What Users Want From Cloud Deletion and the Information They Need: A Participatory Action Study.

KOPO M. RAMOKAPANE, University of Bristol, United Kingdom
JOSE SUCH, Kings College London, United Kingdom
AWAIS RASHID, University of Bristol, United Kingdom

Current cloud deletion mechanisms fall short in meeting users' various deletion needs. They assume all data is deleted the same way—data is temporally removed (or hidden) from users' cloud accounts before being completely deleted. This assumption neglects users' desire to have data completely deleted instantly or prefer to have it recoverable for a more extended period. To this date, these preferences have not been explored. To address this gap, we conducted a participatory study with four groups of active cloud users (five subjects per group). We examined their deletion preferences and the information they require to aid deletion. In particular, we explored how users want to delete cloud data and identify what information about cloud deletion they consider essential, the time it should be made available to them, and the communication channel that should be used. We show that cloud deletion preferences are complex and multi-dimensional, varying between subjects and groups. Information about deletion should be within reach when needed, for instance, be part of deletion controls. Based on these findings, we discuss the implications of our study in improving the current deletion mechanism to accommodate these preferences.

CCS Concepts: • **Human-centered computing** → **User studies**; **Participatory design**; **User centered design**; • **Security and privacy** → *Social aspects of security and privacy*; *Usability in security and privacy*.

Additional Key Words and Phrases: Deletion, Data Deletion, Cloud Deletion, Deletion Preferences, Cloud Computing, Participatory Design, Cloud Storage, Cloud Deletion Mechanisms, Preferences, User studies

ACM Reference Format:
Kopo M. Ramokapane, Jose Such, and Awais Rashid. 2022. What Users Want From Cloud Deletion and the Information They Need: A Participatory Action Study.. 1, 1 (August 2022), 34 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Empowering users regarding data retention and deletion is very important in environments like the cloud, where users do not have direct control over the infrastructure. Clouds can provide significant challenges to users especially verifying the data handling practices of service providers. In fact, the current mechanisms for cloud deletion give users little to no control over how their data is disposed of [25, 44, 53, 54].

This work extends our previous work [54] on cloud deletion. We aim to investigate two issues; users' cloud deletion preferences and the information that may support users' deletion needs in the

Authors' addresses: Kopo M. Ramokapane, marvin.ramokapane@bristol.ac.uk, University of Bristol, 75 Woodlands Road, Bristol, BS8 1UB, United Kingdom; Jose Such, Kings College London, Bush House, 30 Aldwych, London, WC2B 4BG, United Kingdom, jose.such@kcl.ac.uk; Awais Rashid, University of Bristol, Bristol, BS8 1UB, United Kingdom, awais.rashid@bristol.ac.uk.

XXXX-XXXX/2022/8-ART $15.00
https://doi.org/10.1145/1122445.1122456

cloud. We seek to answer the following key research questions: (1) how do users want to delete cloud data? and (2) how do they want to be informed about deletion? Despite evidence [25, 44, 53, 54] that users have various deletion needs, current cloud deletion mechanisms fall short in meeting these needs. For instance, they give users little to no control over how their data should be deleted (i.e., whether data should be completely deleted or be recoverable). Prior studies informing the design of deletion mechanisms have primarily focused on understanding cloud deletion and users' understanding of online deletion. While these efforts are relevant for future designs and the development of deletion mechanisms, they lack insights into deletion preferences or contextual reasons why users may want data to be deleted in a certain way. They have only investigated the narrow issues of data deletion, such as why people delete and the challenges they face. With this focus, they have neglected users' other needs or the underlying reasons they want to delete.

Our previous study [54] has shown that cloud users' have different motivations to delete. For instance, a user may delete to tidy up their account, manage old data or delete to preserve their privacy. These different motivations for deletion may require different ways of deleting, for instance, while deleting just to tidy up one's account may not need to be complete (moving data to the 'trash' folder may be enough), deleting to preserve one's privacy may require deleted data completely removed from one's cloud account. Moreover, when one is unsure of the importance of a file, they may delete it but hope to recover it when needed. Existing literature does not provide insights on these different needs nor the ways to address them. Currently, users are limited regarding how they can delete their cloud data; all data stored in the cloud is temporarily removed from users' locality before being completely deleted. However, this is different to how users may want to delete their data. Cloud providers offer no help or controls to meet these different needs [53].

Regarding information about cloud deletion, our previous work [54] also suggested that users are not satisfied with how information about deletion is shared and distributed. For instance, participants stated that information about deletion, unlike other information like storage size, is only found in privacy policies and is not easily accessible. Nonetheless, privacy policy shortcomings are well understood in the literature. They have surfeit information, are difficult to understand and are usually tailored to demonstrate compliance with legal requirements [23]. Moreover, concerning data deletion, information about deletion is usually compact and short, missing other aspects of data deletion, which may impair transparency. None of the existing studies about the cloud provides insights on what information about deletion users consider important, and when and where it should be shared with them.

To address this gap, we asked the following research questions: (1) how do cloud users classify cloud data, that is, what kinds of data do they treat similar and different with regards to sensitivity and importance? (2) how do users want these data to be deleted, what preferences can be identified from this, considering (a) deletion under individual contexts and (b) social contexts – shared folders? Are these preferences consistent? and (3) is it feasible to design deletion mechanisms that satisfy their deletion preferences? To address the second research question, we examined the following: (1) what information about deletion do users consider important, (2) when do users want this information to be presented to them, and (3) where users prefer to find such information.

In summary, our work makes the following contributions.

- **Cloud data classification.** We identify three categories that users commonly use to classify data stored in the cloud. Cloud users generally categorize data under the following groups: (1) essential and sensitive, (2) less important and less sensitive, and (3) important but less sensitive. *Essential and sensitive* – this is a group of data items that users consider to be necessary or useful and private. *Less important and Less sensitive* – a group of data that users consider less valuable, easy to produce and less private, while *important but less sensitive* data

is the data they consider useful and hard to produce, but they are happy with other people knowing or seeing it.

- **Characterization of cloud user deletion preferences.** We identify and characterize cloud user deletion preferences with regards to deleting in individual and social contexts. Our analysis uncovers four characteristics of cloud user deletion preferences including their complexity and dimensions. Most critically, we discuss how the reason for deletion, the perceived importance (file utility) of the file being deleted, the size of the file, file sharing context, sensitivity and the storage size underpin these preferences.
- **An understanding of users' preferences regarding cloud deletion information.** We find that users consider information on cloud deletion scarce, not useful and that it is usually presented to them at the wrong time through a wrong channel. They prefer that technical information about deletion task (e.g., how deletion is carried out in the cloud) to be made available to them in blogs while information about *who has access to data stored or deleted from the cloud* and *what happens when data is deleted* should also be made available through other channels (e.g., frequently asked questions) other than privacy policies.

The rest of the article is organized as follows: we continue with the discussion of previous studies on cloud deletion in Section 2. Section 3 provides a detailed summary of our previous work on cloud deletion [54]. Section 4 gives a general overview of the approach behind our study. We then present the results of the first activity (data sorting) in Section 5, the second activity (deletion preferences) in Section 6, and the last activity (deletion information) in Section 7. Section 8 presents the discussion and implications of our study and several guidelines for developers of cloud storage platforms.

2 RELATED WORK

2.1 Cloud deletion

Both formal studies and anecdotal evidence suggest that deletion is essential to cloud users. Users delete for various reasons; however, failure to do so can lead to unintended disclosures, clutters, regrets and emotional trauma [35, 44, 53]. While there has been some work focusing on understanding challenges of deleting from other platforms like social media [2, 17, 22, 61], user studies on cloud deletion are novel and sparse [25, 44].

Most usability studies [3, 19, 33, 75] around the cloud focus on other aspects like sharing, synchronization, perceptions, and privacy. Earlier studies argued that not all users intendedly used the cloud storage; they were mainly accidental users—usually surprised to find their data in the cloud. The mismatch between users' expectations and the reality of using the cloud has been cited as the main reason. Users usually misunderstand file synchronization, sharing, and deletion. For example, in Capra et al.'s study [8], some participants manually synced their files to the cloud. Regarding deletion, previous studies have reported mixed perceptions. Ion et al. [19] reported that most users understood that cloud deletion was not instant and permanent; they believed files were still recoverable for some time after deletion. However, younger participants in the cross-generational study [3] believed that deletion was instant and permanent. In our previous study [54], some participants did not understand that deleting from a shared folder would affect all the collaborators. Moreover, Khan et al. explained that the misunderstanding of how shared folders work often leads to users refraining from deleting old files, even when they have the access rights to do so [25].

In a non-cloud context, Murillo et al. found that misunderstandings and unfounded expectations of deletion are limited to user interfaces [44]. Lack of understanding explanation or impact of deletion on the interface can lead to users not correctly assessing the effects of deleting messages [59].

Others [19, 44] also argue that limited understanding of what happens in the backend may lead to unexpected results. Some studies [44, 53] have noted that, in order for users to have a better understanding of deletion, they are expected to understand deletion concepts such as backend, timeliness, backup, derived information, completeness, anonymization, fine-grained and shared copies. However, these concepts are not known among users; after interviewing 36 cloud users, Ion et al. found that users do not understand and know timeliness or data retention [19]. While these studies highlight the need for understanding deletion, none of them suggests how users can acquire such knowledge or what are users' preferences regarding the types of deletion they want concerning particular types of data. Also, it is necessary to study how users delete data in settings when it is shared with others using the cloud service.

Offering complete deletion in the cloud is still a challenge. Hao et al. [14] argue that it is impossible to assure data deletion in the cloud using software-based approaches without physically tampering with the disk. Our previous work [53] has also identified and discussed the challenges of guaranteeing deletion in the cloud, highlighting that its salient infrastructure and operational features make it difficult to completely delete data from the cloud. Despite this, some efforts have been made to assure deletion in the cloud. For example, Tang et al. built a policy-based deployable cloud storage system (i.e., FADE) to protect deleted data [69]. Another scheme that offers fine-grained deletion was proposed by Mo et al. [43]. This scheme is based on a key modulation function (KMF) that allows users to delete individual data items without re-encrypt the rest of the data. Nonetheless, these cryptographic-based solutions work by denying access to the deleted data. They do not remove the deleted data from the cloud, which means the data can still be leaked through other attacks (e.g., brute force attacks or incorrect usage of cryptography libraries during implementation [49]). Moreover, current cloud services do not specify the type of deletion they offer or whether deleted data can still be recovered.

2.2 Classifying Data

Previous works on deletion report that deciding what data to delete or keep is a significant challenge for most users because it involves predicting the future—users are not good at predicting their data preferences over time. Moreover, they must keep in mind the type of data they are deleting, the importance of the data, and whether they will need such data in the future [3, 13, 54, 68, 73, 74]. This process is ongoing and usually lies on two distinctive extremes: hoarding (i.e., keeping data even if it is not valuable) and minimalism (i.e., avoiding storing too much data or regularly deleting data) [73]. Moreover, it is decided on a highly individual level dependent on context, service, and usefulness [44]. In group settings (e.g., shared folders), it is even more challenging because users must consider the future information needs of other collaborators [51].

Deciding on what data to keep or delete from the cloud may require users to understand how the cloud or cloud deletion works [44, 54]. Consequently, several researchers have previously investigated how users decide or classify data. For example, some users classify data through the lens of similarity [7]. Other times, their approaches are contextual [75]. Vioda et al. found that users segmented cloud data into multiple mental places, e.g., work or family [75]. In shared repositories, users usually view data as theirs or belonging to others but rarely as co-owned (i.e., common ownership) [51]. This usually leads to users forgetting about the existence of these data. However, when shown old data, most users tend to delete it rather than adopt other remediations [25]. To help users classify cloud data, Khan et al. [26] proposed a tool to identify if a file was sensitive and useful. Their results showed improvements over the state-of-the-art baselines, from 26% to 159%. This also improved the prediction of whether the user will keep or delete the file by 10%. However, Khan et al. [26] argue that basic metadata and demographic information are not particularly strong predictors of users' file-management decisions.

Other tools have been proposed to help users manage their cloud data. Khan et al. [25] developed a retrospective cloud data management tool and found that 83% of their participants preferred deleting at least one file they saw from their accounts. Will et al. [7] developed a tool to help users manage similar files with the intuition that they should be managed similarly. They found that users were more likely to accept recommendations to delete the files than to move them. While these tools are promising, their performance suggest that automatic cloud file management (including deletion) is still at an early stages. Currently, users are still responsible for classifying and managing their cloud files.

2.3 User Preferences

One of the many motivations for cloud deletion is privacy [25, 44]. Prior research has extensively focused on understanding users' privacy preferences regarding different personal data and technology. Efforts have focused on online social networks [12, 30, 41, 42], smartphone privacy [20, 21], advertising [38, 77], location [1, 34, 39] and data sharing preferences [36, 46, 62]. Results from these studies have led to various privacy mechanisms and improved user interfaces for users to control who can see their posts and avoid regrets in social media, for instance [12, 41, 42]. Also, in other contexts, such as online advertising, users may like to state their preferences not to be tracked or select the kind of adverts they want to see, with Melicher et al. [38] reporting that users found this beneficial and offering them a sense of control. To address users' privacy concerns in smartphones, users can give various applications different access permissions [28, 29]. Regarding deletion, design and longitudinal data management studies point to a growing need for tools and mechanisms that can support users in decisions to keep and discard their data [13, 74]. They emphasize how individuals' practices and preferences differ in their daily lives. For example, users often want to delete similar files even if they were not stored in the same folder [7]. Moreover, they frequently delete files they consider not sensitive and useless. However, they always want to recover work and school-related files if they were deleted by mistake [26]. Studies around deletion preferences concerning cloud deletion are minimal though these two studies evidence some preferences for deletion. In other areas, studies around user preferences have enhanced users' control over their personal data and privacy; insights on deletion preferences are missing, particularly when it comes to deletion from the cloud.

