Exploring User-Suitable Metaphors for Differentially Private Data Analyses

Farzaneh Karegar and Ala Sarah Alaqra, *Karlstad University;*
Simone Fischer-Hübner, *Karlstad University and Chalmers University of Technology*

https://www.usenix.org/conference/soups2022/presentation/karegar

This paper is included in the Proceedings of the Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022).

August 8–9, 2022 • Boston, MA, USA

978-1-939133-30-4

Open access to the Proceedings of the Eighteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

Exploring User-Suitable Metaphors for Differentially Private Data Analyses

Farzaneh Karegar
Karlstad University

Ala Sarah Alaqra
Karlstad University

Simone Fischer-Hübner
Karlstad University,
Chalmers University of Technology

Abstract

Despite recent enhancements in the deployment of differential privacy (DP), little has been done to address the human aspects of DP-enabled systems. Comprehending the complex concept of DP and the privacy protection it provides could be challenging for lay users who should make informed decisions when sharing their data. Using metaphors could be suitable to convey key protection functionalities of DP to them. Based on a three-phase framework, we extracted and generated metaphors for differentially private data analysis models (local and central). We analytically evaluated the metaphors based on experts' feedback and then empirically evaluated them in online interviews with 30 participants. Our results showed that the metaphorical explanations can successfully convey that perturbation protects privacy and that there is a privacy-accuracy trade-off. Nonetheless, conveying information at a high level leads to incorrect expectations that negatively affect users' understanding and limits the ability to apply the concept to different contexts. In this paper, we presented the plausible suitability of metaphors and discussed the challenges of using them to facilitate informed decisions on sharing data with DP-enabled systems.

1 Introduction

Differential privacy (DP) is a mathematically rigorous definition of privacy initially formalized in 2006 by Cynthia Dwork [20] for the calculation of statistics on a dataset. DP places a formal bound on the leakage of information from these statistics about individual data points within dataset. Informally, for each person who submits their data to a differentially private data analysis, DP assures that the output of such analysis will be approximately the same, regardless of the contribution of their data to the data sample under analysis. Differentially private mechanisms perturb data in a controlled manner. This allows quantifying privacy through a privacy loss parameter ε, thereby fulfilling the assurance. Although leading to more privacy, lower privacy loss parameter values negatively affect the accuracy of the results. Consequently, there is a trade-off between privacy and accuracy in differentially private data analyses.

Within the past few years, several large companies, including Apple [42], Google [24], Microsoft [19], Uber [28] and LinkedIn [30], integrated differentially private mechanisms into their systems. The United States Census Bureau also adopted DP to prevent information disclosure in the summary statistics it released for the 2020 Decennial Census [5]. Further, different variants and extensions of DP have been proposed for other types of data analysis scenarios, such as local DP or DP for federated learning. Variants have local or central security models, and the choice of model has a considerable impact on the types of adversarial behaviour the system can tolerate.

Given the growing deployment of differentially private mechanisms in different variants and contexts, there is a need to address the human aspects of DP-enabled systems. In this work, we focus on conveying differentially private data analyses to data subjects who would share their data with systems deploying DP. The data subjects are mainly lay users without any expertise or knowledge about privacy or DP. However, they need to make informed privacy decisions about sharing their data when confronted with DP-enabled systems and services. Usable transparency of the functionality of the underlying differentially private mechanisms could help data subjects form correct mental models of how their data is protected, thus facilitating their decisions. Researchers have shown that how DP is described in practice is insufficient to help users make informed decisions [17]. Therefore, we need to explore how and to what extent the differentially private mechanisms

USENIX Symposium on Usable Privacy and Security (SOUPS) 2022.
August 7–9, 2022, Boston, MA, United States.

should be explained to users. Further, the issues surrounding their understanding and mental models of differentially private data analyses, their perceptions of the privacy provided, and their trust in such systems should be addressed. Therefore, in this work our main objective is to contribute to the body of knowledge on usable DP and to investigate how to effectively explain the underlying differentially private data analyses to lay users to facilitate their decisions.

In explaining a system to users, design models can employ familiar metaphors [14]. When the aim is to explain or represent a complex, abstract or novel domain (i.e. target domain), it is often helpful to use metaphors or analogies and make a vivid comparison to a familiar and often simpler domain (i.e. source domain) that people already understand. Some researchers argue that while in metaphor-based descriptions the mapping of ideas involves transformation, with analogies a direct transferal is made between existing knowledge and a novel domain [9]. However, in this paper, we do not make a distinction between these two concepts. Using an appropriate metaphor, people's understanding can be enhanced by encouraging them to use their pre-existing knowledge of the source domain to structure their thinking about the target of explanation [13]. Consequently, we assume that metaphors can be used to convey the concept of DP and its privacy functionality, as comprehending the complex concept of DP and the privacy protection it provides could be challenging for lay users. Nonetheless, suitable metaphors for complex concepts need to be generated with care and evaluated for their effectiveness in each context.

Previously, Demjaha et al. [18] benefited from the framework proposed by Alty et al. [9] to generate and evaluate explanatory metaphors for E2E encryption. We employed and adapted this framework to generate and evaluate metaphors for DP in the context of different data analyses. Our focus is on pictorial metaphors elaborated with short, simple text. Our approach consists of three phases: 1) metaphor generation, 2) metaphor analytical evaluations based on expert analysis and 3) metaphor empirical evaluations involving lay users. The first two phases resulted in adapted metaphors, their analytical evaluations and a functionality list that can be used to analytically evaluate the suitability of metaphors to convey the privacy functionality of differentially private data analysis to users. The details of the first two phases have been published in [29]. This paper focuses on the third phase but briefly describes the other two phases for clarity. In the third phase, we empirically evaluated the metaphors from phase 2 and addressed the following research questions.

RQ1. What information about the underlying differentially private systems is required by users to decide about using such systems (i.e. sharing their data)?

RQ2. What are users' perceptions about the data privacy provided by the proposed metaphors of DP?

RQ3. To what extent are our proposed metaphors suitable for conveying the concept of DP to lay users in the context of different differentially private data analyses?

To address our questions, we conducted 30 online interviews. We defined three differentially private data analysis scenarios in the context of eHealth for local DP, typical central DP, and central DP for federated learning. Each interviewee was exposed to one of the scenarios and the related adapted metaphor(s) from phase 2. Interviewees responded to questions about their opinions and understanding of the metaphors.

We extended the previous findings on how DP should be explained to data subjects to facilitate their data sharing decisions. Our empirical evaluations provide information on the extent of the suitability of our proposed metaphors to explain DP to lay users and confirm the (plausible) suitability of metaphors while revealing specific challenges that must be addressed.

2 Background

Definition of DP. As defined by Dwork et al. [21], a randomized mechanism A is ε-differentially private, where $0 \leq \varepsilon$, iff for any two data sets D and D' that differ in at most one record, and any set R of possible outputs of A, we have $Pr[A(D) \in R] \leq e^{\varepsilon} * Pr[A(D') \in R]$. The definition prevents an attacker who knows all but one record in a database from inferring the last one after viewing the output. Simply put, DP mechanisms guarantee the stability of the output of a function based on changes that may happen in the input. Such a guarantee can facilitate releasing statistics on a database while preserving individuals' privacy in the database.

Different models. Differentially private mechanisms can be implemented as local or central (aggregate-level) models. In central models, a trusted data analyst (data curator/aggregator) gathers data from individual users and processes the data in a way that satisfies DP before publishing the aggregate statistics, similar to the original definition of DP [21]. In local models, users do not need to trust the entity responsible for analysis because their data get perturbed before being shared. The information disclosure risks differ substantially between these two models. However, in communicating with users, industry and media outlet DP descriptions do not clearly distinguish between central and local models, as reported by Cumming et al. [17]. Therefore, to address the limitation of existing descriptions, we defined three scenarios of differentially private data analyses in the context of eHealth.

Data analysis scenarios. The first scenario (SC1) is related to the local model of DP (Figure 3a in Appendix A) in which user data gets perturbed before being shared with the health company, which might not be trusted. For central models, in one of the scenarios (SC2) we have one data aggregator, a health company that conducts differentially private data analysis on actual information it collects from users and combines (Figure 3b in Appendix A). The other aggregate-level scenario (SC3) is related to differentially private federated learn-

ing where we have several data aggregators (different health companies) that collaboratively make an improved machine-learning model. They train a model collaboratively with the help of an Internet-based analyser (IBA) while preserving the privacy of their users (Figure 3c in Appendix A).

3 Related work: Usable differential privacy

Although technical literature on differentially private mechanisms and how to enhance them abound (e.g. [22,31,33]), just a small body of work focuses on the ethical, legal [15,16,36], and Human-Computer Interaction (HCI) implications of DP [12, 17, 47]. Among the considerable body of work on privacy communications (e.g. [6,34,39,41,48]), only a limited amount of research has focused on the communication of DP with different types of users. For instance, DPComp [26], PSI [25], Overlook [43], DPP [27], and ViP [35] provide interfaces for interacting with DP. However, the target groups of such tools are, for example, data curators/data analysts who may decide about the privacy budget based on their privacy and utility requirements. To the best of our knowledge, only three works have focused on explaining DP to end users (i.e. data subjects), which is more closely related to what we aimed for in this paper.

