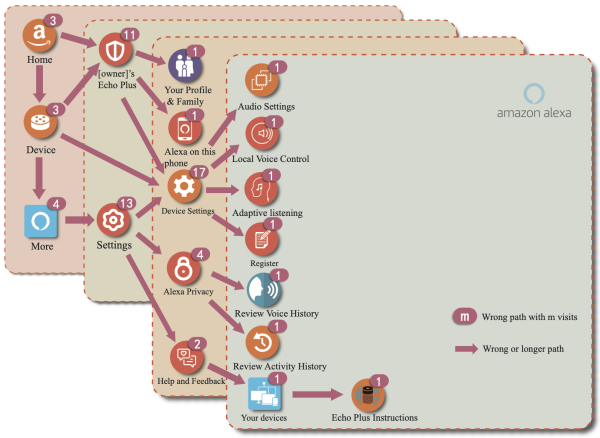


Figure 5: User path of Amazon Alexa users who failed to complete Task 3 using the App. Different colored panes here shows different screens users need to navigate to access different settings.

(4) Which voice commands do you use to awaken your personal assistant? (Please enter your answer)
(5) How long have you been using this device?
  - Less than a month
  - 1-3 months
  - 3-6 months
  - 6 months – 1 year
  - 1 year - 2 years
  - More than 2 years
(6) How often do you use this device per month?
  - I rarely use
  - I use it once a week
  - I use it several times a week
  - I use it every day
(7) How many people in your household use the device?
  - Only me
  - Me and someone
  - 3–5 people
  - More than 5 people
(8) What do you mainly use your device for?
  - Play music
  - Set alarms and reminders
  - Send messages and make calls
  - Shopping
  - Listen to audiobooks and podcasts
  - Manage other smart home devices
  - Other (please specify):
(9) Have you ever deleted data from your device?
  - Yes
  - No
(10) Have you attempted to delete data and encountered any difficulties?
  - Yes
  - No
(11) If you are willing to be interviewed in person, which device would you prefer to use for the interview?
  - Amazon Echo
  - Google Home
  - None
  - Other (please specify):
(12) What is your gender? (Please enter your answer)
(13) What is your age?
  - 18-20
  - 21-25
  - 26-30
  - 31-35
  - 36-40
  - 41-45
  - Above 45
  - Prefer not to say
(14) What is your current occupation?
  - Student
  - Employed full-time
  - Employed part-time
  - Self-employed
  - Temporarily unemployed
  - Full-time homemaker
  - Retired
  - Prefer not to say
  - Other (please specify):
(15) What is your highest level of education?
  - High school or equivalent
  - Completed some university, but no degree
  - Associate degree
  - Bachelor's degree
  - Master's degree
  - Doctorate or higher
  - Professional certification
  - No formal education
  - Prefer not to say

  - Other (please specify):

(16) If invited to take part in an interview study, would you participate?
  - Yes
  - No

(17) (If yes) Please provide your email address. This will only be used to contact you when you are invited to the interview.

D Interview Guide

D.1 Start-up Questions

- Which device do you own?
- What prompted you to start using it?
  - *[Follow-up]* How long have you been using it?
  - *[Follow-up]* How often do you use it?
  - *[Follow-up]* Besides you, who else in your household uses it?
  - *[Follow-up]* Who set up the device?
  - *[Follow-up]* What do you mainly use the device for?
  - *[Follow-up]* Which voice apps or skills do you use often, such as Weather, Uber and Spotify?
- Other than your device, do you own any other smart home devices?
  - *[Follow-up]* Do you use your device to control these other smart home devices?
  - *[Follow-up]* How useful is your smart speaker in terms of controlling your smart home devices?

D.2 User's Understanding of Storage

[Can you please ask the smart assistant a question or say something?]
  - *[Follow-up]* Could you please explain how you believe your command is processed?
- Do you think your voice recordings are saved?
  - *[Follow-up]* How long do you think the recordings are stored for?
  - *[Follow-up]* For what reason?
- Can you access your interaction data?
  - *[If yes, Follow-up]* Where can you access it?

D.3 Deletion Practices and Users' Understanding of Deletion

- What does it mean to delete data in smart speakers?
  - *[Prompt]* How do you define 'deletion' in the context of smart speakers?
  - *[Follow-up]* What do you expect to happen?
  - *[Follow-up]* Do you think it is good to have a confirmation after you delete something?
  - *[If yes, Follow-up]* Can you describe your ideal confirmation notification?
- What happens to data after you delete it?
  - *[Follow-up]* How possible is it to recover data after you have deleted it from the smart speaker?
  - *[If yes, Follow-up]* Who do you think can access deleted data?
  - *[If they don't mention themselves, Follow-up]* Why didn't you mention yourself? Do you think you can access it?
- Have you ever deleted something on your speaker?
  - *[If yes, Follow-up]* How do you normally delete data from the speakers?
  - *[Follow-up]* Is there any reason why you follow this method?
  - *[If no, Follow-up]* Is there any reason you have never tried?
  - *[Follow-up]* What kind of data do you think you could delete?