2.4 Privacy Policies and Terms of Service

Today's online services use notice and choice (i.e., consent) as the paradigm for giving consent online [63]. User Agreements, Terms of Service agreements, and Privacy policies are expected to contain service provider's data practices (i.e., data collection and use) and presumably providing users with sufficient knowledge to help them make informed decisions on whether they should disclose their information or stop using the service [24]. With regards to deletion, they are expected to explain how data is deleted, service provider's deletion practices including retention policy and recovery terms. Nonetheless, the vast amount of literature has reported on the usability problems of current privacy policies. Majority of these are long, unreadable and contain irrelevant information [24, 37, 50]. Some are not aligned with user privacy concerns [11]. Others fail to inform users, leaving them helpless [58]. To actively help users manage their privacy, the authors of [58] suggest that control mechanisms must be relevant, actionable and understandable. They also identify four main dimensions to consider when designing to provide notice: timing, when should a notice be presented; channel, how should the notice be delivered; modality, how the information should be conveyed; and control, how choice options are integrated into the notice. In the context of cloud deletion, improving these notices will help users understand the provider's deletion practices and inform their choices with regards to deletion. In this work, we explore what information about

deletion in the cloud is essential to users, when should it be presented to users (timing), and the channel that should be used.

3 PREVIOUS STUDY - DELETION PRACTICES

In our previous study on cloud deletion [54], we explored several key questions fundamental to usable privacy and security with regards to cloud deletion:

- Motivating factors behind cloud users' need to delete.
- The challenges they face or factors that underpin their failure to delete.
- The strategies that they employ to circumvent the challenges.
- The deletion experience users want.

Using semi-structured interviews (n=26) and grounded theory analysis, we identified four key drivers that motivate users to delete, three key themes that explain deletion failures and various coping strategies that users adapt to deal with failure to delete.

We found that users' motivations to delete are privacy-, expertise-, policy-, and storage-driven.

Privacy driven - This group of participants deleted because they did not trust the cloud provider, wanted to forget certain files and avoid future conflicts.
Expertise driven - Participants' motives to delete were based on users' understanding of the cloud deletion and ability to delete. When participants were confident that they could delete, they were more likely to delete than those who were not confident.
Policy driven - Some participants deleted due to extrinsic policies of their workplaces or perceived value of the files.
Storage driven - Participants were mostly driven to delete to free up storage and avoid clutter.

With regards to deletion failures, our analysis revealed that some deletion failures were due to limited access to information about deletion, interface issues and users' cloud deletion mental models. Participants revealed that very often information about deletion is not enough or not easily accessible as other information such as how to upgrade storage. Others stated that sometimes their cloud application crashes or they struggle to find some deletion features, for example, deleting items from cloud recycle/trash cans. We also found that some deletion failures stem from the misunderstandings and the concepts that some users have about cloud deletion. For instance, some cloud users refrain from deleting from shared folders because they are not sure whether the file, they want to delete will be removed from the visibility of other shared folder members. Others, for example, those who are privacy-driven, may end up having files in their 'deleted items' folder without their knowledge because they may think deletion is permanent.

The study also revealed that cloud users adopt various strategies to address their challenges to delete. For example, some participants prefer to delete from certain devices (e.g., sync folders in their personal computers) or platforms from which they are confident they will be able to delete, while others choose to filter what files get stored in their cloud accounts so that they may not require the need to delete. We also learnt that the choice of the strategy is not consistent, it is always changing, depending on the nature of the challenge, their expertise level, the device they are using and the reason why they want to delete. For instance, participants who delete to free storage favored the strategies that created space (e.g., deleting other files) while those whose motivation to delete is privacy related will adopt a strategy that removed the file from the storage.

Our work revealed that some cloud users desire to have transparency regarding deletion; they wanted information about how service providers delete or handle deleted data to be made freely available. Others expressed the desire to have a complete deletion (i.e., permanent deletion of data and meta-data) as they stated that they were not confident about how data is deleted. Moreover,

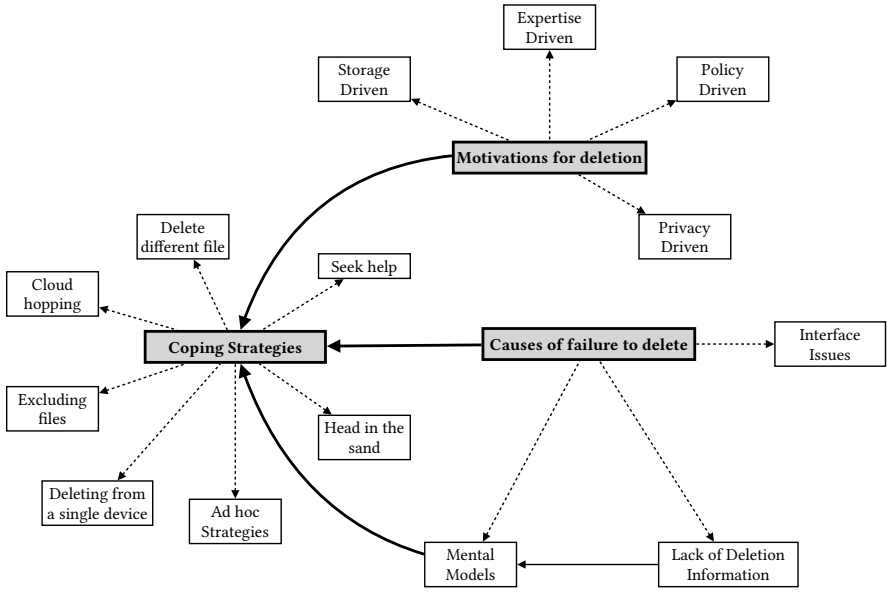


Fig. 1. Adapted. Key Findings from our previous study [54].

some suggested they lacked control over the deletion process and would prefer having control how deletion should be executed, for example, how long deleted files should remain in the 'deleted items' folder. With regards to getting help with deletion, others suggested having dedicated services for answering queries about the deletion. Figure 1 gives a summary of these findings.

Overall our previous work helped us to understand users' cloud deletion practices and how users cope with deletion failures. However, it did not give insights on users' deletion preferences or inclinations concerning the information about the deletion. Participants expressed their desire to have complete deletion and control over deletion. Nonetheless, our work did not explore these views in-depth, for example, we did not ask our participants to explain the scenarios in which they would want complete deletion or whether they would want complete deletion for all their cloud data. With regards to information about deletion, participants expressed the desire to have transparency about the deletion and more information about cloud deletion. However, we did not examine the kind of information about cloud deletion they wanted or where they would expect to find it. In this article, we aim to address these limitations and provide insights on users' cloud deletion preferences, highlighting situations in which users may want complete deletion over other types or deletion. We also provide an understanding regarding information about deletion, what information users consider essential and where and when they would want to have access to it.

4 METHODOLOGY

To investigate users' deletion preferences and information requirements, we used three participatory action research (PAR) [5] tasks. PAR involves participation and action from a group of people who are affected by the same problem and act together to tackle it. As a collaborative research methodology, it offers researchers the opportunity to co-develop or investigate with users. It stresses users' lived experiences, social changes, their construction knowledge which can be useful for solving their everyday challenges [4]. Thus, discovering and developing solutions that are viable and useful to users. Moreover, we chose participatory action research because it is a well-established

method in HCI to explore complex issues with users [5, 48]. Cloud deletion is not an easy topic for all users [44, 54], gathering a group of users to explore it is likely to yield better results than discussing it with individuals on a one-to-one basis. We chose the sorting (i.e., grouping) method for each PAR because sorting is a natural cognitive process routinely used in everyday life on which many evaluations and decision-making processes rely [31].

The rationale behind each PAR.

PAR 1 - Data sorting. We chose data sorting activity because the findings of our previous work [54] suggested that sometimes participants considered the file type when choosing which file to delete, for example, when deleting to free space, participants reported that they usually choose the file they did not consider to be important. In this current study, we wanted to know whether participants' perception of the file (i.e., importance and sensitivity) also influenced their deletion choice. Letting participants group these data types beforehand seemed reasonable before asking them how they would delete such data.

PAR 2 - Deletion preferences. Participants deleted cloud data for various reasons [54], and we hypothesized that this might be linked to deletion types, so asking them to sort data this way seemed an intuitive way to understand how participants want their data to be deleted.

PAR 3 - Information Requirements. We asked participants to categorize information about deletion in terms of importance when they want to see it and the communication channel because prior studies, for example, Murillo et al. suggest users would perceive deletion more accurately if they understood the deletion concepts we used [44]. Thus, we created these concepts and examined whether participants valued some more than others and when and where they would want such information to be made available. Allowing participants to group concepts using our predefined helped us to observe these differences easily.

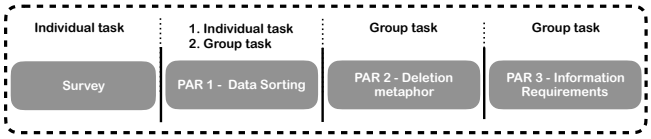


Fig. 2. The survey was taken prior to attending the study sessions. Participants first classified data types, then completed the task based exercises.

Our study involved completing a pre-study survey, and then the three (3) activities/PARs as shown in Fig. 2. These PAR are described in detail in Sections 5, 6, and 7. In this section, we focus on giving details on the overall method, the study procedure, recruitment, data collection, and data analysis.

Study Procedure

Before users could attend a session for task-based exercises, participants were asked to complete an online pre-survey that, in addition to obtaining demographics details, assessed their perception and cloud deletion practices.

Selected participants (criteria explained in the next section) were invited to our labs to complete the rest of the study. The first activity involved sorting data types individually, then as a group. The second activity concerned deletion preferences and the last activity was about deletion information. PAR 3 and 3 activities were group tasks. The lead researcher moderated all the sessions. At the beginning of each session, they explained the general aim of the study, what was required of the

participants in each task, answered participants' queries, started discussions around each PAR (including probing participants to explain their reasoning behind their sorting preferences or ask the other group members to give their opinions), managed time, and ensured that the objectives of the study were addressed. The facilitator also ensured all the study materials were available and that all categorizations were photographed before, during, and after activities.

Code	Group	Gender	Age	Employment	Accounts	Cloud Services
P1	A	Female	21 - 25	Student	2 - 3	Dropbox, iCloud, Google drive
P4	A	Male	31 - 35	full time	4 - 5	Dropbox, Box, Google drive, OneDrive
P11	A	Female	31 - 35	full time	2 - 3	Google drive
P17	A	Male	18 - 20	Student	2 - 3	Dropbox, Google drive, OneDrive
P18	A	Male	31 - 35	full time	6 +	Dropbox, iCloud, Google drive, OneDrive
P2	B	Male	31 - 35	PhD Student	2 - 3	Google Drive Box
P5	B	Female	31 - 35	full time	2 - 3	Dropbox, Google drive
P7	B	Female	18 - 20	Student	1	Google drive
P13	B	Male	26 - 30	full time	2 - 3	iCloud, Google drive
P19	B	Female	41 - 45	full time	4 - 5	iCloud, Google drive, Dropbox
P14	C	Male	26 - 30	full time	2 - 3	iCloud, Google drive, OneDrive
P3	C	Male	26 - 30	PhD Student	2 - 3	iCloud, Google drive
P8	C	Male	21 - 25	full time	2 - 3	Google drive, OneDrive
P9	C	Female	18 - 20	Student	1	Google drive
P20	C	Female	31 - 35	full time	2 - 3	Google drive, OneDrive
P10	D	Male	26 - 30	part time	2 - 3	Dropbox, iCloud, Google drive
P6	D	Female	26 - 30	Unemployed	2 - 3	Dropbox, Google drive
P12	D	Female	26 - 30	PhD Student	1	Google drive
P15	D	Female	31 - 35	Student	4 - 5	Google drive, OneDrive
P16	D	Male	36 - 40	full time	2 - 3	Google drive, Box, Amazon Cloud Drive

Table 1. Summary: Study Demographics.

Recruitment, Ethics and Data collection

After obtaining an ethics clearance, we recruited participants through social media, word of mouth, and adverts around the university and the city center. Interested respondents were encouraged to complete a screening form. The purpose of this questionnaire was to identify active cloud users who were 18 or older meeting three or more of the following:

- having deleted from the cloud through more than one device or interface (so that they could provide their experiences based on more than one interface),
- having more than one cloud account (so that they could share their experiences based on more than one provider),
- sharing some folders (to be able to gather their deletion preferences from shared files),
- experienced some challenges when deleting (to learn about how they would solve their challenges),
- interested in cloud deletion, and
- being able to attend the participatory study.

Sixty-five (65) people (40% identified as male) responded to our adverts and completed the screening questionnaire. Sixty stated that they have deleted in the cloud, 17 of which experienced some challenges while deleting, 76.9% sharing folders, 46.2 % had more than one account, and only three people stated they could not attend to do the study.