Bullek et al. [12] used a virtual spinner to describe the randomized response technique (RRT), a specific local DP technique, to users. They investigated whether users trust the RRT mechanism and if they adjust their privacy decisions when they see more details of the privacy promises implied by the RRT. Bullek et al. [12] reported that users vastly preferred the most anonymous spinner, although some participants preferred the most truthful spinner because they thought it minimized the ethical consequences of lying. We also use the spinner metaphor to describe DP in a local model. However, our spinner metaphor conveys the privacy-accuracy trade-off. Consequently, our results regarding users' preferences for which spinner to use differ from what Bullek et al. reported. In addition, our study reveals the shortcoming of the spinner metaphor for lay users and how it can be improved.

Xiong et al. [47] analysed the effects of using different short textual descriptions to inform users that their information is protected with DP on their willingness to share different types of information (sensitive and nonsensitive). They slightly modified and adapted descriptions from the companies/organizations that deployed and communicated DP to the public. Their results show that although users struggled to understand the DP descriptions, the descriptions explaining implications, that is, what happens if the aggregator's database is compromised, could facilitate people's data-sharing decisions and their comprehension of the local and central models.

Cummings et al. [17], in a series of online surveys, exposed their people to short verbal DP descriptions derived from publicly available descriptions of DP and investigated respondents' privacy expectations of DP-enabled systems and their willingness to share data in such systems and showed that common privacy concerns can be addressed by DP. However, how DP is described in the real world haphazardly raises privacy expectations that may mislead users about the systems' privacy features. Results of studies in [17,47] show the need for better DP descriptions for users.

To the best of our knowledge, no attempts have yet been made to generate, test and compare metaphors for conveying the underlying differentially private data analysis (both local and central model) to lay users.

4 Method

Figure 1 shows an overview of our approach which is based on a framework proposed by Alty et al. [9]. The framework provides suitable tools and techniques for metaphor design for interactive systems. Demjaha et al. [18] previously adapted the framework and analytically and empirically evaluated the efficacy of their explanation metaphors for E2E encryption. Due to contextual differences, to reach our objective, we applied the adapted and extended version of the framework. Particularly, two rounds of analytical evaluations are included and the steps related to the integration of metaphors into the user interfaces of real systems are excluded. More details on phases 1 and 2 and the design of interviews as part of phase 3 are provided in Section 4.1 and Section 4.2, respectively.

4.1 Phase 1 and Phase 2

Phase 1: metaphor generation. We used both the *extension* and the *design metaphor* techniques proposed by Alty et al. [9] to generate metaphors in our work.

To begin with, we reviewed literature and media outlets to see how others conveyed the concept of DP to users using metaphors or analogies. Our literature review uncovered that, for the first time, Warner [45] proposed randomization of responses by a spinner to improve the reliability of them to sensitive questions. The spinner metaphor was later used by Bullek et al. [12] to convey DP to lay users. The spinner has also been used in media outlets to convey how DP works [2]. We searched the Web for *differential privacy* alone and in combination with the keywords users, people, definition and introduction. We examined each of the first five pages of the results to find explanations (in any format, including videos) describing the concept at a high level. Our search on media outlets showed that DP is explained to people using an example of tossing a coin for changing user responses [1], noisy sound waves of radio channels [4] and a noisy portrait [3]. Investigating how companies described DP to their users did not result in any other metaphors we could use in our study.

In phase 3, we monitored and analysed users' language when they talked about their perception, and opinions of DP and the metaphors to which they were exposed to see whether further metaphors could be derived.

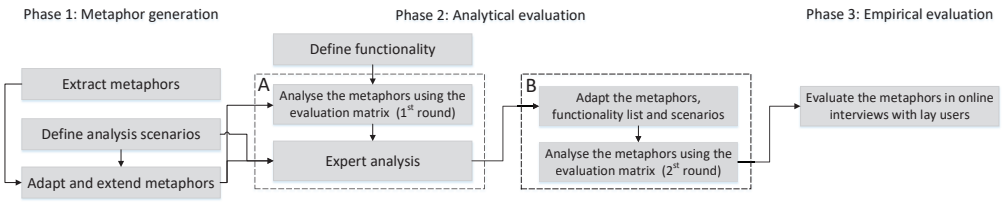


Figure 1: The steps followed to address the research questions.

Phase 2: analytical evaluation. Metaphor-system pairing is the step in Alty et al.'s framework [9] to analytically evaluate metaphors for which system functionality should be defined and then compared to features implied or conveyed by a metaphor. Therefore, we defined general privacy features of differentially private analyses and analysed metaphor-system pairings based on a metaphor evaluation matrix adapted from [18]. The matrix helps categorize the comparison of DP features with the features a metaphor (M) supports into desirable (DP+M+), undesirable (DP+M-), and very undesirable (i.e. conceptual baggage – DP-M+) groups of features. The template of the metaphor evaluation matrix we used is provided in [29].

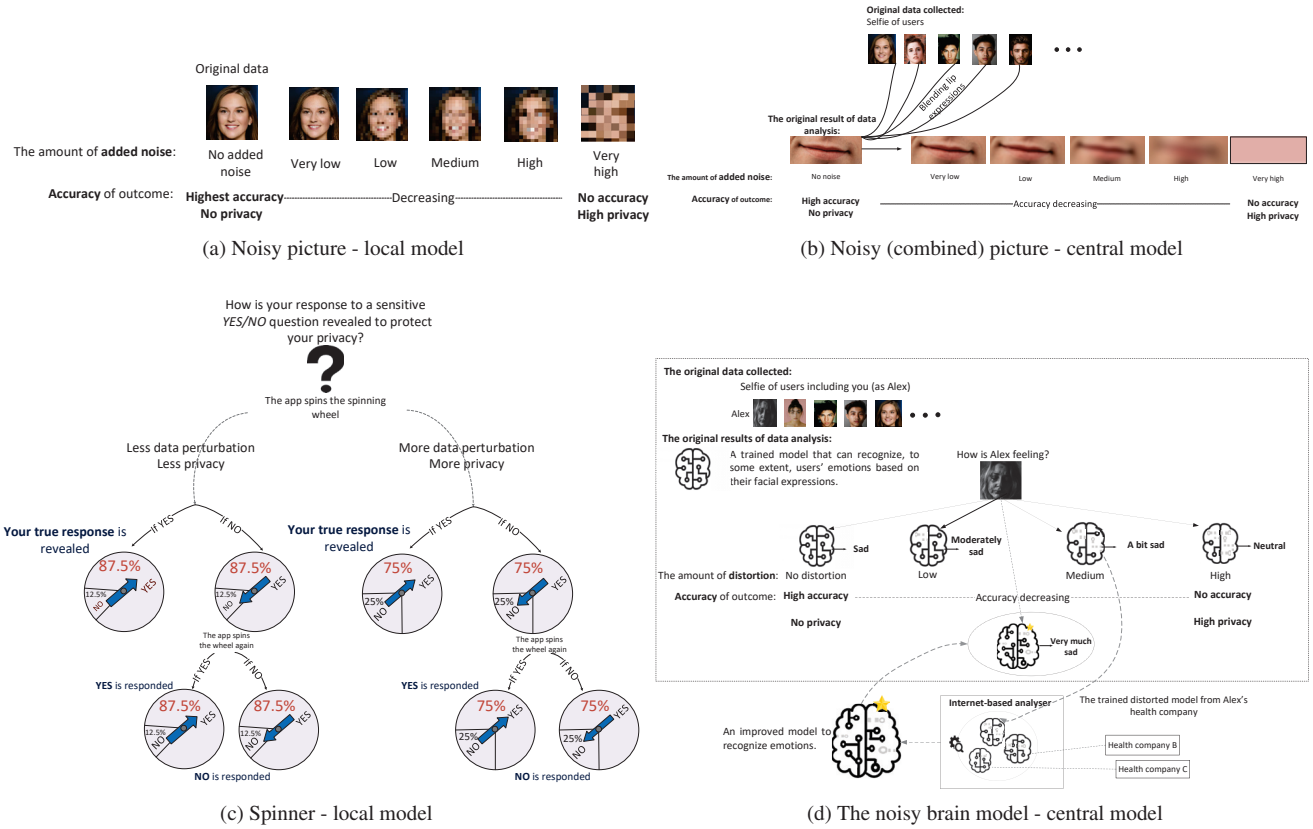
Eight privacy experts knowledgeable about DP from academia and industry reviewed our materials in step A of phase 2 (see Figure 1), including the description of scenarios, the original functionality list, the resulting metaphors in phase 1 and our first round of analytical evaluation. The purpose of the expert review was to improve the validity and to check the authenticity of our materials. We reached the experts through personal contacts and ongoing collaborations within joint projects. Based on the reviews, we adapted the metaphors, functionality list and scenarios and re-analysed the metaphors (step B of phase 2). Section 5 presents our functionality list and briefly describes the resulting metaphors from phase 2, depicted in Figure 2, which we tested in our interviews.