D.4 Hands-on Deletion Tasks

Task 1: Delete voice recordings

- How would you delete the voice recordings?
  - *[Follow-up]* Why do you want to do in this way?
  - Please try it.
- (If Alexa asks them to record a voice id, stop it, and ask)
  - *[Follow-up]* Do you think you need to record it? What would happen?
- (If they struggle with it)
  - *[Follow-up]* What would you normally do when you face some difficulties?
  - *[Follow-up]* Can you try it?
  - *[Prompt]* If it doesn't work, can you try other ways?
- Please describe the process you just tried for deleting the recordings.
  - *[If it is easy]* What do you think makes it easy?
  - *[If it is tricky]* What do you think makes it difficult?
  - *[Follow-up]* How clear or useful were the voice instructions?
  - *[Follow-up]* Do you think the explanations in small text were clear under the title?
  - *[Follow-up]* Were the prompts/confirmation clear during your deletion process?
  - *[Follow-up]* Did you notice the auto-delete feature?

Task 2: Delete the transcript of your interaction history

- How would you delete transcripts if you are to delete them?
  - *[Follow-up]* Why do you want to do in this way?
  - Please try it.
- Did you see the text transcript of your interaction history in this process?
  - *[Follow-up]* Where is it?
  - *[Follow-up]* What is the difference between voice recording and text transcript?
- How can this process be improved?
- can you describe your ideal scenario?

Task 3: Resetting the speaker to default settings

- What would you do if you were asked to reset the speaker to factory settings?
  - *[Follow-up]* What is your expected outcome?
  - Please try it.
- Please describe the process you just tried.
  - *[If participant only finds one approach, Follow-up]* Do you know other ways to do it?

- Do you think this process is easy to operate?
  - *[If it is easy]* What do you think makes it easy?
  - *[If it is tricky]* What do you think makes it difficult?
  - *[Follow-up]* How useful were the voice responses?
- How can this process be improved?

D.5 Conclusion Questions

- Now that you've tried these features, do you think you will use them in the future?
- What information do you think is useful to know about deleting from smart speakers?
- What kind of information needs to be shared with users to help them understand the deletion process?
- What is your opinion about this study?

E Interview Codebook

This codebook represents a series of themes and the corresponding codes observed in the study.

E.1 Data and Data Deletion Definitions in SPAs

Participants defining data deletion in the context of SPAs.

Deletion in terms of purpose.

- Removing all personal preferences and habits
- Removing data to protect it (abiding by regulations)
- Removing or erasing data for space
- Restoring the device to initial state

Deletion in terms of impact.

- Everything is completely erased permanently

Deletion is removing access to data.

- Removing access to data

Deletion in terms of recoverability.

- Data is still recoverable
- Data is not completely removed from the device

Voice recordings vs Textual scripts.

- Voice and scripts are useful
- Voice and textual scripts linked together and there are no differences between them
- Voice recordings contain more information

E.2 Data Deletion Practices, Perceptions and Expectations in SPAs

Deletion practices. Participants discussing reasons why the delete or not delete data from SPAs.

- Reasons for deletion - clear space
- Reasons for no deletion practices - not aware
- Reasons for no deletion practices - no privacy concerns
- Reasons for no deletion practices - does not use the speaker for sensitive activities

Perceptions of the deletion process. Participants discussing what happens when they request for data deletion.

- Data is completely deleted
- Data recovery is possible
- Data remains for a short while after deletion

Expectations after deletion request. Participants discussing what they expect after they request for data to be deleted.

- Data to be permanently deleted
- Device with deleted data (cleared or freed memory)
- Deleted data to be inaccessible
- Retention and recoverable
- Information of deletion
- Prompts and confirmation of deletion

E.3 Mental Models

Participants' various mental models of data deletion.

Perceptions of the concept of deleted data.

- Data should be removed completely from the system
- Deleted data is only marked deleted

Feeling about recovery of deleted data.

- Recovery is impossible
- Recovery should be controlled (Desire)
- If recycle bin exist, data recovery is possible
- Do not know whether data can be recoverable

Understanding of the access to deleted data from SPAs.

- Deleted data should not be accessed
- No one can access deleted data
- Who can access data after deletion - Service providers and developers
- Who can access data after deletion - Users
- Who can access data after deletion - Not sure

E.4 Concerns Around Data Deletion

Participants' concerns around data deletion.

Deletion not being complete and permanent.

- Data may not be anonymized properly after deletion
- Confirmation and transparency - Feedback on the data that has been deleted
- Service providers retain data without consent

Complete deletion is uncertain.

- Data is not deleted
- Cannot verify deletion
- Possible recovery after deletion
- No confidence in data deletion

Access to deleted data.

- Accessing deleted data without consent
- Accessing deleted data violates privacy

Lack of control over deletion. -

- Deleted data being used without their consent
- Deletion is out their control
- Service providers have power to decide when deletion
- Concern about limited SPAs mechanisms