In the end, 20 (50% females) participants were invited to take part in the study. We divided them into four equal groups. Each participant was given a day and time when the study will take place and was asked to confirm their availability. While the initial group allocation was random, for diversity purposes, we ensured that each group contained at least three non-student participants, not more than three people of the same gender, and diverse age groups. We considered this sample sufficient for the study and the complexity of the topic. Deletion as a topic is not as popular as privacy. Consequently, a smaller number of participants allowed us to probe and ask follow-up questions

during the session. Complex topics are easy to explore using the participatory method [5, 48]. Moreover, Oates and Alevizou [45] suggest that conducting studies with groups of five to eight is sufficient to generate discussions and provide valuable insights. Table 1 lists the demographics of all the participants. We obtained consent to record audio and take pictures (i.e., pictures of the props without participants' faces) during the sessions. Each participant received compensation worth $7.00 for their time.

In total, we collected 224min worth of audio from all the group sessions and took a total of 94 images (188MB). Sessions took an average of 63min excluding breaks.

Pilot studies

After obtaining ethics clearance, we ran three pilot study sessions with three different groups of participants (i.e., Four participants per group). These sessions were used to understand how long each task would take to complete, assess whether the research protocol is realistic and workable and whether the study props were enough to complete the studies. Pre-tests were run as if they were the final study; participants completed the ethics process, and the researcher conducted the study as if the results were going to be used for the final report. However, during the pilot study, the researcher noted how the study was going, noting challenges and how participants engaged with the props. At the end of each session, the participants were asked how they found the study, what they struggled with and what they would remove from the study.

As a result of the pilot study, we removed some data types because participants suggested they would not usually have such data in the cloud and felt similar to other data types we already had. In the end, we removed 12 data types. Some data types were removed to reduce the time to complete PAR1 and PAR2 tasks. Initially, we did not provide participants with deletion types, but we discovered that without deletion types, participants only decided on two options, to delete and not to delete. Consequently, this did not provide any variation on how they would delete their data. However, after introducing deletion types in the second session, we noticed differences in how they wanted data to be deleted. This also led to rich discussions about deletion types. Moreover, we initially had one break between PAR1 and PAR2. However, after the first pre-test, we allowed participants to take breaks after each session if they wanted to. Data collected during the pilot studies was not used in the final results.

Analysis

After transcribing all the audio and photos, we performed qualitative analysis—a thematic theory approach. To generate a codebook, a lead researcher independently coded all the data from the first group. The first stage of analysis involved the identification of various data classification from photos from PAR1 using the open coding technique [9, 16]. This first focused on the individual sorting then the group sorting. The lead researcher recorded the characteristics of each group and classified similar groups since they had overlapping features. For example, groups that participants labeled personal, private, important were grouped together because they overlapped. Then, the PAR 1 section of the transcripts was coded, mainly to understand the groups' reasons, which helped inform the decision behind grouping certain groups together. This process led to the first codebook. After identifying the high-level groups, the second researcher analyzed the same photos and transcripts to confirm whether they agreed with the first coder. The two researchers then discussed the codebook, especially the high-level groups identified from the data.

The lead researcher then coded the second and third PARs, following the same process with the second coder confirming and discussing the codebook. The only difference in analyzing PAR2 and PAR3 was that we used closed coding [9, 16]. Initial codes for PAR2 and PAR3 were the categories we used in the activities. For example, PAR2 categories were the four deletion types we used in

the deletion activity. After compiling a codebook, two researchers independently coded the rest of the scripts from the remaining groups using a single codebook. The Cohen's Kappa coefficient agreement was found to be 0.72, showing a high degree of agreement between the two researchers. After the independent step, the two researchers collaborated to refine the disputed codes which resulted in disagreements. We found that this was due to researchers interpreting some codes differently, so the codebook was refined to clarify them. After the initial coding was complete, further analysis revealed themes and categories about users' deletion preferences.

After coding the first two sessions, we did not see any variation within the groups that participants generated in PAR1. In some cases, the contents of the grouping would have more data types or different data types, but the definition of the groups by the participants remained the same. However, we continued with the rest of the group session to confirm whether we had reached saturation point.

5 PAR1 – DATA SORTING

Data types			
Medical report/ information	Music videos	Old birthday video	Children photos (Family)
Rifle licence	Honeymoon photos	Genetic information	Facebook downloaded data
Immigration documents	WhatsApp backup	Family photos	Research data
Personal information	Meme videos	Job application letter	4MB video clip
Biometric data	Meme images	Legal documents	Business contracts
Passport copy	3GB wildlife video	E-books	Friends photos
Old bank statements			Pet care information

Table 2. List of data types used in the data sorting activity.

Activity Design

To explore the deletion preferences, we developed a sorting task which required participants to sort various types of data (e.g., meme videos, passport copies) individually and as a group. We adopted a free-sorting technique and asked participants to categorize given data types according to how they perceive them so that similar data types are gathered together. The categories were not predetermined so participants could create as many groups as they find fit. There were 26 data types in total, selected from various research work and online sources [15, 18, 27, 40, 64]. We listed the suggested data types from all our sources then generated more data types like those indicated in the listed sources. Based on this list, we then discussed which data types to include or exclude from our study. These were data types that users were familiar with, particularly around the context of cloud usage. We wanted a list of data types that covered most aspects of an individual's life and may pose privacy reasons for participants if not handled well. Table 2 shows the list of all data used in our study. The purpose of this task was to visualize and identify participants' intuitive categories and investigate whether these categories would have any influence on their deletion preferences. On average, this activity took each participant 5 min and around 13 min when part of a group.

PAR1 – Findings

Following open coding, we identified three themes which show how participants categorised cloud data.

We found that people classify data differently—they create various groups, sometimes overlapping. Participants frequently categorised data according to sensitivity, utility, content, and use. Common

groups included: sensitive, personal, important, less important and less sensitive, miscellaneous, and sensitive and important. Some participants highlighted that *important data* is the data that they consider useful (utility), hard to get, and do not want to lose while *sensitive or personal data* is the data they consider private and can be used by others to identify them. For consistency purposes, we grouped similar groups and recorded them as a single group. As a result, we ended up with three groups: important and sensitive, less important and less sensitive, and important but less sensitive.

Important and sensitive: This group contained data that participants described as private, personal and they did not want to share it with unknown or unauthorised people.

Less important and less sensitive: This consisted of data that participants considered less useful, easy to reproduce and less private.

Important but less sensitive: This group contained data that participants considered useful and hard to reproduce but were happy about other people knowing about it.

	Individual sorting			Group Sorting		
Data Types/Files	Unimportant & less sensitive	Important but less sensitive	Important & sensitive	Unimportant & less sensitive	Important but less sensitive	Important & Sensitive
Medical report	0	2	18	0	0	4
Rifle licence	5	7	8	0	1	3
Immigration documents	1	2	17	0	1	3
Personal infromation	1	1	18	0	0	4
Biometric data	1	0	19	0	0	4
Passport copy	3	3	14	0	2	2
Old bank statements	4	0	16	0	0	4
Business contracts	0	7	13	0	1	3
Music videos	20	0	0	4	0	0
Honeymoon photos	3	7	10	0	0	4
WhatsApp backup	10	0	10	2	0	2
Meme videos	19	0	1	4	0	0
Meme images	19	0	1	4	0	0
3GB Wildlife video	19	0	1	4	0	0
Children photos (Family)	1	7	12	0	1	3
Friends photos	4	7	9	0	4	0
Facebook downloaded data	14	1	5	2	0	2
Old birthday video	6	9	5	0	2	2
Genetic information	0	2	18	0	0	4
Family photos	3	7	10	0	0	4
Job application letter	6	2	12	0	1	3
Legal documents	0	4	16	0	0	4
E-books	16	4	0	4	0	0
Pet care information	8	9	3	4	0	0
4MB video clip	17	2	1	4	0	0
Research data	2	13	5	0	3	1

Table 3. A heat map summarising the results of the individual and group sorting tasks. Each cell reports how individuals and groups perceived each data type. Color code: Red represents more individuals/groups while green represents less numbers.

We observed some differences in the data sorting activity between individuals and groups. The results of our sorting activity are shown in Table 3. We discuss these differences below.

Individual sorting. Individually, participants generally classified data into four to five groups. These most common groups were personal, sensitive and important, not important, miscellaneous and work. Other groups were named entertainment, family and less sensitive and less important. Similar groups (according to their description or properties) were joined together, and we found that participants had fewer data types categorised as important but less sensitive. The important

and sensitive group included data that was about them (e.g., copy of a passport) or data related to their families (e.g., family photos). Data types perceived to be less important and less sensitive included music videos and Facebook downloaded data [1].

Group sorting. In group settings, participants categorized data into three or four groups. The groups were fewer in number and contained more data types. After grouping these collections into our three broad groups, the less important and less sensitive group had fewer items while the important and sensitive group had the most items. Some data types were split between the two groups. For example, two groups classified WhatsApp data as unimportant and less sensitive, while the other two categorized it as Important and Sensitive. When referring to photos, P5 noted:

> *"This is about family, private things. Like children's photos."* P5 Group B

P8 stated that meme videos and images were more suitable under unimportant data:

> *"I would have these under entertainment, something like that. This is not important..."* P8 Group C

Individual vs Group context. Comparing the sorting between the individuals and groups, the number of data types classified as Important and Sensitive increased in the group sorting than it was during individual sorting. Furthermore, most data types classified as important but less sensitive during individual sorting ended up in important and sensitive in group settings. We posit that the differences in the results may be because, in group settings, participants discussed different risks concerning each type of data which may have influenced their choice. For instance, participants discussed different threats that could affect disputed data types or how such data can merely be misused (e.g., data being used to impersonate the owner).

> *"People can impersonate you, something like that. I can use [business contracts], I can claim to be you..."* P16 Group D

These discussions may have impacted individual users' perception of some data because all the data types which were discussed in this manner were generally moved to the sensitive and important group. For instance, during individual data sorting task, some group A participants classified WhatsApp backup data as not important and less sensitive but during group sorting where the risk was discussed participants agreed that such data should be classified as sensitive and important because WhatsApp data may contain personal and private information.

> *"Do you know they backup everything... WhatsApp backup would contain a lot things I assume... messages, pictures, am not sure... maybe contacts. You don't want people to know that... I would say private."* P18 Group A

We discuss the implications of this finding further in the discussion section.

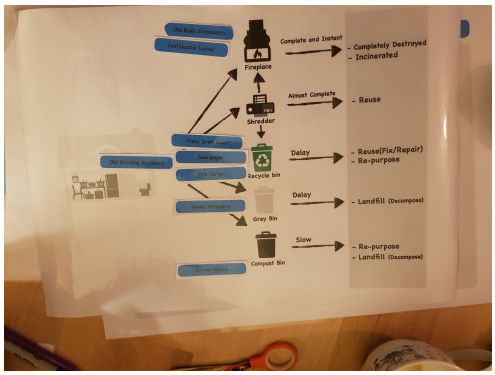
6 PAR2 – DELETION METAPHOR

Activity design

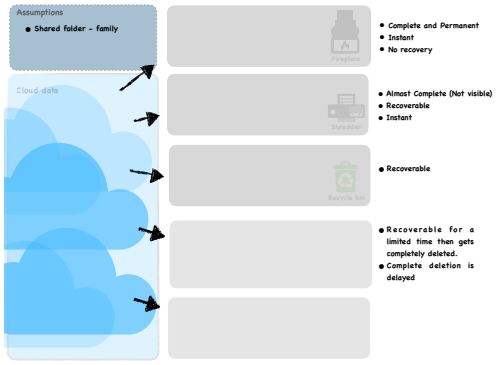
To help elicit cloud deletion preferences, we developed a garbage collection metaphor. Using metaphors is a well-known HCI technique to help users think about digital objects as they would think about real-world objects so as to increase their familiarity with them [6, 10]. We also used our metaphor to minimize participants introducing their own metaphors to the study which may lead to inconclusive results. We learnt from our previous study [54] that users can possess incomplete or incorrect mental models of cloud deletion. This PAR was divided in two parts.

The first part of this activity focused on household waste management. Using an A3 paper, we presented a diagram of a house and five empty boxes depicting different ways of managing

[1] After Cambridge Analytica scandal [32] Facebook allowed their users to download all the data they have shared on the network. https://www.facebook.com/help/212802592074644



(a) Participants completing the metaphor task



(b) Shared folder scenario

Fig. 3. Tasks: (a) Before completing the deletion preferences task, a metaphor task was used to help users appreciate cloud deletion — users exploring different ways of managing waste, and (b) During the deletion preferences task, users classified how they would delete data from a shared folder.

household waste. The five boxes had a picture of a fireplace, shredder, green bin, grey bin, and compost bin. Each box had a description of what the box represents, and a list of properties associated with the represented method of waste management. We then created eight (8) labels representing household waste. These included: old bank statement, confidential letter, Fizzy or Soda can, newspaper, milk carton, rotten apples, candy wrappers, and old working computer keyboard. The task was for participants to place each type of waste on the appropriate box (bin). We chose garbage metaphor because it somehow depicted deletion, i.e., disposal of unwanted material with which most participants are familiar. Using this metaphor also helped us further build rapport. Fig. 3a shows participants completing the metaphor task.