4.2 Phase 3: Interviews

To evaluate the metaphors in Figure 2 and address our RQs, we conducted online interviews (via Zoom) with lay users. The interviews differed based on the data analysis scenarios (SC1 to SC3 described in Section 2) and the related metaphor(s) to which the interviewee was exposed. The interviews had three stages: 1) a prelude session, 2) a main session and 3) an epilogue session. In the prelude session, after a brief introduction to the study the interviewees were asked for their consent and were provided with a link to answer optional demographic questions (age group, gender, educational background). The main session had two parts: 1) scenario introduction and gauging expectations and opinions (before exposure to metaphors) and 2) metaphor introduction and gauging perceptions and opinions.

The first part of the main session started with an introduction of a persona followed by the data analysis scenario. A persona was used to avoid the disclosure of personal information. The interviewees were exposed to the related pictorial presentation of a data analysis scenario (Appendix A). Simultaneously, the interviewer read the scenario description. The interviewees were informed about the general privacy problem in the scenario and the existence of DP to mitigate the problem. However, the information on how DP would work was not revealed yet. The description of SC1 read to participants is reported as an example in Appendix B. After the scenario introduction, participants played the role of the persona and were asked to make a decision on sharing their data based on the scenario. They then answered the interview questions and provided input on reasons for their decisions, requirements for further information on DP, perceptions of the benefits and risks if they agreed to share data, expectations of privacy protection in the scenario and factors that would help them their trust in DP to protect their privacy.

In the second part, each interviewee was exposed to the related pictorial metaphor(s) for the scenario (Figure 2). At the same time, the interviewer read the simplified description of DP as the accompanying information defining the metaphor. The exact accompanying information for each metaphor is provided in Appendix C. Afterwards, participants were requested to review the decision they previously made regarding sharing their data and to elaborate on their decisions to see how the DP description could have affected their decisions. They then provided their opinions on the understandability of the metaphor and how it could be improved. Questions about users' perceptions of distortion/perturbation and privacy provided by DP were asked. This was followed by questions to check whether the metaphor could convey the features in the functionality lists, including the privacy-accuracy trade-off and users' perceptions and preferences. Participants were prompted to elaborate on their opinions about the remaining privacy risks, whether they would trust DP to protect their data, to describe DP in their own words, and to suggest alternative descriptions of DP. The main session ended here for the SC2 and SC3 interviews. However, for SC1, half of the participants were first exposed to the spinner metaphor and then to the noisy picture metaphor, while the other half were exposed in the opposite order. After being exposed to the second metaphor, participants answered questions about their perceptions of the second metaphor. They also answered



(a) Noisy picture - local model

(b) Noisy (combined) picture - central model

(c) Spinner - local model

(d) The noisy brain model - central model

Figure 2: The metaphors to which participants were exposed in our interviews.

questions about which of the metaphors was easier to understand, conveyed privacy-accuracy trade-off in a better way, associated perturbation/distortion with privacy protection in a better way and which they preferred to be exposed to for DP description.

The interview was concluded in the epilogue session in which we asked participants to reflect on any issues we did not discuss in the interview. All interviewees who completed the interview were compensated with 20 GBP. The interview guide is provided in Appendix D.

Sampling and recruitment. We recruited 30 interviewees (10 for each scenario) from the Prolific platform. We used Prolific prescreening filters to recruit people whose current countries of residency were EU countries, EEA countries, the UK and Switzerland due to the scope of our funders. To sample lay participants without knowledge of DP, we excluded those with an educational background related to engineering, computing (IT) and computer science. Furthermore, at the beginning of the interview, we asked a few questions to gauge participants' initial familiarity with privacy-enhancing tools (PETs) and DP. We also conducted three pilot interviews (one for each scenario), which did not result in any major changes. Our participants' demographics are reported in Section 6.

Data analysis. We analysed our empirical data using thematic analysis [11]. Our data included approximately 36.6 hours of audio recordings from the interviews (SC1=13.6 h, SC2=13 h, SC3=13 h). A research assistant manually transcribed the audio recordings. All authors read and familiarized themselves with the content of the transcripts. Using NVIVO (software for qualitative data analysis), one author analysed the transcripts; this resulted in 1257 excerpts for SC1, 1255 for SC2 and 1142 for SC3. The other authors then reviewed the excerpts. All authors met afterwards for a workshop to discuss the codes and the code book and to agree on terminologies and resolve conflicts (disagreement percentage was 2%). We then finalized the code books for each scenario that were used as a basis for the following rounds of the analysis. Because we went through several iterations to discuss the codes and resolve conflicts, our disagreement percentage was calculated based on the final round of codes. Next, each author independently led a preliminary categorization (thematic analysis) of one scenario, and then reviewed the other two scenarios. All authors met afterwards in a second workshop to discuss the main themes of all scenarios. Following agreement abot the main themes, the second round of categorization of codes into the agreed-upon themes and further sub-themes was conducted. A third workshop was conducted to finalize the analysis and results.

Ethical considerations. The study was approved by the ethical advisor at Karlstad University. Interview data, including Zoom session recordings, were collected on the legal basis of informed consent given by the participants, and all data were processed in compliance with the General Data Protection Regulation (GDPR). Participants were instructed to use a non-identifying pseudonym as their Zoom name and to turn off their cameras during the recording to prevent the leakage of any identifying information. During the interviews, we introduced the scenarios in terms of a fictional user (persona) called Alex and asked the interviewees to answer questions from the perspective of Alex or in general and thus NOT to reveal any sensitive personal data, such as data related to their personal health conditions.

5 Functionality list and adapted metaphors

The following is the adapted functionality list after receiving feedback from experts:
(F1) A differentially private analysis, that is, a mechanism, bounds and quantifies the probability of additional privacy risk any individual would face because of their participation in a data analysis.
(F2) The privacy of a differentially private analysis is controlled by tuning a privacy loss parameter.
(F3) The smaller the value of the privacy loss parameter, the better the privacy guarantee for an individual.
(F4) The smaller the value of the privacy loss parameter, the less accurate the results of data analysis.
(F5) A differentially private analysis randomly perturbs data on an aggregate level (i.e. the results of the analysis) or individual level (i.e. the input data), depending on the context.
(F6) The amount of perturbation is controlled by the underlying differentially private analysis.
(F7) A differentially private analysis is resistant to privacy attacks based on auxiliary information, i.e. any past, present and future information an attacker may have.
(F8) A differentially private analysis does not promise unconditional freedom from privacy risks.

F1 can be interpreted in different ways. For the central model, it should convey that the results of a differentially private data analysis do not significantly depend on any particular individual's data. F1 can also be rephrased in terms of plausible deniability for a particular data record in the local model and participation in data analysis in the central model. Although a metaphor may not directly convey F1, it may imply one of its interpretations.

Considering our target group, we did not focus on the privacy loss parameter but on the role of perturbation in providing privacy and the effects of perturbation on the accuracy of the results. Therefore, if a metaphor conveys that more perturbation leads to better privacy but less accuracy we assume it covers F3 and F4. Further, we avoided including additional details to keep the features simple.

Table 1: Features of functionality list covered or not (Y/N) by each metaphor.

Metaphor\feature	F1	F3	F4	F5	F6	F7	F8	Context
Spinner	Y	Y	Y	Y	Y	Y	Y	Scenario 1
Noisy single picture	N	Y	Y	Y	Y	N	Y	Scenario 1
Noisy picture - combined	Y	Y	Y	Y	Y	Y	Y	Scenario 2
Distorted brain model	Y	Y	Y	Y	Y	Y	Y	Scenario 3

The development stages of our metaphors are defined in [29] in detail. In sum, in phase 1 we adapted and extended our initial metaphors (described in Section 4.1) because they were not necessarily suitable for all scenarios and all models of DP. In phase 2, based on the results of the expert analysis and our analytical evaluation, we excluded the metaphor of noisy sound waves of a radio channel due to features categorized as conceptual baggage and adapted our preliminary spinner metaphor to better communicate F3, F4 and F6.

The metaphors shown in Figure 2 were all defined for an eHealth application where users' stress levels are determined by analysing their face or lip expressions in pictures (selfies) contributed by the users. For SC1, we chose the metaphor of a *noisy picture* showing different levels of added noise with different degrees of pixelation (Figure 2a) and an adapted spinner metaphor showing two spinners with different biased outcomes mediating different levels of perturbation (Figure 2c). For SC2, a noisy combined picture metaphor was used to convey that noise is added to the aggregated data (i.e. the combination of pictures with lips is pixelated, Figure 2b) and not directly to individual records (users' selfies). For SC3, we used a *distorted brain*, for which some of the neural connections are greyed out, as a metaphor of a differentially private trained model (Figure 2d).

Table 1 shows whether each of our adapted metaphors conveys or implies the features in the functionality list, although it is subject to the validation of users. Features F3 to F6 and F8 are conveyed by all four metaphors. Until completely distorted, we can still have a useful analysis that may carry a risk of revealing information about individuals. F1 is implied by the spinner metaphor. However, the noisy picture metaphor for the local model (Figure 2a) does not cover F1 and F7. The noisy combined picture metaphor may convey F1 and F7, depending on the combination of all pictures picked for that metaphor. In addition, users' understanding of, for example, how much the aggregate-level picture might be revealing and if and how the added noise can circumvent privacy leakage from a combined picture may play a significant role. The distorted brain metaphor (Figure 2d) is quite abstract, and whether it conveys F1 and F7 is greatly dependent on what users know or understand from the concept of a model.

6 Interview results: Phase 3

Among the 30 participants (P1–P30), 13 identified themselves as female (SC3=2, SC2=7, SC1=4), 16 as male (SC3=8,

SC2=3, SC1=5), and one did not answer demographic questions. Our interviewees were relatively young; 18 were aged 18–25, 8 were aged 26–35 and 3 were aged 36–45. They had diverse academic backgrounds, including medicine, chemistry, psychology, cooking, international business and architectural design. Most were pursuing higher education studies. However, four of the interviewees indicated they were high school graduates or students. While the participants were not generally knowledgeable about PETs, some were aware of technologies or tools that help protect privacy. Encrypted messaging in specific communication apps, cookie consent forms, basic pseudonymization with reference to what Prolific does to hide users' real identity (e.g. using codes instead of emails/full names) and virtual private networks (VPNs) were mentioned in all three groups. Further, the participants were not knowledgeable about DP and had not heard about it before, meaning they were non-experts in privacy.

In total, our analysis resulted in 12 main themes. The main themes are indicated by a (T) and a unique number (X). Subthemes follow the format TX.X. When reporting a scenario-specific sub-theme, the scenario number follows the theme number in this format: (TX.X-SCX). If a sub-theme was common between all scenarios we omitted the scenario number. An overview of all themes is provided in Appendix E. We skip the theme number in the number format of a sub-theme when we report a sub-theme in this section for readability purposes. For example, instead of reporting (T1.1) we simply report (1). For SC1, the order of being exposed to the two metaphors (spinner/noisy picture first) had no significant difference in the results. The first four themes (T1–T4) are *pre-explanation* themes and the rest (T5–T12) are *post-explanation* themes. The explanation refers to the introduction of the DP metaphors relating to each scenario (see Section 4.2).

Information needed for trust and data sharing: Themes T1–T4 and T6–T8 address RQ1, as they shed light on the information affecting users' trust in and decisions to share their data with a DP-enabled system. The results show that the mere presence of a privacy technique is seemingly enough to persuade users to share their data. However, lack of transparency about DP leads to varied expectations and interpretations of who gets access to actual (raw) data, different assumptions (correct/incorrect) about DP and negative impacts on willingness to share data with and trust in a DP-enabled system. Most participants required usable transparency of DP, for example, to know how DP works, protects, and uses personal data and to know about the risks of identification.

T1: Factors affecting sharing of data. In all three scenarios, participants mentioned positive (1) and negative (2) factors affecting their decision to share their data with DP-enabled systems. Positive factors are the existence of a protection technique (1.1), transparency of DP (1.2), providing reassurance regarding data safety and reliability (1.3), the specific type of data and data processing purposes (1.4), good reputation/location of the company (1.5-SC2,3), the existence of DP as a trust factor (1.6-SC2,3), contribution to the improvement of the health app (1.7-SC1,2) and being anonymous (1.8-SC1). The claimed existence of a privacy technique was important and enough for several participants to decide to share their data. In SC1, where the company does not have to be trusted, anonymity was mentioned more often than in SC2 and SC3, where the reputation of the company mattered for trust. P36 mentioned the following reason for deciding to share data: *"Because the site has a good reputation so I- I think my data is safe"*. Participants had concerns about different kinds of privacy risks that negatively affected their sharing decisions, including the involvement of third parties and data/purpose misuse risks (2.1), identification risks (2.2-SC1,2) and data leakage/security risks (2.3-SC1,2). In addition, incorrect assumptions about DP (2.4-SC1,2), such as being reversible, negatively affected the decision to share. Before being exposed to how DP works, participants had the opportunity to make assumptions about its functionality (see also T3). Although the existence of a protection technique motivates people to share their data, the lack of transparency regarding DP (2.5-SC1,2) negatively affects their decisions to share data. Other hindering factors included not trusting the company (2.6-SC1), not trusting DP to protect privacy (2.7-SC1,2) and a general lack of trust (2.8-SC3) due to the belief of the persistent possibility of data leakage.

T2: Expressed needs for more privacy information. Across all three scenarios, most participants expressed a need for more information related to privacy protection (1) and more specifically related to DP (2) that should be provided in an understandable way (usable transparency) (3). P3 indicated that concrete examples should be given to illustrate the protection and risks of using DP: *"I might want to know what exactly they would protect, like what goes under the protection model and what doesn't [...] the data that they do protect is sleep cycles, but they don't protect the um... information about maybe the steps I'm taking"*. The main needs were for information about the provided privacy functionality (1.1), further specific privacy protection information (1.2), data storage information (1.3), whether sensitive data is processed and with whom it is shared (1.4), information about anonymity/re-identifiability when sharing data (1.5) and information about protection against breaches and risks (1.6). Furthermore, the need for more DP-related information (2) was also expressed, including information on how DP works (2.1), how DP uses and protects data (2.2), the accuracy of personal data that the company receives (2.3-SC1) and information on how trustworthy DP is (2.4-SC3).

T3: Expectation of claimed protection (data access). Our results show that the mere claim that DP protects data without further information on how it works can lead to different assumptions about DP (1) and its privacy features. It can also lead to varied expectations and interpretations regarding access to actual (raw) data (2) by different entities involved in data analysis. Such assumptions and expectations may prevent

users from sharing their data if they incorrectly assume that a specific entity (e.g. the health company in SC1) gets access to their data as they disclose them. DP has been associated with anonymization/pseudonymization (1.1) or with encryption (1.2-SC1,2). Several participants still perceive the risks of identification or data leakage/security risks (1.3-SC1,2) even with DP in place, and/or (incorrectly) think that DP is reversible (1.4-SC2) or assume that analysing data requires access to actual raw data (1.5-SC2,3) or simply associate DP with lower accuracy of data (1.6-SC1).

Assumptions about DP (1) played a significant role in participants' perceptions and expectations of the claimed protection and access to data by different entities. In all scenarios, participants who associated DP with pseudonymization expected that the raw data would be shared with different entities, depending on the context. For example, P2 stated that when sharing with the health company: *"I just assumed that some more personal information would be anonymous, and the rest would be like the raw data"*. In SC2 and SC3, participants assumed that medical researchers and the IBA needed access to raw data to analyse data, which is a false assumption. Likewise, in SC1, doubts about where the protection technique comes into force and the fact that the app is provided by and belongs to the company contributed to users' confusion and wrong incorrect assumption about access to raw data.

T4: Expressed trust factors of DP protecting data. In all scenarios, transparency of DP (1), transparency of data processing types and purposes (2) and good reputation of the company and its history of securing data (3) appeared as factors affecting users' trust in DP to protect their privacy. Trust factors also include being legally (GDPR) compliant (4-SC1,2), having unlinkability features (5-SC1,2), the existence of different privacy assurances and guarantees (6-SC2,3), trusting the company (7-SC1,3), having accountability measures in place (8-SC2) and being a standardized technique (9-SC3). Interestingly, although for the local model (SC1) the health company does not have to be trusted, the trustworthiness of the company appeared as a trust factor. P9 elaborates: *"if I see that the company itself has been trustworthy for several years and has not had major controversy with previous products"*. In addition, incorrect assumptions about DP impacted user trust (10-SC1,2). This included the assumption of DP being reversible, which negatively impacted trust, and associating it with encryption, which positively impacted trust.