In the second part, we tailored our metaphor and used its main concepts to design a deletion preference activity which required participants to sort out how they would delete cloud data. Rather than boxes depicting different bins, the boxes now depicted different four(4) types of deletion and their properties.

We considered the four main types of deletion or deletion properties according to previous literature [53, 57] and named them for easy understanding.

> **Complete deletion** — deletion that removes all copies of data permanently from the cloud,
> **Soft deletion or partial deletion** — nominal deletion where some parts of the deleted data can still be recovered,
> **Camouflage deletion** — deletion that removes data from the users' locality but not removed and can always be recovered, and
> **Trash can deletion** — deletion that allows recovery for some time before data is completely erased.
> **Not delete** — this where an individual does not want to delete.

For this task, we had twenty-eight data labels (from the data sorting task) for participants to use for the deletion task. Each group was required to categorize how they would want such data to be deleted under two different conditions; (1) when deleting from a personal account, and (2) deleting from shared folders.

PAR 2 – Findings

After open coding (explained in Section 4), we proceeded to identify themes and relationships between our data, especially those that explained users' deletion preferences. We identified four themes that explained their desires and needs regarding deletion in the cloud. We further compared these preferences to the results from the data sorting task and found that deletion preferences (or needs) are not always aligned with how participants perceive data. Participants' deletion preferences are multi-dimensional, different and not consistent within individual and social contexts.

Data	Complete Deletion	Soft Deletion	Camouflage Deletion	Trashcan Deletion	Not to delete
Medical report/information	A, B, C		D		
Riffle licence	A, B, C				D
Immigration Documents	A, B	C			D
Personal information	A, B	C			D
Biometric data	A, B, D	C			
Passport copy	A, B, C				D
Old bank statements	A, B, D		C		
Business contracts	A, B		C	D	
Music videos	C, D	A	B		
Honeymoon photos	A	B			C, D
WhatsApp backup	A			B, C	D
Meme videos	C, D	A	B		
Meme images	C, D	A	B		
3GB Wildlife video	B, C, D	A			
Children photos	A	B	D		C
OS installation file	C, D	A, B			
Friends photos	A	B	D		C
Facebook downloaded data	A, C, D	B			
Old birthday video		A, B			C, D
Genetic information	A, B, C				D
Family photos		B	A, D		C
Application letter	B	A		C, D	
Legal documents	A, B	C	D		
E-book (pdfs)	C	A	B	D	
Pet care information	C, D	A	B		
4 MB video clip	C, D	A	B		
Research information		A	B, C	D	

Table 4. Summary: Data deletion preferences for deleting from an individual account by each group.

Individual contexts.

Comparing the deletion preferences to the data sorting activity, we found that individuals' preferences are often not aligned with how they classify or perceive data. Individuals may desire a deletion type that may not necessarily be a conventional way to delete certain data types. For instance, data classified as personal or private may not necessarily be disposed of by a deletion type that destroys it completely and instantly to prevent it from falling into the wrong hands or being abused. Deletion preferences are not fixed to any particular group of data. Participants explained that data types could not be tied to a specific type of deletion because deletion needs (requirements) are different and constantly changing. For instance, Group C and D preferred Complete and Permanent deletion when deleting data perceived as unimportant and less sensitive, arguing that such data is unnecessary and might not be needed in the future. However, in some cases, Group C did not opt for complete and permanent deletion for the same data. Table 4 shows the deletion preferences of deleting various data types from an individual cloud account for all the groups.

Within individual contexts, participants **deletion preferences vary and depend on various factors.** During the preference tasks, participants expressed distinct deletion preferences. We found that participants' deletion needs concerning a particular group of data or individual data types

vary. We observed that these preferences were different even when the group had discussed and agreed on the risks or how to classify a particular data type, each group would argue for a different type of deletion. For instance, during the sorting task all the groups classified photos belonging to a friend as important but less sensitive, however, during the deletion preferences task, each group decided to have this type of data deleted differently.

> *"Friend's photo can be very important. I say lets keep it unless we have another copy. "* P9 Group C

> *"Think about this, why would I have my mate's private photo. I know it's not private, but I am sure it would be something funny. I would delete it just so I can always get it back..."* P10 Group D

During coding, we also found suggestions that **deletion preferences within individual contexts change over time.** Changes in participants' lives or how they use the cloud may impact their deletion preferences. We found evidence that when cloud usage, file utility, social life, storage needs, and privacy needs changes, participants' preferences may also change.

Cloud usage. We found that when cloud users change their use of the cloud, it may also change how they want data to be deleted. Some cloud accounts start as personal but may end up being used for other purposes, such as storing work-related data. Participants then mentioned that this, in many cases, changes how they would want the cloud provider to delete their data.

> *"I used to work as a photographer. I never deleted anything. Now, I just delete."* P4 Group A

File utility. We also found that when the value or utility of the file changes, its deletion needs may also change. For instance, at the time of the study, a PhD student from Group C (P3) explained that they could not delete their research data. However, after ten years, their university policy requires complete and permanent deletion of such data if it is no longer in use. They also highlighted that after a certain period has elapsed, the data may not have the same value as before, prompting a different type of deletion. For instance, honeymoon photos may have a different value after divorce.

> *"I would never completely delete [research data]. The university can delete it. We are not allowed to delete data before 10 years unless its special data."* P3 Group C

Storage size and needs. Some users' preferences change when they need more storage, especially for those that use services that count deleted items as part of user's storage quota. Participants stated that when deleting from such services, one may prefer complete deletion than soft deletion which may leave such data as part of the account.

> *"it depends on the service provider, if they treat recycle bin as separate space then I can choose trash can. In not, I am completely deleting it."* P18 Group A

Privacy needs. Other participants stressed that as their privacy needs change or when data becomes irrelevant, their deletion preferences for such data also changes. For instance, when private or personal information change, they may no longer be required to entirely or permanently delete files containing such information.

> *"We need to know why we are deleting first. We will always have different results depending on context. If today I have cancer, I want to keep that to myself, permanently delete. Tomorrow, I am well and I don't care who knows about [it]. "* P2 Group B

We also found that **deletion preferences are complex and multi-dimensional** in nature; they depend on different factors including the reason for deletion, file sensitivity, file utility and size.

Reasons for deletion. Participants noted that the reason why they want to delete a file plays a vital role in the type of deletion they would prefer. For instance, when deleting to prevent others from seeing or having access to a file, (e.g., health-related data) participants desired a deletion process that is instant, complete and allows no recovery. However, one group argued that medical reports, unlike genetic information, might require recovery as past medical conditions may not be reproducible while genetic information can be obtained all the time.

> *"How much will I pay to get this info; I will keep it [genetic info] if it is expensive, maybe... again it never changes. If its cheap, I will permanently delete it."* P12 Group D

When the reason for deletion is to tidy the account (e.g., deleting memes and music videos which they do not consider important or sensitive), they prefer a quicker method (i.e., soft deletion), citing that completeness in such a case is not essential. Despite this reason, some participants stated that non-essential data might require complete deletion since that data may not be necessary for the future.

> *"This is permanent deletion, why would you keep this. Unless they are the videos you made. I will never need this."* P7 Group B

Concerning space, participants preferred soft deletion when the deleted files do not count towards the storage quota but complete deletion when they count.

> *"I didnt know deleted folder can count. Maybe thats why I used to buy storage all the time."* P19 Group B
>
> *"There is no point to do trash can, if it is not freeing your storage."* P18 Group A

File sensitivity. We also found that file sensitivity affects the type of deletion participants may desire to choose. When participants considered a file to be sensitive or private, they highlighted the need for complete deletion, but when it is not, they preferred other types of deletion. Group A emphasized complete deletion for most data because they could not always trust that the cloud provider is not malicious.

> *"Its immigration documents, where do I put them. If someone can recover them, you are in trouble. This is the cloud; I prefer to immediately remove it completely."* P13 Group B
>
> *"If applying for visa extension, am concerned that this [data] can be stolen but I need this readily available. But after I get my visa, I may want to delete [it] quickly"* P3 Group C
>
> *"I remember stories about DropBox. Deleted files came back... it shows they dont delete. permanent deletion for me, please."* P16 Group D

File utility. Participants also discussed how the importance and purpose of a file might influence how it is deleted. We found that participants generally do not favor deleting files they consider useful, especially those they cannot easily reproduce. Nevertheless, if they do so, they would prefer soft deletion so they can always recover such files. During the deletion activity, we observed that groups preferred not deleting data that held fond memories, such as honeymoon photos and old-day birthday videos.

> *"Birthday videos are great. My mom would be mad at us for deleting them. I would only delete them if I have a copy somewhere. "* P7 Group B

File size. Some deletion preferences may be influenced by the size of the file that is being deleted. However, not on its own, but with other factors discussed earlier in this section. Smaller files merited soft deletion over complete deletion, particularly when the file does not contain sensitive content or the reason for deleting is not to create free space. Groups B, C, D preferred complete deletion for a 3GB random wildlife video.

> *"Its big, random. Why keep it? I would completely delete it."* P8 Group C

Deletion interface. Participants highlighted that sometimes their choice of deletion might be influenced by the device they use to access the cloud. They argued that some interfaces provide better usability, influencing their deletion choice. For instance, Group B and C argued that they would mostly want complete deletion if they are using a sync folder in the computer or a computer browser than when using a mobile application. Some participants reasoned that it is easier and effortless to identify mistakes and recover files using web browsers or sync folders than mobile applications.

> *"Have you ever tried to delete from your phone. iPhone used to be so small. I would permanently delete useless things which don't matter. If its works [related], maybe not."* P20 Group C

Data	Complete Deletion	Soft Deletion	Camouflage Deletion	Trashcan Deletion	Not to delete
Medical report/information		A, B	C,D		
Riffle licence			A, B, C		D
Immigration Documents			B		A, C, D
Personal information			A, B		C, D
Biometric data	A, B, D	C			
Passport copy			B		A, C, D
Old bank statements	A, B, C, D				
Business contracts			A, C	B, D	
Music videos	C, D	A, B			
Honeymoon photos			B		A, C, D
WhatsApp backup	A, B, C, D				
Meme videos	C, D	A, B			
Meme images	C, D	A, B			
3GB Wildlife video	C, D		A, B		
Children photos					A, B, C, D
OS installation file	A, B, C, D				
Friends photos					A, B, C, D
Facebook downloaded data	A, B, C, D				
Old birthday video					A, B, C, D
Genetic information	A, B, D	C			
Family photos					A, B, C, D
Application letter	A, B	C, D			
Legal documents	A, B	C, D			
E-book (pdfs)	A, C, D				
Pet care information			A, B, C, D		
4 MB video clip	C, D	A, B			
Research information	A, B, C, D				

Table 5. Summary: Data deletion preferences for deleting from a family shared folder.

Social context.

Regarding social context or deleting from shared folders, we found that despite data classification groups, participants preferred soft deletion or not to delete files than complete deletion. We also found that, unlike in individual contexts, where preferences vary a lot, in a social context, preferences differ but are highly dependent on the situation. Participants' choices mostly fluctuated between not deleting, camouflage deletion (i.e., always recoverable) and soft deletion. Table 5 shows the deletion preferences of deleting various data types from a family shared folder for all the groups.

Participants' **deletion preferences over shared folders also change over time**. However, we found that their deletion preferences are influenced by the type of relationship between shared folder members, social life, the total number of users involved, authorship status and perceived trust.

Type of relationship. Participants highlighted that their choice of deletion in a shared folder might be influenced by how the shared folder members are associated. They prefer deletion that allows

recovery when it concerns close relationships such as family members. Participants explained that such groups usually include non-tech savvy users who may accidentally delete useful files thus the need for a non-destructible method of deletion. One participant stated that without the option to recover deleted data in the family shared folders, some family members might delete photos of themselves (e.g., when there were children) which they deem embarrassing and such memories may be lost forever.

> *"My brother would permanently delete everything. He doesn't like his old pictures. With family, I would say I want to recover every time."* P7 Group B

Regarding shared folders among work colleagues, participants preferred not to delete because some members of the group may still be using the concerned data. However, when all group members have authorized the deletion, they preferred complete deletion with no recovery.

> *"It happened at work recently. They deleted everything and our manager was not happy. We were lucky we could still recover them."* P14 Group C

Social life. Participants also highlighted that when their offline life or relationships change, they may change how they would delete certain files. When their offline social links become weaker, they may desire more destructible type of deletion than when they are stronger. Some participants stated that when they exit a shared folder, they may want to completely delete some of the data they contributed to that shared folder.

> *"My group went to the lecturer and said I didn't do anything. I changed groups and went into our shared folder and deleted everything. I would want permanent deletion for that. "* P17 Group A

Number of members. We also found that deletion preferences are also affected by the number of people sharing a folder. When fewer people (e.g., 5 or less) share a folder, they sometimes prefer complete deletion; however, when the number is more significant, they prefer deletion that allows recovery. Participants reported that it is usually easier to get permission to delete data from a smaller number of participants than a larger one. They explained that they prefer deletion types that allow recovery for larger groups because they can quickly recover files if other group members complain.

> *"If you delete something, maybe you should have everyone in the shared folder agree. Imagine twelve people. Two is easy. If there are two many people. You just delete to hide it from everyone. if they realize, you bring it back."* P18 Group A

> *"Communication is important. If everyone agrees then why not [delete permanently]. It works well with few people."* P3 Group C

Creators and authorship. We also discovered that participants consider who created the shared folder or uploaded/created files before deciding how they can be deleted. All the groups agreed that they would easily consider complete deletion for the file or data they uploaded to a shared folder. However, regarding data they did not create or upload, they would prefer deletion types that allow recovery. This suggests that when participants are confident about the data they are deleting, they prefer complete deletion. However, when not certain, prefer to have a recovery option.

> *"If I uploaded it. I can easily delete it. If its permanent, I can always upload it again."* P17 Group A

Trust. Our last theme discusses trust, that a member of a shared folder has the technical skills to delete without mistakes (confidence in the ability of others) and trust that people have on others (interpersonal trust). We found that for shared folders where some members are not trusted to have enough technical skills, deletion that allows recovery was the most preferred type of

deletion. However, when members are trusted, complete deletion is allowed particularly for fixing mistakes such as deleting data uploaded by mistake. Regarding interpersonal trust, our coding suggests that when the interpersonal trust is high (e.g., a couple sharing a folder), participants are comfortable with complete deletion. We also found there are instances where both forms of trust are considered; for example, tech-savvy partners are more often trusted with a complete deletion that non-tech-savvy ones.

"Our holiday folder is no delete folder. I would say no delete or deletion which I can recover." P6 Group D

Deletion Concepts	
Accountability	Who has access to deleted data in the cloud
	Who has access to all data stored in the cloud
Backend, Anonymisation	What happens to deleted data
Backend	How data is deleted
	How is data stored in the cloud
	How much storage size is left
	The extent of data deletion (e.g., Complete deletion, Soft deletion)
	The location where data is stored
	The number of copies of data stored in the cloud
	What happens when a user deletes their cloud account
	How copies of data created by the provider are deleted.
Shared Folders	How data is deleted from a shared folder
	How shared folders work
	In whose account does data in shared folders reside
	How deletion from shared folders works (e.g., who has the right to delete)
Time	The time it takes the provider to completely delete from the recycle bin
	The time it takes to completely delete data from the cloud
	The time it takes for all copies of data to be deleted from the cloud
	The time it takes to completely delete a cloud account
User Interface	How data is deleted from a web interface
	How data is deleted from a "sync folder" or "cloud folder" in my computer
	How data is deleted from the cloud using a smartphone
Data Recovery	Data recovery after data has been deleted from 'deleted folder' or 'trash can'

Table 6. Summary: Deletion concepts considered and used in the study.

7 PAR3 – INFORMATION REQUIREMENTS

Activity design

This activity focused on information about the deletion. We created twenty-three (23) labels containing information related to the cloud and deletion. This information covered deletion concepts suggested important by prior research [44] such as time, shared folders, copies, back-end and user interface (UI). Table 6 shows the list of the information used in our study. Participants were asked to categorize these in three ways: (1) the order of importance with regards to deletion, (2) the point (time) when they would prefer to see the information, and (3) where they would expect or prefer to find that information (channel). However, unlike the data sorting activity (PAR1) where categories were not given, in this task, categories were given to the participants beforehand. The categories were given to simplify and reduce task time. Despite being given categories, participants were encouraged to add more categories they considered applicable.

Deletion Concepts	Critical	Important	Not critical	Before	During	After	Privacy Policy	Blog Pages	Pop-up Dialogs	Adverts	FAQs	Dashboard/UI
Who has access to deleted data in the cloud	4	0	0	3	0	1	4	0	0	0	0	0
Who has access to all data stored in the cloud	4	0	0	4	0	0	3	0	0	0	0	1
What happens to deleted data	2	0	2	4	0	0	2	1	1	0	0	0
How data is deleted	2	2	0	3	2	0	2	0	0	0	2	0
How is data stored in the cloud	0	0	4	4	0	0	1	1	0	0	2	0
How much storage size is left	2	0	2	1	2	1	0	0	2	0	0	2
The extent of data deletion	3	1	0	3	2	0	3	0	2	0	0	0
The location where data is stored	2	1	1	3	1	0	2	1	0	1	0	0
The number of copies of data stored in the cloud	0	2	2	3	0	1	0	2	1	0	0	1
What happens when a user deletes their cloud account	2	2	0	1	2	1	2	1	0	0	2	0
How copies of data created by the provider are deleted	1	2	1	3	1	0	1	3	0	0	0	0
How data is deleted from a shared folder	3	1	0	4	0	0	0	2	1	0	1	0
How shared folders work	2	2	0	3	1	0	1	1	0	0	2	0
In whose account does data in shared folders reside	1	2	1	4	0	0	3	1	0	0	0	0
How deletion from shared folders works (e.g., who can delete)	4	0	0	2	3	0	3	0	1	0	0	0
The time it takes the provider to completely delete from the recycle bin	1	1	2	1	2	1	0	1	3	0	0	0
The time it takes to completely delete data from the cloud	3	1	0	3	0	1	1	0	3	0	0	0
The time it takes for all copies of data to be deleted from the cloud	2	2	0	2	0	2	1	2	0	0	1	0
The time it takes to completely delete a cloud account	0	4	0	3	0	1	2	1	1	0	0	0
How data is deleted from a web interface	1	3	0	4	0	0	0	2	0	0	2	0
How data is deleted from a "sync folder" in my computer	1	2	1	4	0	0	2	1	0	0	1	0
How data is deleted from the cloud using a smartphone	1	3	0	4	0	0	0	2	0	0	2	0
Data recovery after data has been deleted from "deleted folder"	3	1	0	1	3	0	0	2	1	0	2	0
	Importance			Time			Channel of Communication					

Table 7. Deletion information preferences. In some instance, participants desired to have certain information before signing up or before deleting and during deletion. Regarding channel of communication, there were some cases where participants desired to have information made available through several channels.

Regarding importance, we gave participants three categories: the most important info, less important information and neutral group. Regarding *Time*, they could choose between *Before signing up or deleting*, *During usage or deletion* and *After deletion*. And lastly, with regards to communication channels, participants could choose from *Privacy Policies*, *Blog pages*, *Interactive Dialogs*, *Adverts*, *Frequently Asked Questions (FAQs)*. We chose these channels because they are mostly used to distribute information about cloud services.

To ensure common ground for deletion concepts, the lead researcher allowed participants to familiarise themselves with concepts and explained them to participants, especially those that confused them. For example, some participants were not familiar with the "sync folder."

PAR3 – Findings

We first present the results of PAR3 in Table 7. We show how participants categorized data with regards to importance, time and channel of communication. Table 8 discusses and summarizes these categorizations. We conclude this section by presenting the lessons learnt from this activity.

Following PAR3 and open coding (described in detail in Section 4), we learnt that

- information about deletion is less visible,
- while deleting or when considering deletion, users prefer to have relevant information that will inform their decisions,
- presenting users with information at the right time will inform users' decisions better,
- the channel of communication is essential. Important information should be made available in different places, and
- essential information should not be limited to privacy policies.

Deletion Concept	Importance	Time	Channel
Accountability	Participants considered info about who has access to their deleted data critical.	They prefer to know this information before using the cloud. In some cases, they preferred this info after they have deleted their data.	They mostly expect this information to be in privacy policies.
Anonymisation Backend	Participants had a divided opinion over the importance of knowing what happens to deleted data. Some participants considered it critical while others said it was not.	They preferred having this knowledge before using or deleting from the cloud.	Participants expect to be able to have access to this info through privacy policies, blog pages and interactive dialogs.
Backend	All the groups perceived info about how data is stored not essential. However, they mostly perceive backend info important if not critical with regards to deletion.	They generally want to know this info before using the cloud or attempting to delete. Some argued that they would want to know much storage is left and the extent of deletion while deleting or after deleting.	They prefer this info to be shared through privacy policies, blog pages and interactive dialogs. Participants stated they would expect info about the extent of deletion in privacy policies and interactive dialogs. Some participants explained that info about the deletion of copies of data should be made available in blog pages. They also suggested getting info about the amount of storage left and the number of copies of data existing in the cloud to be fully communicated in the account page or dashboard. One group stated they would expect info about where data is stored to be made available in advertisements.
Shared folder	Participants considered info about shared folders to be critical, particularly, how data is deleted from shared folders and who can delete from a shared folder. One group argued that info about whose account shared data resided was not critical to know compared to how data is deleted from shared folders.	They prefer getting information about deleting from shared folders before using the cloud or deleting. However, they highlighted they would want to know who can delete from shared folders while using the cloud or when attempting to delete.	Participants preferred having information about shared folders in privacy policies, blog pages, interactive pop-ups. For instance, they would expect to know who can delete through interactive dialogs when attempting to delete. Furthermore, they would also expect info about how data is deleted from shared folders to be distributed through FAQs and blog pages.
Time	Info about how long it takes to delete is considered at very least important if not critical to know. Participants mostly considered knowing the duration of deleting data completely from the cloud critical.	Participants expressed that info concerning duration should mostly be made available before the sign up to use the cloud. However, info about how long it will take to delete all copies of data from the cloud should be made available before and after deleting.	Participants mostly expected info about the duration to be found in Privacy policies when it is about completely deleting an account and deleting a “sync” folder from their local machine. However, they mostly expect info about the duration to be in blog pages, interactive dialogs and FAQs. They did not expect any of this info to be shared through their cloud dashboard/account pages or adverts.
Recovery	Participants perceive info about whether they can recover data from the cloud to be critical.	They expect to know this info while using the cloud or when they are deleting data from the cloud. Some participants argued that such info is better known before signing up or attempting to delete.	Info about data recovery was mainly expected to be found in blogs and FAQs.

Table 8. Summary. Deletion information preferences with regards to deletion concepts.

Information about deletion is less visible

Activity 3 revealed that most participants were unaware of what information about cloud deletion was available or where they could access it. For instance, during PAR3, many participants asked us whether all information used in the study was indeed available. They reported not having seen some of the information but expressed their desire to have access to such information. For example, all the participants from one group debated whether cloud providers have information about data recovery after deleting from a "trash can" or "deleted item" folder on their website. They mostly agreed that such information is not available in the cloud. We also learned that most participants generally assume that all important information about deletion will be found in privacy policies.

> *“These are the things I would never think of, this would be important for the company not me. Anyway, this is probably in the privacy policy of Dropbox.”* P6 Group D

Users prefer precise and relevant information

Participants do not want to have all the information about deletion through one channel of communication; they explained that having too much information in one place can be overwhelming and may cause people not to be interested. We learnt that participants prefer only to be exposed to information critical for decision making while using or deleting from the cloud. For instance, information about how data is stored in the cloud was deemed not important, but all the groups highlighted how deletion from shared folders works was very critical to know so they would expect to have easy access to such information.

> *"People like you would love knowing this kind of things. I don't really think its important to me. I don't care how it is stored. Can I get it when I want it yes... deletion maybe.. sometimes you want to know what happens when you delete something. "* P7 Group B

> *"For its like... can I recover this file. If yes, great. Google should get me to that information easily."* P15 Group D

> *"People should care about this and how to delete from shared folders. Everything down there is okay in my opinion."* P18 Group A

Presenting information at the right time informs users' decisions

Participants prefer to have the information they consider critical or important presented to them before signing up for cloud services or before attempting to delete it. They preferred knowing information that concerns 'how' and 'who' handles data before they sign up for a service (e.g., information about who has access to all data stored in the cloud). Participants also highlighted that they need to be informed about the status of the system after they have deleted it. We also learned that most participants prefer knowing how long it will take for copies of data to be deleted after they have requested deletion.

> *"yes, it may influence my decision to sign up. You don't want to sign up, then later find that I cannot delete my data. But, again, I don't do really do this when I want to register [to a cloud service], it's too much [information]."* P18 Group A

> *"In Mac[OS] after you install something, they ask you if you want to keep a file, I always delete it. Thats cool because I need that information then"* P2 Group B

> *"You don't want to delete and then be asked if you are sure the following day. I would panic. Information should just be there at the right time."* P6 Group D

Importance and channel of communication

We learnt that participants generally prefer to have most of the information they consider critical in privacy policies, though, this is contextual. However, unlike our previous study [54] where participants mentioned the desire to find some information in adverts, in this study advertisements were least preferred for sharing information. We also found that information considered least important may also be preferred in privacy policies and interactive dialogs.

> *"This is the kind of information you find in privacy policies. Lets put it under there. Imagine seeing this on your screen. Its overwhelming..."* P19 Group B

Beyond privacy policies

While privacy policies usually have information about deletion, and participants expected most of the info to be in privacy policies, we learnt that users do not always want to have all the information about deletion in the privacy policies (See Table 7). Our activity revealed that users occasionally prefer to have some information made available through some channels such as provider's blog

posts, account dashboard, adverts, FAQs, and interactive dialogs. We learnt that participants often expect technical information (i.e., info about how data deletion is achieved or completed) in blogs and FAQs. For example, all the participants highlighted that information about how to delete using the web interface or mobile application should be made available in FAQs and blogs. We also learnt that information about 'who' and 'what' is expected to be found in privacy policies while information about duration is expected to be instant and be made available through interactive dialogs, blogs and FAQs.

> *"I know we put this under privacy policies. But I don't read them. Sometimes, you want this kind of information somewhere nice. A short blog explaining what happens to deleted data would be useful."* P13 Group B

8 DISCUSSION

Our study revealed various insights concerning cloud deletion and information preferences. In this section, we review our activities and discuss the implication of our findings regarding deletion mechanisms with respect to deleting in individual and social settings. We conclude this section by further informing policy around cloud deletion.