T6: Varied impact of DP descriptions on decisions to share. The exposure to metaphoric descriptions of how DP works had a varied impact on the participants' willingness to share their data. The metaphoric DP descriptions either supported/increased the willingness to share (1) or decreased the willingness to share (2). Some participants indicated that privacy concerns are not critical for the decision to share. For example, P7 stated: *"considering I agreed earlier on my data to be shared, I don't think that would be that much of a problem but this would be at the back of my mind"*. There were four participants in SC1, six in SC2, and five in SC3 who decided to share their data and persisted in sharing after exposure to the related metaphor. A few participants decided to share, contrary to their previous decisions, or became more inclined to do so (three in SC1, one in SC2 and two in SC3). Trust in having privacy protection and safety due to DP (1.1), the existence of distortion for privacy protection (1.2), transparency of DP (1.3-SC1,3), trusting the company receiving the data (1.4-SC2), the type of data requested (not perceived as sensitive) (1.5-SC1) and perceived common good benefits of sharing (1.6-SC2) were the factors that supported/increased willingness to share data. Interestingly, misconceptions about DP can also have a positive impact on data sharing (1.7-SC2). The perception of aggregation being secure enough for privacy protection increased the willingness to share in SC2. For example, P11 stated that: *"the first image with no noise is a mixture of the selfies [...] there is some sort of privacy cause it's not my actual picture."* Trading accuracy for privacy (2.1) and the type of data requested (2.2) were the factors in all scenarios that negatively impacted the willingness to share. Participants were mostly not happy to share the type of data they considered very personal. Many voiced the need for more information (2.3-SC1,2) or concerns about the risk of identification (2.4-SC1,2), which were other factors that decreased their willingness to share. Further, misconceptions about DP once again negatively impacted users' perception of its privacy protection (2.5-SC1). For example, after being exposed to the metaphor, P7 stated: *"I cannot guarantee about the privacy which I'm letting it go [...] I mean if I had some noise it's already blurred, but there are many ways which we can, you know, remove the noises."*

T7: Perceptions of information provided/missing. Most participants (eight in SC1, eight in SC2 and five in SC3) perceived the metaphors as easy to understand. In SC3, participants expressed confusion about the model and distortion. They desired more information about what distortion is, how it happens and what its role is in privacy protection and more concrete examples. In all scenarios, most participants expressed interest in receiving more detailed information on how DP works but in simple and clear language. For all three metaphors, people thought there was a lack of information on how distorted/perturbed data can be useful for the analysis and wanted to know if they would have control over the level of distortion. For the noisy picture metaphor, they specifically wanted to know if the process was reversible and thought the levels of accuracy/privacy shown needed elaboration. Participants also suggested some improvements for the spinner. Some indicated that the "YES/NO" on the first spinner was confusing and suggested replacing it. P4 stated: *"YES/NO you're not sure what they're talking about...that can maybe be mistaken as yes or no question"*. In SC1, most participants believed the noisy picture was easier to understand compared to the spinner metaphor; it was appreciated because of its brevity, clarity, simplicity and graphical visualization.

T8: Expressed trust factors (post-explanation). Most participants stated they would generally trust DP to protect their privacy. Transparency of DP (1), type of data/purposes of processing (2), accuracy (accurate results) (3) and understanding of protection provided by DP (4) were the common trust factors in all scenarios. Having control of the distortion level (5-SC1,2), a balanced trade-off (6-SC2) and aggregated data (7-SC3) were also factors indicated to enhance users' trust. Misconceptions about DP were reported to negatively impact users' trust (8-SC2). Many shared the misconception of DP being reversible, which led to distrusting DP. P16 stated: *"I don't trust this because it's very easy to reverse it...it can be made by humans so we can reverse it"* and P6 stated: *"pixels themselves are related to the maths and how the math ... aids the encryption and I'd be worried if it's done by maths can the process be reversed"*.

Perceptions of privacy features of DP and the extent of the suitability of metaphors: Themes T5 and T9–T12 relate to RQ2 and RQ3, as they specifically reveal users' perceptions about the claimed data protection of DP and their understanding of its different privacy features implied or conveyed by our metaphors. In sum, participants correctly perceived that perturbation leads to privacy protection. They also understood, to varying degrees in all scenarios, that perturbation protects against identifiability and provides plausible deniability. However, in all scenarios most of the participants understood the trade-off between accuracy and privacy protection. An analysis of users' perceptions of privacy features of DP revealed several misconceptions, including reversibility of the process (e.g. data distortion) and the perception of DP as encryption. People also had varied perceptions about protection against adversaries with auxiliary information, preferences for the level of distortion and acceptance of and perceptions about remaining risks across all scenarios.

Table 2, which is an updated version of Table 1 based on the themes relating to RQ3, summarizes the extent of the suitability of our metaphors. Y in the table implies that the feature was understood by the majority of participants (80% or more), while N means that the feature was not understood by most of them (20% or less). P shows diversity in understanding, that is, an indication that the feature was perceived by some of the participants. P* means that although the auxiliary information was perceived to be of no help for re-identification by some participants, the reasons behind it were related to the misconception that aggregation would sufficiently protect their privacy.

Table 2: Features of functionality list understood (or not) by data subjects: Yes (Y), No (N), Partially (P)

Metaphor\feature	F1	F3	F4	F5	F6	F7	F8	Context
Spinner	Y	Y	Y	Y	Y	Y	P	Scenario 1
Noisy single picture	P	Y	Y	Y	Y	N	P	Scenario 1
Noisy picture - combined	P	Y	Y	P	Y	P*	P	Scenario 2
Distorted brain model	P	Y	Y	P	Y	P*	P	Scenario 3

T5: Perceptions of claimed protection of DP. Analysing users' perceptions of claimed privacy protection that they assumed was conveyed by the metaphors revealed their misconceptions of DP (1) and their perception of claimed protection by distortion (2). The only common misconception among all scenarios was the perception of DP (noise addition/perturbation) being reversible (1.1). However, in SC1 the reversibility of DP was triggered by the noisy picture metaphor and not by the spinner metaphor. Other common misconceptions, at least between two of the scenarios, include the perception of DP enabling selective disclosure (1.2-SC1,2), the perception of perturbation on individual data records instead of on the aggregate level in SC2 or on the model in SC3 (1.3- SC2,3). Further, there was the perception that aggregation provides enough privacy (1.4-SC2,3). For example, P15 stated: *"I believe that the picture is safe enough because it is a combination and it's not linked to any specific person"*. Some participants believed that distortion would selectively add noise to parts of data or exclude sensitive parts of data and share the rest; for example, P14 stated: *"But since they can't hide everything using this system some of my other data probably, which are not this important, can be probably leaked"*. In SC1, based on the noisy picture metaphor, the description was taken literally (1.5-SC1) and led to the perception of distortion as pixelation of data, or as P9 expressed it: *"I think they also try to either blur or in this case the classic mosaic censorship"*. Further, the pixelated picture metaphor led to the perception of DP as encryption (1.6-SC1). In SC2, it was assumed that how DP works was a secret, which led to the misconception that knowledge of DP by someone could reveal information about individuals (1.7-SC2) if that person accessed differentially private results of analysis. For example, P12 stated: *"Because they know the algorithms and the mathematical equation that are needed to get this level of distortion. They could reverse it"*.

Almost all participants in SC1 and half of the participants in SC2 and SC3 perceived that perturbation protects privacy (2.1). However, in SC1 participants' opinions varied regarding the metaphor that better conveyed that distortion protects privacy. While almost half of the participants believed that the noisy picture better showed the amount of distortion and how it protected privacy, two believed that the spinner metaphor better communicated the unidentifiability feature.

Further, distortion was believed to protect against identifiability or to provide plausible deniability (2.2) to a varying degree in all scenarios. While in SC1 the majority understood it well, in SC2 and SC3 few perceived it correctly. However, using the example of having a unique feature in a population resulted in helping participants (almost all in SC2) to perceive the need for distortion even when aggregation is in place and that it can protect against identifiability, even with unique features (see also 7). The metaphor in SC3 led to confusion about distortion and privacy protection (2.3-SC3). The brain icon often contributed to participants' confusion and was partly misinterpreted and taken literally as images

of users' brains. People had different perceptions of what a model was and what it meant to distort a model. In SC1, a comparison of opinions on two metaphors revealed there were different perceptions on the level of privacy protection based on the metaphors (2.4-SC1). The spinner was perceived to provide better privacy protection. This was among the reasons why almost half of the participants expressed a preference to be exposed to a system illustrated by the spinner metaphor than to one illustrated by the noisy picture metaphor. Interestingly, the results in SC1 revealed that the perception of distortion (gained from the metaphor) is not easily transferable/applicable to other contexts (2.5-SC1). Although it generally made sense to the participants that distortion could protect privacy, it was hard to understand what distortion was and how it would affect data and its accuracy if we had data types other than pictures or YES/NO questions.

Perceptions about the claimed protection after exposure to the metaphors showed varied perceptions about data access by different actors (3) across all scenarios. Understanding of what the company could access contributed to people's correct perception about data access by different actors in SC1. However, in SC2 and SC3 the misconception of how DP works and confusion about the concept of a model and its distortion resulted in only about half of the participants having a correct perception about data access by different actors.

People also had various perceptions about protection against adversaries with auxiliary information (4) across all scenarios. In SC1, based on the noisy picture metaphor, the auxiliary information was perceived to be helpful for the identification (4.1-SC1). However, based on the spinner metaphor, the auxiliary information was perceived to be of no help for re-identification (4.2) of users, given users could lie in the answers perturbed by the spinner. Almost all participants (9/10) in SC1 believed that no one could distinguish actual and random answers from each other. In SC2 and SC3, auxiliary information was mostly perceived to be of no help for re-identification. However, the reason for this perception was the misconception that aggregation would sufficiently protect their privacy and no one with or without extra information about users could identify them.