8.1 Data Sorting Activity

Our data sorting (or data perception) task showed that data sensitivity varies and depends on individuals' and groups' understanding of risk pertaining to the concerned data. This compliments prior works [40, 60]. This finding further re-establishes that users' perception (including evaluation of risk) of data is subjective, particularly narrow in individual settings than groups. Somehow this is expected; groups discuss risk better. However, it does highlight the possibilities of users' underestimation of risk leading to data being deleted in a manner that leaves it vulnerable or damaging consequences.

The sorting task also suggests that data utility may be as valuable to users as sensitivity when choosing how data should be treated. This is not to say there is a trade-off between the two, but rather it is considered important as sensitivity. From our study, data considered important tend to influence group composition or properties. Participants grouped these data types despite some having varied perceived sensitivity. It may be interesting to investigate how much data utility influences perceived sensitivity or whether the two are mutually exclusive.

Our findings also revealed that risk was not the only aspect that influenced sorting; some groups discussed content, context, and other groups touched on the element of trust and safety. For example, all groups classified honeymoon photos as "important and sensitive." However, in some cases, participants mentioned it was not that the data was "important" utility-wise but because they aimed to protect the dignity of what the photo might contain. Concerning children's photos, some participants argued that while the photos themselves might be of value to them, they may not contain sensitive content. However, they are protecting the safety of children. Participants also discussed the element of trust during friends' photos. All the groups agreed that friends' photos were likely to contain common content between them and their friends; therefore, they may be important only in terms of "value," which may include an emotional attachment. Some participants mentioned that their friends would trust that they would keep such data safe.

Moreover, during group sorting, we also learned that data types were more likely to be moved from other groups to "important and sensitive" group than vice versa. While discussing such data, participants who had initially categorized data in a different way were more inclined to agree to have the data in important and sensitive than they initially thought. However, when discussing Pet care information, participants who had initially classified this as sensitive or important data

agreed to have it under unimportant and less sensitive. They mostly argued this information was more about the animal's health than them, though they know this information is linked to them. Prior studies [55, 71, 72] on Pet wearable have found that pet owners do not always understand that their pet wearable devices may also collect sensitive information about them.

8.2 Deletion preferences - How do users want to delete cloud data?

Users' desire to delete data vary and is dependent on various factors. While existing literature [52, 69, 70] has mostly accepted that perceived high sensitivity drives the choice for the most destructible deletion method, our findings suggest this is not held in some cases. In fact, deletion preferences do not entirely rely on how data is perceived or categorized. For instance, data perceived as not sensitive may warrant a destructible method of deletion. This suggests that the choice of deletion in the cloud goes beyond sensitivity; it is multi-dimensional and depends on the context of deletion. In individual contexts, deletion preferences are complex, diverse, and change over time as users' lives change, while in social contexts, they seem to be dependent on relationships, trust, creators, authorships, and their offline life situations. These differences suggest that users should be given choices regarding how they may delete data in the cloud.

8.3 Deletion information - How do they want to be informed about deletion?

Our results suggest that users want to know more about deletion than how data is stored in the cloud. This reaffirms our earlier study's findings that users are interested in knowing what happens to their data during deletion [54]. All the groups in our study only suggested that how data is stored in the cloud was not critical to know, somehow highlighting that most of the information about deletion is essential. Moreover, this discovery also supports the idea that users need to understand the deletion concepts better to make informed decisions. We also learned that most participants prefer knowing most information before deleting, sometimes even before signing up. This finding supports our argument that the user's choice of a service provider may depend on how the provider deletes data or communicate about the deletion. Surprisingly, most participants preferred to have most of the information about deletion in privacy policies despite a surplus of literature stating that users do not read privacy policies. We posit this may stem from users mixing where they expect to find information and where it should be placed. Nonetheless, our findings suggest that participants consider deletion information essential to contribute to how service providers should handle their data, or maybe participants searched for such information in privacy policies and could not find it. However, this may also suggest that users do not know what information about deletion exists in privacy policies; therefore, they make assumptions about where it should be.

8.4 Deletion Controls - Individual settings

Based on the results we obtained, we envision next-generation deletion controls with the following properties:

Intelligent and Personalized. In comparison with sharing preferences [36, 46, 62], deletion preferences also dependent on various factors: file attributes, relationships between owners, and mental models. We found that deletion preferences are not fixed; they change based on the context and time, suggesting that deletion needs or requirements cannot be generalized. There is no one-size-fits-all solution for meeting these preferences.

Our results also suggest that users sometimes prefer more than one type of deletion for particular data depending on the context. For instance, users may choose complete deletion or soft deletion over less essential and less sensitive data. This suggests that users do not only use one type of deletion for certain data but may use a different type depending on the context. Thus, the

complete deletion may not only be used for privacy purposes but also for destroying data they deem unnecessary. Moreover, we found that participants do not always choose the destructive method of deletion for data they consider private or confidential. We learnt that they also consider how easy it is to get or reproduce the data. This suggests that deletion preferences are complicated and do not always reflect users' privacy concerns.

The findings above suggest that translating users' deletion preferences into deletion controls will be challenging because the preferences are many, different, and dependent on many factors. There is a need for intelligent deletion controls that can adapt and accommodate different deletion preferences. Designers may have to employ AI techniques (such as machine learning) to implement smart deletion controls that can actively assist users in their deletion tasks—for instance, automatically clustering data types (as we observed people do themselves) with similar deletion preferences. Previous studies [7, 26] show some promising results regarding data classification concerning similarity, sensitivity, and usefulness. Cloud systems could automatically learn the deletion preferences for particular data per-user basis and establish suitable deletion defaults.

Interface-aware. As people use different interfaces (i.e., web interfaces, mobile apps, and sync folders) to access the cloud depending on their need, sharing audience, cost, and accessibility, designers should account for these factors. This suggests that users may be willing to incur the cost of setting their deletion preferences or switching devices to delete easily. Our previous work [54] attests to this; users switch between devices to avoid facing some deletion challenges. Designers should, therefore, consider whether one interface can contain all the necessary features required to offer deletion preferences or whether some features may be excluded. For example, to minimize effort when deleting through mobile devices, deletion mechanisms could be simplified to avoid complexity and cost. However, for web interfaces, users may be presented with more detailed controls.

Layered. Our results suggest that the level of detail of deletion preferences differ across users; fine-grained controls may interest some users while others may find them demanding. We, therefore, call on designers to develop deletion mechanisms that are layered; not only limited to fine-grained deletion choices but also coarse-grained ones for users who may find them too demanding. Fine-grained controls could focus on file properties and sharing context to account for costs and other use cases like associating a deletion type with a particular device. Coarse-grained deletion controls could include default settings that are easier to understand.

Retrospective. Our results suggest that deletion preferences may change over time. This suggests that users may have specific requirements for deleting old data or data that have not been modified for over a long period. As stated by prior work [25], users have an interest in managing old data; therefore, deletion mechanisms should cater for the deletion of old data. Users could be given a chance to define when data should be considered old and specify how it should be deleted. For instance, users could be notified when data has become old and be requested to take action on how it should be deleted.

8.5 Deletion controls - Social Contexts (Multi-party deletion mechanisms)

Multi-party issues are well understood in other domains like social media [56, 65–67] but are very limited in cloud computing [76]. In this section, we discuss how multi-party deletion in the cloud could be understood or improved.

Multiuser-aware. We found that participants refrained from deleting files they did not create or author from shared folders. Designers should, therefore, consider implementing deletion controls that take into account the multiuser nature of shared folders. For instance, members of a shared

folder can agree beforehand which folders or files could be deleted. A file creator can specify whether others can delete a file before and during upload. This may reduce conflicts that may arise from deleting from shared folders or choosing the type of deletion (e.g., recoverable or not) to use.

Conflicts. Deleting from shared folders may cause conflicts as deletion preferences can be different (or not align). The research could focus on understanding collaborative decision-making regarding deletion and what conflicts could arise from shared folders. Also, investigate what mechanisms could be implemented to reduce such conflicts. Current mechanisms allow uploaders to define privacy settings (including deletion) of the content, but there is no support for other members of the shared folder who may disagree or have different deletion preferences. Moreover, designers should consider providing reactive approaches to enable users to deal with conflicts or re-establish deletion rules or preferences. An intriguing endeavor would include investigating how group decisions actually translate to the choice of deletion in practice.

Ownership. Due to varying deletion requirements that arose from deleting from social contexts activity, we think ownership within shared folders may need to be defined. In a shared folder, one may assume the uploader is the owner, therefore, has the right to determine how an item should be deleted. However, there are cases where an item may be considered co-owned by others, e.g., photos. Understanding co-ownership is critical to design tools for managing deletion in the cloud.

Context aware mechanisms. This work and our previous work [54] did not investigate conflicts and consequences of deleting from social contexts. It is not clear whether these preferences include other people or if they considered others before deleting. Moreover, if they did, what they actually considered. Context-aware mechanisms could be used to help users consider others before executing their deletion preferences. Audience visualization could be employed to help users understand who will be impacted by their deletion decision. This can reduce the chances of accidental deletions (and conflicts) that may affect other people.

8.6 Deletion Information

Relevant information at the right time. Prior study [44] suggests that information about deletion should cover at least six areas: back-end, time, backup, derived information, anonymization and shared copies. While they do not suggest how and where such information could be presented to users, our study suggests this information should not only be constrained to privacy policies but be made available through other channels as well. Our results also suggest that sharing these concepts at the right time may improve users' understanding of deletion. For instance, explaining the retention period (i.e., time) immediately after the user has just deleted may help them realize that data is not entirely removed from the cloud. Similarly, users may also understand that the deletion process may sometimes include anonymization before data is removed entirely from the cloud.

Our previous work [54] suggested that users seek help (i.e., look for information on deletion) in order to delete (i.e., accomplish a task). However, this study further confirms that users seek information to delete only but not to improve their knowledge about the deletion. They only search for information about deletion when they need it. Designers could, therefore, categorize information about deletion into two groups: primary information and secondary information. Primary information could contain information that is essential for deletion and is needed by users to make confident decisions about deletion, while secondary information could focus on information that is not contextualized but may be necessary for users. Our results suggest critical information should be easily accessible, particularly during deletion, while other information could be made available to users upon request or through other channels. Moreover, research should not

only focus on improving privacy-related information on policies only but should investigate what deletion information is relevant for privacy policies.

Deletion status and summaries. Our results highlighted the importance of giving users the status of their deletion action; users want to know whether deletion completed successfully and what it means with regards to deleted data. Designers could inform users about the effects of their deletion; file moved to "deleted items folder" or how long it will take to remove the file from the cloud completely. Moreover, deleted data could have a timer showing when it is going to be completely removed from the deleted item folder. This would help users understand retention better; which data is still recoverable, and which one is not. For those who delete for privacy reasons, they would also get assurance that their data is removed from the cloud entirely.

Our results also suggest the need for deletion summaries. This feature would allow users to see their deletion records and help them audit their accounts, hence allowing them to reverse their decisions. For instance, through summaries, users may see which data has been deleted, how it was deleted and whether it is recoverable. Users could set how frequently they want to receive these reports or get them on-demand.

Limitations

One disadvantage of participatory action research is domination. It is possible that one or two participants may have dominated the group and overruled others. This may have limited the ability of others to freely express their views. To mitigate this, the researcher conducting the fieldwork stepped in to the discussions and encouraged quiet group members to share their views.

This study was exploratory and mainly qualitative in nature. While the sample was varied and roughly balanced across different demographic variables like gender, education and employment, other variables like age were varied but less balanced (18-45). We encourage additional studies with a larger and more diverse cloud user population, though getting older participants may be challenging as they tend to engage less with new technologies [47]. Moreover, while four groups with five participants are sufficient to create insightful discussions, it also limits the total range of experiences. Five people may have fewer experiences than a group of ten [45]. Moreover, it is also possible that in some cases, minority opinion can be one which may lead to the opinion being discarded easily. We attempted to minimize this by having an odd number of participants per group to create a chance of having a minority opinion of more than one. The hypothesis that we raised over the preferences on deletion and information about deletion should be quantitatively tested in a subsequent confirmatory study.

Moreover, despite explaining all the concepts to all the participants during each session, it is possible that some concepts were misunderstood or been interpreted differently. Our prior study [54] showed that users have different understandings of cloud and cloud deletion. We attempted in this study to create a shared understanding of all the concepts we used. However, we still posit that some concepts may have been misunderstood or conflated, for example, camouflage deletion and soft deletion or dashboard/UI and pop-up dialogs. We also noticed that during PAR3, some groups ignored the dashboard/UI category. This category may have been ambiguous and confused the participants. We also noticed this when analyzing advertisements; in our prior study [54], participants expressed the lack of information about deletion on cloud storage advertisements. However, during this study, only one group suggested that some information should be in the advertisement. Future studies could clarify these discrepancies.

While participants expressed different preferences for deleting data from cloud-based storage, it is possible that, in practice, they may not invest in such deletion preferences. Moreover, as our study revealed, these preferences may change over time (not stable), suggesting that the deletion

behavior or the choice of deletion may also change. Also, some participants mentioned that they would not delete some of the data we presented to them, suggesting that users may not delete their data despite the availability of deletion mechanisms/operations. This is highlighted by some studies in digital hoarding, for example, Sweeten et al. [68]. Future studies should explore the differences between the deletion preferences and the actual deletion behavior.