T9: Perceptions of the accuracy-privacy trade-off. There were various perceptions about the accuracy-privacy trade-off of DP (1) among participants in all three scenarios. Most participants in all scenarios understood the trade-off. In SC1, everyone understood the trade-off for the noisy picture metaphor, and the majority stated that the trade-off is better conveyed by the noisy picture than by the spinner; that is, it shows a clearer progression of noise and its effects on accuracy. However, problems in understanding different terminologies or trade-off elements (2) were reported, which contributed to the misunderstanding of the trade-offs. There were different perceived consequences of trade-offs (3) among the participants in all three scenarios. Several consequences of a lack of accuracy regarding the expense of privacy protection were perceived, including misguided or inaccurate information (3.1-SC1, SC2, SC3), service dissatisfaction (3.2- SC1,3), unreliable recommendations (3.3- SC1,3), application uselessness (3.4- SC1,3) and trust concerns (3.5-SC3). In addition, it was noted for SC1 that the context matters when it comes to trade-offs. For example, P4 stated: *".. because this is health issue so it's not always good to share the wrong information"*. Furthermore, in SC3 it was noted that distortion in long term could lead to false results and would provide no benefits.

T10: Preferences about distortion levels. The general preferences about distortion levels varied across the scenarios. In SC1, participants' preferences regarding the noisy picture varied from no noise to high noise. However, there was a consensus in SC1 regarding the spinner picture; all of the participants preferred the spinner with less probability of revealing true responses. In SC2, four participants indicated that a balance between privacy and utility is important. For example, regarding distortion preferences P12 stated: *"in the medium distortion I think there is the perfect balance"*. Likewise, in SC3 five participants indicated a preference for a medium level of distortion to balance privacy against utility. In addition, it was indicated in SC1 that the level of perturbation/distortion depends on context (i.e. health) and the amount of data to be shared.

T11: Varied acceptance/perceptions of remaining risks. There were five, six and five responses in SC1, SC2 and SC3, respectively concerning the remaining risks the participants perceived (1). Many indicated their perceptions of risks were part of their general perception of privacy risks online, such as through hacker attacks or through the possible misuse of the vast amount of personal data collected about people. For example, P9 stated: *"Every single minute of our life.. we are being tracked be it by the Internet browsing history or Google Maps... It's a privacy concern but nothing new"*. However, when it came to accepting the remaining risks, only one in SC1 refused to accept the remaining risks. There were five responses from SC1 and seven responses from each of SC2 and SC3 that indicated the participants would accept the remaining risks (2). There were three, one and three participants, in SC1, SC2 and SC3, respectively, who indicated that they either have no concerns about or no knowledge of any remaining risks (3) and that they trust the mechanism to protect their data. However, most participants across all scenarios indicated that information about the remaining risks is needed for decision making (4).

T12: Users' input/suggestions for DP alternatives. Distortion was described in different ways (1). In all scenarios, several participants described distortion as the change of original data to protect privacy (1.1). Distortion was also described as something that masks/hides data (1.2 - SC1,2) or filters/removes data (1.3-SC3). Nonetheless, how people described the privacy features of DP varied in different scenarios (2). In SC1, all those who were exposed to the spinner

metaphor first and asked to described DP in their own words highlighted the plausible deniability without referring to the privacy-accuracy trade-off (2.1-SC1,2,3). However, most of those who were exposed to the noisy picture metaphor first referred to the trade-off and the effects of distortion on the accuracy of data (2.2-SC1,2). In SC2, half of the participants referred to the trade-off (2.2-SC1,2). The rest just highlighted the privacy protection features of distortion. They confirmed the importance of including the trade-off feature in their description only after being prompted by the moderator. In SC3, most participants highlighted the protection/security that DP provides and did not mention the trade-off (2.1-SC1,2,3). Participants were confused about the meaning of distortion in the context. For example, P29 said: *"They won't send our face but what they are sending?"*. When asked to describe DP, four participants still referred to distortion on individual selfies than distortion on an aggregate level.

The participants' alternatives to the metaphor to describe DP (3) include DP as pseudonymization (3.1) in SC1–SC3, DP as a generalization (e.g. using ranges instead of single data points) (3.2- SC1,2), and DP as encryption or a technique that mixes data in SC1 and SC2 (3.3-SC1,2). Further, participants in SC1 and SC3 suggested text-based metaphors/examples (3.4-SC1,3) to describe distortion and the trade-off between privacy and accuracy. For example, P30 stated: *"That some of the words are... made completely meaningless"*. Analysing users' descriptions and suggested alternatives did not result in any suitable new metaphors for DP.

7 Discussion

Metaphors can influence how people think about a wide range of issues (e.g. [40]), concepts and experiences [23, 37]. However, metaphorical descriptions may come with specific problems. For instance, metaphorical mappings are partial. They highlight some features of a target domain and de-emphasise others [44] or imply features that do not exist [9]. Cognitive, affective and social-pragmatic factors also moderate the power of metaphors [44]. Our interview results showed the plausible suitability of our metaphors, each to a varying degree (see Table 2), to convey the privacy features of DP to lay users. However, at the same time, our study reveals and confirms several challenges that require further attention if we intend to use metaphors:

Privacy-accuracy trade-off in focus. Because the feature of accuracy loss is prominently demonstrated by the metaphors, some participants defined DP as accuracy loss and/or emphasised the accuracy loss characteristic more than the privacy protection features of DP. This also contributed to participants' accuracy loss-related concerns regarding DP and was a factor for not trusting DP.

Earlier work on *differential identifiability* [10, 32] suggests that information on identification risk reduction is of more relevance for policy makers than information on how to approach the trade-off; therefore, it should be in focus when explaining DP in an understandable way. Our interview results confirmed that identification risks are of special interest and are a general concern even for lay users. Therefore, we recommend future research on the effects of DP explanations (in metaphoric or other forms) that emphasise the reduction of identification risks when explaining DP to different groups of users. This is in line with the recommendation of Wu et al. [46] based on a related study on mental models of encryption. They suggest improved risk communication focusing on the *why* in terms of benefits for the user rather than on *how* the technology works. Our metaphors mainly convey how the technology works by showing privacy protection through the addition of noise. In addition to communicating the benefits of reduced identification risks (and thus emphasising the "why"), users should be guided regarding adequate identification risks per context and the implications (similar to what is also suggested by [35]).

Conveying the feature of plausible deniability. In contrast to the privacy-accuracy trade-off, other features of DP, such as plausible deniability (F1), were not as clearly conveyed to the participants, with an exception of the spinner metaphor. Plausible deniability can be perceived as a benefit by users for accepting deferentially private data analyses. Therefore, it should preferably be communicated to users in accordance with Wu et al.'s recommendation of focusing on the benefits for users [46]. However, an illustrative example for the noisy picture metaphors (pixelating pictures) provided in a follow-up question (to Q24) during the interviews helped several participants understand that even people with uniquely identifiable features should not stick out in differentially private data releases. The noisy picture metaphors could be improved by directly integrating the following illustrative example as a metaphor extension for SC2: *"One of the pictures shows a person with a unique characteristic (e.g. a spot on the lips), which is still visible in the combined picture, while not visible any longer in the perturbed combined picture"*. This extension helps clarify plausible deniability and also shows that aggregation alone is not sufficiently protective. Therefore, our suggested improvement can also address another common misconception and incorrect threat model concerning statistical inference attacks that several participants had for SC2. Previously, Wu et al. [46] recommended explaining the strength of a technology in terms of the capabilities of likely attackers; our proposed improvement follows this recommendation.

Misconceptions based on digital world analogies. Misconceptions about DP are likely triggered by participants' knowledge of security technologies that they are familiar with and that they relate to DP by assuming that DP would have the same features. For instance, the noisy picture metaphor based on pixelation could be related to encryption and could lead to the assumption that DP is reversible, a misconception that largely appeared for the noisy picture and brain metaphors in

all scenarios but was not observed for the spinner metaphor. Similarly, two of the participants (out of four) who were familiar with VPNs and their feature of hiding IP addresses perceived DP as selectively hiding data ("*black out*", "*filter out*" data), and one participant who heard of firewalls understood DP as a means of access control. Similar issues with digital world analogies that are made and that may impact the users' mental models of new privacy technologies they are unfamiliar with were observed earlier [7, 8]. Hence, besides considering real-world analogies, DP metaphors should address the challenge of catering for digital world analogies that users may make.