Lastly, during some exercises, few participants showed signs that might have led to task exhaustion. To mitigate this and the negative impact this might have in the results—e.g., participants rushing through the tasks, we introduced breaks in between the tasks.

9 CONCLUSION

As technology continues to evolve, features such as the deletion of data should also change and become flexible to meet customers' needs. This study lays the foundation for better cloud deletion controls and interfaces. We investigated cloud deletion preferences and the information that supports deletion in the cloud. We have shown that users have different deletion requirements—different data requires different types of deletion. Our results also show that deletion preferences differ significantly across people depending on their needs—there is no one size fits all solution; controls should be intelligent and personalized, interface-aware, layered, retrospective, and multiuser-aware. Our activities have shown how complex deletion preferences are, highlighting the challenges of designing controls for deletion. Our results also provide useful insights on how information about deletion can be improved to inform users better when deleting. In particular, we show what information about deletion users consider important, where they want it to be made available, and at which point in time of their use of the cloud.

With our work, we hope to bring attention to and inspire positive change in the space of data deletion in the cloud, pushing towards a distinct approach and truly user-centered deletion mechanisms. A plethora of existing literature has established that users' perception heavily influences their behavior; thus, it is essential to consider how users perceive various types of cloud deletion and their effects. This may help policymakers better understand how users perceive risk and how to ensure that users are protected from harms that may result from deletion. Our findings can also help service providers reflect on the type of information users view as important. Providers showing users what types of deletion methods are available may give users the impression that their providers understand their concerns and care about data deletion. There is also a need to have more deletion awareness campaigns, particularly in the cloud—its uses and consequences since more data is in the cloud. Relevant agencies should use relevant deletion terms to help people understand that deletion methods are different and have different outcomes.

ACKNOWLEDGMENTS

We would like to thank all the participants who took part in the study, and the editor and reviewers who provided us with valuable feedback.

REFERENCES

[1] Hazim Almuhimedi, Florian Schaub, Norman M. Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. 2015. Your Location has been Shared 5, 398 Times!: A Field Study on Mobile App Privacy Nudging. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, Seoul, Republic of Korea, April 18-23, 2015*, Bo Begole, Jinwoo Kim, Kori Inkpen, and Woontack Woo (Eds.). ACM, New York, NY, USA, 787–796. https://doi.org/10.1145/2702123.2702210

[2] Hazim Almuhimedi, Shomir Wilson, Bin Liu, Norman M. Sadeh, and Alessandro Acquisti. 2013. Tweets are forever: A large-scale quantitative analysis of deleted tweets. In *Computer Supported Cooperative Work, CSCW 2013, San Antonio, TX, USA, February 23-27, 2013*, Amy S. Bruckman, Scott Counts, Cliff Lampe, and Loren G. Terveen (Eds.). ACM, London, UNITED KINGDOM, 897–908. https://doi.org/10.1145/2441776.2441878

[3] Benett Axtell and Cosmin Munteanu. 2019. Back to Real Pictures: A Cross-generational Understanding of Users' Mental Models of Photo Cloud Storage. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–24.

[4] Marlyn Bennett. 2004. A review of the literature on the benefits and drawbacks of participatory action research. *First Peoples Child & Family Review* 1, 1 (2004), 19–32.

[5] Jarg Bergold and Stefan Thomas. 2012. Participatory research methods: A methodological approach in motion. *Historical Social Research/Historische Sozialforschung* Vol. 37 (2012), 191–222.

[6] Alan F Blackwell. 2006. The reification of metaphor as a design tool. *ACM Transactions on Computer-Human Interaction (TOCHI)* 13, 4 (2006), 490–530.

[7] Will Brackenbury, Galen Harrison, Kyle Chard, Aaron J. Elmore, and Blase Ur. 2021. Files of a Feather Flock Together? Measuring and Modeling How Users Perceive File Similarity in Cloud Storage. In *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*, Fernando Diaz, Chirag Shah, Torsten Suel, Pablo Castells, Rosie Jones, and Tetsuya Sakai (Eds.). ACM, New York, NY, USA, 787–797. https://doi.org/10.1145/3404835.3462845

[8] Robert Capra, Emily Vardell, and Kathy Brennan. 2014. File synchronization and sharing: User practices and challenges. *Proceedings of the American Society for Information Science and Technology* 51, 1 (2014), 1–10.

[9] Donald R Cooper, Pamela S Schindler, and Jianmin Sun. 2006. *Business research methods*. Vol. 9. Mcgraw-hill New York, New York.

[10] Alan Dix, Janet Finlay, Gregory D Abowd, and Russell Beale. 2004. *Human-computer interaction*. Pearson Education, 80 Strand, London.

[11] J. B. Earp, A. I. Anton, L. Aiman-Smith, and W. H. Stufflebeam. 2005. Examining Internet privacy policies within the context of user privacy values. *IEEE Transactions on Engineering Management* 52, 2 (May 2005), 227–237. https://doi.org/10.1109/TEM.2005.844927

[12] Liang Gou, Michelle X Zhou, and Huahai Yang. 2014. KnowMe and ShareMe: understanding automatically discovered personality traits from social media and user sharing preferences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, ACM, New York, NY, USA, 955–964.

[13] Hana Habib, Sarah Pearman, Jiamin Wang, Yixin Zou, Alessandro Acquisti, Lorrie Faith Cranor, Norman M. Sadeh, and Florian Schaub. 2020. "It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices. In *CHI '20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, April 25-30, 2020*, Regina Bernhaupt, Florian 'Floyd' Mueller, David Verweij, Josh Andres, Joanna McGrenere, Andy Cockburn, Ignacio Avellino, Alix Goguey, Pernille Bjøn, Shengdong Zhao, Briane Paul Samson, and Rafal Kocielnik (Eds.). ACM, New York USA, 1–12. https://doi.org/10.1145/3313831.3376511

[14] Feng Hao, Dylan Clarke, and Avelino Francisco Zorzo. 2016. Deleting Secret Data with Public Verifiability. *IEEE Trans. Dependable Secur. Comput.* 13, 6 (2016), 617–629. https://doi.org/10.1109/TDSC.2015.2423684

[15] Alan Henry. (Last accessed Jun 22, 2022). *Scan and Save Images of Your Passport and Prescriptions When Traveling*. Life hacker. https://lifehacker.com/scan-and-save-images-of-your-passport-and-prescriptions-927527185

[16] Judith A Holton. 2007. The coding process and its challenges. *The Sage handbook of grounded theory* 3 (2007), 265–289.

[17] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. 2015. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, October 12-16, 2015*, Indrajit Ray, Ninghui Li, and Christopher Kruegel (Eds.). ACM, New York, NY, USA, 781–792. https://doi.org/10.1145/2810103.2813603

[18] WhatsApp Inc. (accessed Jun 11, 2022). *Backing up to Google Drive*. Meta. https://faq.whatsapp.com/en/android/28000019/

[19] Iulia Ion, Niharika Sachdeva, Ponnurangam Kumaraguru, and Srdjan Capkun. 2011. Home is safer than the cloud!: privacy concerns for consumer cloud storage. In *Symposium On Usable Privacy and Security, SOUPS '11, Pittsburgh, PA, USA - July 20 - 22, 2011*, Lorrie Faith Cranor (Ed.). ACM, New York, NY, USA, 13. https://doi.org/10.1145/2078827.2078845

[20] Qatrunnada Ismail, Tousif Ahmed, Kelly Caine, Apu Kapadia, and Michael Reiter. 2017. To Permit or Not to Permit, That is the Usability Question: Crowdsourcing Mobile Apps' Privacy Permission Settings. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 119–137.

[21] Qatrunnada Ismail, Tousif Ahmed, Apu Kapadia, and Michael K. Reiter. 2015. Crowdsourced Exploration of Security Configurations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, Seoul, Republic of Korea, April 18-23, 2015*, Bo Begole, Jinwoo Kim, Kori Inkpen, and Woontack Woo (Eds.). ACM, New York, NY, USA, 467–476. https://doi.org/10.1145/2702123.2702370

[22] Maritza L. Johnson, Serge Egelman, and Steven M. Bellovin. 2012. Facebook and Privacy: It's complicated. In *Symposium On Usable Privacy and Security, SOUPS '12, Washington, DC, USA - July 11 - 13, 2012*, Lorrie Faith Cranor (Ed.). ACM, New York, NY, USA, 9. https://doi.org/10.1145/2335356.2335369

[23] Farzaneh Karegar, John Sören Pettersson, and Simone Fischer-Hübner. 2020. The dilemma of user engagement in privacy notices: Effects of interaction modes and habituation on user attention. *ACM Transactions on Privacy and Security (TOPS)* 23, 1 (2020), 1–38.

[24] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. 2009. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security, SOUPS 2009, Mountain View, California, USA, July 15-17, 2009 (ACM International Conference Proceeding Series)*, Lorrie Faith Cranor (Ed.). ACM, New York, NY, USA, 12. https://doi.org/10.1145/1572532.1572538

[25] Mohammad Taha Khan, Maria Hyun, Chris Kanich, and Blase Ur. 2018. Forgotten But Not Gone: Identifying the Need for Longitudinal Data Management in Cloud Storage. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 543, 12 pages. https://doi.org/10.1145/3173574.3174117

[26] Mohammad Taha Khan, Christopher Tran, Shubham Singh, Dimitri Vasilkov, Chris Kanich, Blase Ur, and Elena Zheleva. 2021. Helping Users Automatically Find and Manage Sensitive, Expendable Files in Cloud Storage. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael Bailey and Rachel Greenstadt (Eds.). USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, 94710, 1145–1162. https://www.usenix.org/conference/usenixsecurity21/presentation/khan-mohammad

[27] Nancy J. King and V.T. Raja. 2012. Protecting the privacy and security of sensitive customer data in the cloud. *Computer Law & Security Review* 28, 3 (2012), 308 – 319. https://doi.org/10.1016/j.clsr.2012.03.003

[28] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Zhang, Norman M. Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. 2016. Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions. In *Twelfth Symposium on Usable Privacy and Security, SOUPS 2016, Denver, CO, USA, June 22-24, 2016*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, 94710, 27–41. https://www.usenix.org/conference/soups2016/technical-sessions/presentation/liu

[29] Rui Liu, Jiannong Cao, Kehuan Zhang, Wenyu Gao, Junbin Liang, and Lei Yang. 2016. When privacy meets usability: unobtrusive privacy permission recommendation system for mobile apps based on crowdsourcing. *IEEE Transactions on Services Computing* 11 (2016), 864–878.

[30] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. 2011. Analyzing facebook privacy settings: user expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, ACM, New York, NY, USA, 61–70.

[31] Katharina Lobinger and Cornelia Brantner. 2019. Picture-sorting techniques: Card sorting and Q-sort as alternative and complementary approaches in visual social research. *The Sage Handbook of Visual Research Methods, 2nd Revised and Expanded Edition* 1 (2019), 309–321.

[32] Alexandra Ma and Ben Gilbert. 2019 (accessed June 16, 2022). Facebook understood how dangerous the Trump-linked data firm Cambridge Analytica could be much earlier than it previously said. https://www.businessinsider.com/cambridge-analytica-a-guide-to-the-trump-linked-data-firm-that-harvested-50-million-facebook-profiles-2018-3?r=US&IR=T#what-did-cambridge-analytica-do-1.

[33] Cathy Marshall and John C. Tang. 2012. That syncing feeling: early user experiences with the cloud. In *Designing Interactive Systems Conference 2012, DIS '12, Newcastle Upon Tyne, United Kingdom, June 11-15, 2012*. ACM, New York, NY, USA, 544–553. https://doi.org/10.1145/2317956.2318038

[34] Paolo Massa, Chiara Leonardi, Bruno Lepri, Fabio Pianesi, and Massimo Zancanaro. 2015. If You Are Happy and You Know It, Say "I'm Here": Investigating Parents' Location-Sharing Preferences. In *Human-Computer Interaction - INTERACT 2015 - 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part III (Lecture Notes in Computer Science, Vol. 9298)*, Julio Abascal, Simone D. J. Barbosa, Mirko Fetter, Tom Gross, Philippe A. Palanque, and Marco Winckler (Eds.). Springer, Bamberg, Germany, 315–332. https://doi.org/10.1007/978-3-319-22698-9_20

[35] Tara Matthews, Kathleen O'Leary, Anna Turner, Manya Sleeper, Jill Palzkill Woelfer, Martin Shelton, Cori Manthorne, Elizabeth F. Churchill, and Sunny Consolvo. 2017. Stories from Survivors: Privacy & Security Practices when Coping with Intimate Partner Abuse. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, May 06-11, 2017*, Gloria Mark, Susan R. Fussell, Cliff Lampe, m. c. schraefel, Juan Pablo Hourcade, Caroline Appert, and Daniel Wigdor (Eds.). ACM, New York, NY, USA, 2189–2201. https://doi.org/10.1145/3025453.3025875

[36] Michelle L Mazurek, JP Arsenault, Joanna Bresee, Nitin Gupta, Iulia Ion, Christina Johns, Daniel Lee, Yuan Liang, Jenny Olsen, Brandon Salmon, et al. 2010. Access control for home data sharing: Attitudes, needs and practices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, ACM, New York, NY, USA, 645–654.