Usable transparency: challenges and possible solutions. Transparency about various aspects related to DP was named a trust factor by participants, who demanded information in a clear and easy-to-digest form. However, not all the aspects of interest, including all essential privacy features of DP, can be conveyed well by a single metaphor. In addition, interest in transparency of DP may vary significantly among different people. While about half of our participants stated their interest in how DP works, others were only interested in its privacy functionality, remaining risks and consequences. Further, our results showed that individuals had varied and not always correct perceptions of the different privacy features of DP (e.g. F5, F7, F8). In addition, our findings confirmed the problem that individuals lack clear and correct mental models of threats, which was also highlighted by [38] in a study on metaphors for E2E encryption.

Therefore, we suggest complementing metaphor illustrations with additional information when suitable. The additional information should highlight important aspects not sufficiently conveyed by the metaphor and should allow users to easily access additional information of their choice (e.g. by using multi-layered policy statements with links to subpages with various information and varying details on DP). In conformance with the recommendation by [38], future work should focus on finding information and complementing metaphoric illustrations that can change mental models and correct persistent misconceptions that individuals commonly have.

Metaphorical explanations: a quandary. Finally, our study also demonstrated and confirmed that metaphoric explanations inherently suffer from several shortcomings that we need to consider and counteract when we use metaphors to explain privacy technologies to users. Complementing metaphors with suitable additional information, as suggested above, can be one way to counteract these shortcomings.

Problems to abstract: Users might either take metaphors literally or have problems applying the explained features to another context. For instance, our study revealed that the noisy picture metaphor for distortion was generally understood for pictures as data types. However, when asked to apply the concept of distortion to numbers, several participants literally applied it by hiding/blurring numbers.

Different perceptions of the level of privacy protection across metaphors: Two metaphors for the same concept may result in different perceptions of the level of privacy protection. Half of our participants in SC1 preferred to be exposed to the spinner metaphor because they assumed it provided a better privacy-accuracy trade-off, although almost all believed that this trade-off was easier to understand with the noisy picture. The diverse levels of abstractions of the underlying system as a result of using different metaphors impose the risk of different (inaccurate) perceptions of privacy protection.

Conceptual baggage: Our interviews confirmed that metaphors may convey negative or positive features that the system does not have. Such features, if positive, may create an incorrect sense of privacy protection or, if negative, may affect trust and data sharing decisions. For example, our interview results revealed that a noisy picture metaphor conveyed that people with auxiliary information could identify users. Our results likewise revealed that adding noise to pictures could have resulted in the perception that this process was reversible. This conceptual baggage of the noisy picture metaphor negatively impacted our participants' trust and data sharing.

Limitations. We conducted the interviews online with participants' cameras turned off to preserve their privacy, which could have hindered our observations of their attentiveness. However, all participants appeared to be very engaged and attentive in the interviews. Further, our sample mainly consisted of young educated participants, which could have contributed to their understanding of the privacy technique described. However, it could have also negatively impacted their understanding of metaphors, that is, misconceptions due to associating DP with other familiar techniques.

8 Conclusion

This article presents our investigation of the suitability of metaphors to explain differentially private data analyses to lay users to facilitate their informed decisions. We highlight that there is a high interest in usable transparency of DP and privacy protection in general, with different preferences for various aspects (privacy functionality and/or structural information on how DP works) and levels of detail. Our results showed the plausible suitability of the metaphors presented to explain some privacy features of DP to users. We also discuss the misconceptions that result from the metaphors and the challenges of using them. While some of the issues can be addressed by improving the metaphors, others are rooted in the inherent limitations of metaphors. Further research is needed to address these challenges and investigate the type of information that should be provided to lay users to complement metaphoric illustrations to explain the functionalities of DP and correct common misconceptions.

Acknowledgments

This work was funded by the H2020 Framework of the European Commission under Grant Agreement No. 786767 (PAPAYA project) and by the Swedish Knowledge Foundation (TRUEdig project). The work was also partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

References

[1] Differential privacy - simply explained. https://www.youtube.com/watch?v=gI0wk1CXlsQ. Accessed: 2022-02-16.

[2] Differential privacy, an easy case. https://accuracyandprivacy.substack.com/. Accessed: 2022-02-16.

[3] Explaining differential privacy in 3 levels of difficulty. https://aircloak.com/explaining-differential-privacy/. Accessed: 2022-02-16.

[4] What is differential privacy? https://www.youtube.com/watch?v=-JRURYTfBXQ. Accessed: 2022-02-16.

[5] John M Abowd, Gary L Benedetto, Simson L Garfinkel, Scot A Dahl, Aref N Dajani, Matthew Graham, Michael B Hawes, Vishesh Karwa, Daniel Kifer, Hang Kim, et al. The modernization of statistical disclosure limitation at the US Census Bureau. https://www2.census.gov/cac/sac/meetings/2017-09/statistical-disclosure-limitation.pdf.

[6] Idris Adjerid, Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Sleights of privacy: Framing, disclosures, and the limits of transparency. In *Ninth Symposium on Usable Privacy and Security (SOUPS)*, pages 1–11, 2013.

[7] Ala Sarah Alaqra, Simone Fischer-Hübner, and Erik Framner. Enhancing privacy controls for patients via a selective authentic electronic health record exchange service: Qualitative study of perspectives by medical professionals and patients. *Journal of medical Internet research*, 20(12):e10954, 2018.

[8] Ala Sarah Alaqra, Bridget Kane, and Simone Fischer-Hübner. Machine learning–based analysis of encrypted medical data in the cloud: Qualitative study of expert stakeholders' perspectives. *JMIR human factors*, 8(3):e21810, 2021.

[9] James L. Alty, Roger P. Knott, Ben Anderson, and Michael Smyth. A framework for engineering metaphor at the user interface. *Interacting with computers*, 13(2):301–322, 2000.

[10] Daniel Bernau, Günther Eibl, Philip W Grassal, Hannah Keller, and Florian Kerschbaum. Quantifying identifiability to choose and audit ε in differentially private deep learning. *arXiv preprint arXiv:2103.02913*, 2021.

[11] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

[12] Brooke Bullek, Stephanie Garboski, Darakhshan J. Mir, and Evan M. Peck. Towards understanding differential privacy: When do people trust randomized response technique? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, page 3833–3837. ACM, 2017.

[13] Herbert H Clark. *Using Language*. Cambridge University Press, 1996.

[14] Louise Clark and M. Angela Sasse. Conceptual design reconsidered: The case of the Internet session directory tool. In Harold Thimbleby, Brid O'Conaill, and Peter J. Thomas, editors, *People and Computers XII*, pages 67–84. Springer, 1997.

[15] Aloni Cohen and Kobbi Nissim. Towards formalizing the GDPR's notion of singling out. *Proceedings of the National Academy of Sciences*, 117(15):8344–8352, 2020.

[16] Rachel Cummings and Deven Desai. The role of differential privacy in GDPR compliance. In *FAT'18: Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2018.

[17] Rachel Cummings, Gabriel Kaptchuk, and Elissa M Redmiles. "I need a better description": An investigation into user expectations for differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 3037–3052, 2021.

[18] Albese Demjaha, Jonathan M Spring, Ingolf Becker, Simon Parkin, and M Angela Sasse. Metaphors considered harmful? an exploratory study of the effectiveness of functional metaphors for end-to-end encryption. In *Proc. USEC*, volume 2018, 2018.

[19] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In *Advances in Neural Information Processing Systems*, pages 3571–3580, 2017.

[20] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming*, pages 1–12. Springer, 2006.

[21] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography*, pages 265–284, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[22] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

[23] Kristen C Elmore and Myra Luna-Lucero. Light bulbs or seeds? how metaphors for ideas influence judgments about genius. *Social Psychological and Personality Science*, 8(2):200–208, 2017.

[24] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, pages 1054–1067, 2014.

[25] Marco Gaboardi, James Honaker, Gary King, Jack Murtagh, Kobbi Nissim, Jonathan Ullman, and Salil Vadhan. PSI (Ψ): A private data sharing interface. *arXiv preprint arXiv:1609.04340*, 2016.

[26] Michael Hay, Ashwin Machanavajjhala, Gerome Miklau, Yan Chen, Dan Zhang, and George Bissias. Exploring privacy-accuracy tradeoffs using DPComp. In *Proceedings of the 2016 International Conference on Management of Data*, pages 2101–2104, 2016.

[27] Mark F St John, Grit Denker, Peeter Laud, Karsten Martiny, Alisa Pankova, and Dusko Pavlovic. Decision support for sharing data using differential privacy. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pages 26–35. IEEE, 2021.

[28] Noah Johnson, Joseph P Near, Joseph M Hellerstein, and Dawn Song. Chorus: Differential privacy via query rewriting. *arXiv preprint arXiv:1809.07750*, 2018.

[29] Farzaneh Karegar and Simone Fischer-Hübner. Vision: A noisy picture or a picker wheel to spin? exploring suitable metaphors for differentially private data analyses. In *European Symposium on Usable Security 2021*, EuroUSEC '21, page 29–35, New York, NY, USA, 2021. ACM.

[30] Krishnaram Kenthapadi and Thanh TL Tran. Pripearl: A framework for privacy-preserving analytics and reporting at Linkedin. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 2183–2191, 2018.