[37] Aleecia M McDonald and Lorrie Faith Cranor. 2008. The cost of reading privacy policies. *ISJLP* 4 (2008), 543.

[38] William Melicher, Mahmood Sharif, Joshua Tan, Lujo Bauer, Mihai Christodorescu, and Pedro Giovanni Leon. 2016. (Do Not) Track me sometimes: users' contextual preferences for web tracking. *Proceedings on Privacy Enhancing Technologies* 2016, 2 (2016), 135–154.

[39] Kristopher Micinski, Daniel Votipka, Rock Stevens, Nikolaos Kofinas, Michelle L. Mazurek, and Jeffrey S. Foster. 2017. User Interactions and Permission Use on Android. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 362–373. https://doi.org/10.1145/3025453.3025706

[40] George R Milne, George Pettinico, Fatima M Hajjat, and Ereni Markos. 2017. Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs* 51, 1 (2017), 133–161.

[41] Gaurav Misra and Jose Such. 2016. How socially aware are social media privacy controls? *Computer* 49, 3 (2016), 96–99.

[42] Gaurav Misra and Jose Such. 2017. REACT: REcommending Access Control decisions To social media users. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, Sydney, Australia, July 31 - August 03, 2017*, Jana Diesner, Elena Ferrari, and Guandong Xu (Eds.). ACM, New York, NY, USA, 421–426. https://doi.org/10.1145/3110025.3110073

[43] Zhen Mo, Yan Qiao, and Shigang Chen. 2014. Two-Party Fine-Grained Assured Deletion of Outsourced Data in Cloud Systems. In *IEEE 34th International Conference on Distributed Computing Systems, ICDCS 2014, Madrid, Spain, June 30 - July 3, 2014*. IEEE Computer Society, New York USA, 308–317. https://doi.org/10.1109/ICDCS.2014.39

[44] 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, Baltimore, MD, 329–339. https://www.usenix.org/conference/soups2018/presentation/murillo

[45] Caroline J Oates and Panayiota J Alevizou. 2017. *Conducting focus groups for Business and Management students*. SAGE, 55 City, London UK.

[46] Judith S. Olson, Jonathan Grudin, and Eric Horvitz. 2005. A study of preferences for sharing and privacy. In *Extended Abstracts Proceedings of the 2005 Conference on Human Factors in Computing Systems, CHI 2005, Portland, Oregon, USA, April 2-7, 2005*, Gerrit C. van der Veer and Carolyn Gale (Eds.). ACM, New York, NY, USA, 1985–1988. https://doi.org/10.1145/1056808.1057073

[47] Katherine E. Olson, Marita A. O'Brien, Wendy A. Rogers, and Neil Charness. 2011. Diffusion of Technology: Frequency of use for Younger and Older Adults. *Ageing International* 36, 1 (01 Mar 2011), 123–145. https://doi.org/10.1007/s12126-010-9077-9

[48] Rachel Pain and Peter Francis. 2003. Reflections on participatory research. *Area* 35, 1 (2003), 46–54.

[49] Nikhil Patnaik, Joseph Hallett, and Awais Rashid. 2019. Usability Smells: An Analysis of Developers' Struggle With Crypto Libraries. In *Fifteenth Symposium on Usable Privacy and Security, SOUPS 2019, Santa Clara, CA, USA, August 11-13, 2019*, Heather Richter Lipford (Ed.). USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, 94710, 245—-257. https://www.usenix.org/conference/soups2019/presentation/patnaik

[50] Irene Pollach. 2007. What's wrong with online privacy policies? *Commun. ACM* 50, 9 (2007), 103–108.

[51] Emilee J. Rader. 2009. Yours, mine and (not) ours: social influences on group information repositories. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, April 4-9, 2009*. ACM, New York, NY, USA, 2095–2098. https://doi.org/10.1145/1518701.1519019

[52] Arthur Rahumed, Henry C. H. Chen, Yang Tang, Patrick P. C. Lee, and John C. S. Lui. 2011. A Secure Cloud Backup System with Assured Deletion and Version Control. In *2011 International Conference on Parallel Processing Workshops, ICPPW 2011, Taipei, Taiwan, Sept. 13-16, 2011*, Jang-Ping Sheu and Cho-Li Wang (Eds.). IEEE Computer Society, Manhattan, New York, U.S, 160–167. https://doi.org/10.1109/ICPPW.2011.17

[53] Kopo M. Ramokapane, Awais Rashid, and Jose M. Such. 2016. Assured Deletion in the Cloud: Requirements, Challenges and Future Directions. In *Proceedings of the 2016 ACM on Cloud Computing Security Workshop, CCSW 2016, Vienna, Austria, October 28, 2016*, Edgar R. Weippl, Stefan Katzenbeisser, Mathias Payer, Stefan Mangard, Elli Androulaki, and Michael K. Reiter (Eds.). ACM, New York, NY, USA, 97–108. https://doi.org/10.1145/2996429.2996434

[54] Kopo M. Ramokapane, Awais Rashid, and Jose M. 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, Santa Clara, CA, USA, July 12-14, 2017*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, 94710, 241–256. https://www.usenix.org/conference/soups2017/technical-sessions/presentation/ramokapane

[55] Kopo M. Ramokapane, Dirk van der Linden, and Anna Zamansky. 2019. Does my dog really need a gadget?: What can we learn from pet owners' amotivations for using pet wearables?. In *ACI'19: Sixth International Conference on Animal-Computer Interaction, Haifa Israel, November 12-14, 2019*. ACM, New York, NY, USA, 6:1–6:6. https://doi.org/10.1145/3371049.3371054

[56] Yasmeen Rashidi, Tousif Ahmed, Felicia Patel, Emily Fath, Apu Kapadia, Christena Nippert-Eng, and Norman Makoto Su. 2018. "You don't want to be the next meme": College Students' Workarounds to Manage Privacy in the Era of Pervasive Photography. In *Fourteenth Symposium on Usable Privacy and Security, SOUPS 2018, Baltimore, MD, USA,*

August 12-14, 2018, Mary Ellen Zurko and Heather Richter Lipford (Eds.). USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, 94710, 143–157. https://www.usenix.org/conference/soups2018/presentation/rashidi

[57] Joel Reardon, David A. Basin, and Srdjan Capkun. 2013. SoK: Secure Data Deletion. In *2013 IEEE Symposium on Security and Privacy, SP 2013, Berkeley, CA, USA, May 19-22, 2013*. IEEE Computer Society, Manhattan, New York, U.S, 301–315. https://doi.org/10.1109/SP.2013.28

[58] F. Schaub, R. Balebako, and L. F. Cranor. 2017. Designing Effective Privacy Notices and Controls. *IEEE Internet Computing* 21, 3 (May 2017), 70–77. https://doi.org/10.1109/MIC.2017.75

[59] Theodor Schnitzler, Christine Utz, Florian Farke, Christina Pöpper, and Markus Dürmuth. 2020. Exploring user perceptions of deletion in mobile instant messaging applications. *J. Cybersecur.* 6, 1 (2020), tyz016. https://doi.org/10.1093/cybsec/tyz016

[60] Eva-Maria Schomakers, Chantal Lidynia, Dirk Müllmann, and Martina Ziefle. 2019. Internet users' perceptions of information sensitivity–insights from Germany. *International Journal of Information Management* 46 (2019), 142–150.

[61] Manya Sleeper, Justin Cranshaw, Patrick Gage Kelley, Blase Ur, Alessandro Acquisti, Lorrie Faith Cranor, and Norman M. Sadeh. 2013. "I read my Twitter the next morning and was astonished": a conversational perspective on Twitter regrets. In *2013 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI '13, Paris, France, April 27 - May 2, 2013*, Wendy E. Mackay, Stephen A. Brewster, and Susanne Bødker (Eds.). ACM, New York, NY, USA, 3277–3286. https://doi.org/10.1145/2470654.2466448

[62] Manya Sleeper, William Melicher, Hana Habib, Lujo Bauer, Lorrie Faith Cranor, and Michelle L. Mazurek. 2016. Sharing Personal Content Online: Exploring Channel Choice and Multi-Channel Behaviors. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*, Jofish Kaye, Allison Druin, Cliff Lampe, Dan Morris, and Juan Pablo Hourcade (Eds.). ACM, New York, NY, USA, 101–112. https://doi.org/10.1145/2858036.2858170

[63] Robert H Sloan and Richard Warner. 2014. Beyond notice and choice: Privacy, norms, and consent. *J. High Tech. L.* 14 (2014), 370.

[64] Harsha Srinivas. (last accessed June 11, 2022). *Is it safe to store personal IDs like scanned copies of passport, on Google drive?* Quora. https://www.quora.com/Is-it-safe-to-store-personal-IDs-like-scanned-copies-of-passport-on-Google-drive

[65] J. Such and N. Criado. 2016. Resolving Multi-Party Privacy Conflicts in Social Media. *IEEE Transactions on Knowledge and Data Engineering* 28, 7 (2016), 1851–1863. https://doi.org/10.1109/TKDE.2016.2539165

[66] Jose Such and Natalia Criado. 2018. Multiparty Privacy in Social Media. *Commun. ACM* 61, 8 (July 2018), 74–81. https://doi.org/10.1145/3208039

[67] Jose Such, Joel Porter, Sören Preibusch, and Adam Joinson. 2017. Photo Privacy Conflicts in Social Media: A Large-Scale Empirical Study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3821–3832. https://doi.org/10.1145/3025453.3025668

[68] George Sweeten, Elizabeth Sillence, and Nick Neave. 2018. Digital hoarding behaviours: Underlying motivations and potential negative consequences. *Computers in Human Behavior* 85 (2018), 54–60.

[69] Yang Tang, Patrick P. C. Lee, John C. S. Lui, and Radia J. Perlman. 2010. FADE: Secure Overlay Cloud Storage with File Assured Deletion. In *Security and Privacy in Communication Networks - 6th Iternational ICST Conference, SecureComm 2010, Singapore, September 7-9, 2010. Proceedings (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 50)*, Sushil Jajodia and Jianying Zhou (Eds.). Springer, Berlin, Heidelberg, 380–397. https://doi.org/10.1007/978-3-642-16161-2_22

[70] Yang Tang, Patrick P. C. Lee, John C. S. Lui, and Radia J. Perlman. 2012. Secure Overlay Cloud Storage with Access Control and Assured Deletion. *IEEE Trans. Dependable Secur. Comput.* 9, 6 (2012), 903–916. https://doi.org/10.1109/TDSC.2012.49

[71] Dirk Van Der Linden, Matthew Edwards, Irit Hadar, and Anna Zamansky. 2020. Pets without PETs: on pet owners' under-estimation of privacy concerns in pet wearables. *Proc. Priv. Enhancing Technol.* 2020, 1 (2020), 143–164.

[72] Dirk Van Der Linden, Anna Zamansky, Irit Hadar, Barnaby Craggs, and Awais Rashid. 2019. Buddy's wearable is not your buddy: Privacy implications of pet wearables. *IEEE Security & Privacy* 17, 3 (2019), 28–39.

[73] Francesco Vitale, Izabelle Janzen, and Joanna McGrenere. 2018. Hoarding and Minimalism: Tendencies in Digital Data Preservation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI 2018, Montreal, QC, Canada, April 21-26, 2018*, Regan L. Mandryk, Mark Hancock, Mark Perry, and Anna L. Cox (Eds.). ACM, New York, NY, USA, 587. https://doi.org/10.1145/3173574.3174161

[74] Francesco Vitale, William Odom, and Joanna McGrenere. 2019. Keeping and Discarding Personal Data: Exploring a Design Space. In *Proceedings of the 2019 on Designing Interactive Systems Conference, DIS 2019, San Diego, CA, USA, June 23-28, 2019*, Steve Harrison, Shaowen Bardzell, Carman Neustaedter, and Deborah G. Tatar (Eds.). ACM, New York USA, 1463–1477. https://doi.org/10.1145/3322276.3322300

[75] Amy Voida, Judith S. Olson, and Gary M. Olson. 2013. Turbulence in the clouds: challenges of cloud-based information work. In *2013 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI '13, Paris, France, April 27 - May 2, 2013*, Wendy E. Mackay, Stephen A. Brewster, and Susanne Bødker (Eds.). ACM, New York, NY, USA, 2273–2282. https://doi.org/10.1145/2470654.2481313

[76] Dominik Wermke, Nicolas Huaman, Christian Stransky, Niklas Busch, Yasemin Acar, and Sascha Fahl. 2020. Cloudy with a Chance of Misconceptions: Exploring Users' Perceptions and Expectations of Security and Privacy in Cloud Office Suites. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA, 94710, 359–377. https://www.usenix.org/conference/soups2020/presentation/wermke

[77] Yu Xu and Michael J. Lee. 2018. Shopping as a Social Activity: Understanding People's Categorical Item Sharing Preferences on Social Networks. In *Joint Proceedings of the ACM IUI 2018 Workshops co-located with the 23rd ACM Conference on Intelligent User Interfaces (ACM IUI 2018), Tokyo, Japan, March 11, 2018 (CEUR Workshop Proceedings, Vol. 2068)*, Alan Said and Takanori Komatsu (Eds.). CEUR-WS.org, New York, NY, USA, 12. http://ceur-ws.org/Vol-2068/humanize4.pdf