[31] Jong Wook Kim, Kennedy Edemacu, Jong Seon Kim, Yon Dohn Chung, and Beakcheol Jang. A survey of differential privacy-based techniques and their applicability to location-based services. *Computers & Security*, 111:102464, 2021.

[32] Jaewoo Lee and Chris Clifton. Differential identifiability. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, page 1041–1049, New York, NY, USA, 2012. ACM.

[33] Meng Li, Liehuang Zhu, Zijian Zhang, and Rixin Xu. Achieving differential privacy of trajectory data publishing in participatory sensing. *Information Sciences*, 400:1–13, 2017.

[34] Vivian Genaro Motti and Kelly Caine. Towards a visual vocabulary for privacy concepts. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 60, pages 1078–1082. SAGE Publications Sage CA: Los Angeles, CA, 2016.

[35] Priyanka Nanayakkara, Johes Bater, Xi He, Jessica Hullman, and Jennie Rogers. Visualizing privacy-utility trade-offs in differentially private data releases. *arXiv preprint arXiv:2201.05964*, 2022.

[36] Daniel L Oberski and Frauke Kreuter. Differential privacy and social science: An urgent puzzle. *Harvard Data Science Review: HDSR*, 2(1):1–21, 2020.

[37] Shani Robins and Richard E Mayer. The metaphor framing effect: Metaphorical reasoning about text-based dilemmas. *Discourse Processes*, 30(1):57–86, 2000.

[38] Leonie Schaewitz, David Lakotta, M Angela Sasse, and Nikol Rummel. Peeking into the black box: Towards understanding user understanding of E2EE. In *European Symposium on Usable Security 2021*, pages 129–140, 2021.

[39] Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. A design space for effective privacy notices. In *Eleventh Symposium on Usable Privacy and Security (SOUPS)*, pages 1–17, 2015.

[40] Aaron M Scherer, Laura D Scherer, and Angela Fagerlin. Getting ahead of illness: Using metaphors to influence medical decision making. *Medical Decision Making*, 35(1):37–45, 2015.

[41] Awanthika Senarath, Nalin AG Arachchilage, and Jill Slay. Designing privacy for you: A practical approach for user-centric privacy. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*, pages 739–752. Springer, 2017.

[42] ADP Team et al. Learning with privacy at scale. *Apple Mach. Learn. J*, 1(9), 2017.

[43] Pratiksha Thaker, Mihai Budiu, Parikshit Gopalan, Udi Wieder, and Matei Zaharia. Overlook: Differentially private exploratory visualization for big data. *arXiv preprint arXiv:2006.12018*, 2020.

[44] Paul H Thibodeau, Teenie Matlock, and Stephen J Flusberg. The role of metaphor in communication and thought. *Language and Linguistics Compass*, 13(5):e12327, 2019.

[45] Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

[46] Justin Wu and Daniel Zappala. When is a tree really a truck? exploring mental models of encryption. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS)*, pages 395–409, 2018.

[47] Aiping Xiong, Tianhao Wang, Ninghui Li, and Somesh Jha. Towards effective differential privacy communication for users' data sharing decision and comprehension. In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 392–410. IEEE, 2020.

[48] Yixin Zou, Shawn Danino, Kaiwen Sun, and Florian Schaub. You 'might' be affected: An empirical analysis of readability and usability issues in data breach notifications. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

A Data analysis scenarios - Figures

Figure 3 shows the three data analysis scenarios in our study.

B Example of scenario description

Scenario 1. "The app notifies its users, including Alex, that it is possible to receive supportive recommendations to help them cope with stressful conditions if they want and agree.

To do so, the health company needs to: a) receive stress-related information from different users. Stress-related information may include, for example, users' responses to daily questions about their moods or users' selfie pictures on different occasions when they feel stressed or not. b) Combine and analyse the information it collects from different users to gain insights into stressful conditions and provide remedies and assistance to cope with stress.

In this scenario, Alex trusts the wearable device and her phone but not the health company. Therefore, the information the health company receives from users through the app can negatively affect Alex's (and other users') privacy. Thus there is a privacy problem. The health company may learn about Alex's stress problems and stressful situations.

To protect users' privacy and mitigate the privacy problem, the app applies a privacy mechanism on Alex's (or any other user's) input data before the personal data leaves Alex's (or any other user's) device. This satisfies so-called differential privacy, a formal notion of privacy that provides provable privacy assurances. To a certain extent, this differentially private mechanism prevents leakage of Alex's (or any other user's) actual stress records."

C Metaphor descriptions

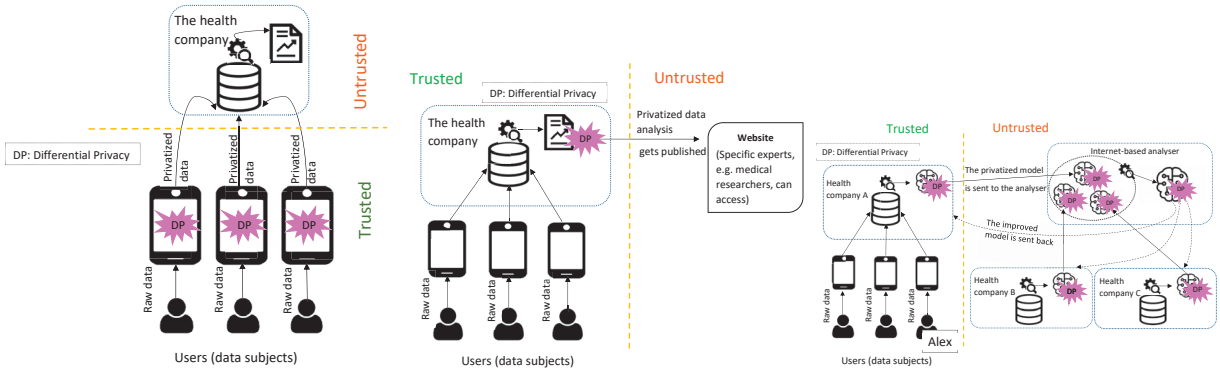
The metaphors were accompanied by descriptions that the moderator read to the interviewees.

Description of metaphor in Figure 2c: "Now imagine, in the scenario described, that the health app requests its users including you (as Alex) to answer some sensitive YES/NO questions about their stress conditions. The health company then receives the responses and can use the responses to analyse, for example, the proportion of users who say YES or NO to each question. The health app will protect users' privacy by using a differentially private mechanism, as depicted in this figure.

Your health app on your phone uses a spinner wheel to perturb (change) your responses to the questions with a controlled and known probability based on the underlying mechanism before sharing them with your health company. The app spins the wheel. If it lands on YES, your true response will be revealed. If it lands on NO, it will spin the wheel again. If it lands on YES the second time, it will reveal YES and if it lands on NO, it will reveal NO regardless of your true responses. The purpose of perturbing your responses is to assure your privacy. The mechanism guarantees that what the health company can infer about your true responses is limited and negligible. You can deny, to a certain extent defined by the mechanism, if a given YES or NO response is your true response.

This figure is not a precise representation of the underlying mathematical mechanism that perturbs users' data; it is just a simplified example of what perturbation means. Note that the outcome of the spinner, whether it lands on YES or NO, remains hidden from the health company. Although the health app deliberately perturbs its users' responses, the health company can still benefit from the collected responses to infer the proportion of users who said YES or NO to each question."

Description of metaphor in Figure 2a: "Now imagine, in the scenario described that the health app requests its users including you (as Alex) to share their selfies with the health company. The health company can then use the selfies and analyse, for example, the most common facial expressions of users when they are stressed. The health app will protect users' privacy by using a differentially private mechanism, as depicted in this figure. First, you share your selfie with the



(a) Scenario 1 - local model (b) Scenario 2 - central model (c) Scenario 3 - central model

Figure 3: Data analysis scenarios.

app. Then your health app perturbs (carefully distorts) the details of your picture by adding a specific amount of noise to it (e.g. a medium amount of noise) based on the underlying mechanism before sharing it with your health company. The purpose of distorting your selfie is to assure your privacy. The mechanism guarantees that what the health company can infer about your true selfie is limited. You can deny, to a certain extent defined by the mechanism, if a shared selife is your true selfie.

This figure is not a precise representation of the underlying mathematical mechanism that perturbs (distorts) users' data; it is just a simplified example of what perturbation means.

Although the health app deliberately adds noise to its users' selfie pictures, the health company can still benefit from the collected selfies to infer the most common facial expressions of users when they are stressed."

Description of metaphor in Figure 2b: "Now imagine your health company wants to analyse what the common lip expression is when people do (not) feel stressed and it requests users, including you as Alex, to share their selfies so it can combine and analyse them. The health company will protect its users' privacy by using a differentially private mechanism to analyse their data, as depicted in this figure.

Before revealing the common lip expression, the health company perturbs (carefully distort) the details of the common lip expression based on the underlying mechanism by adding a specific amount of noise to it (e.g. a medium amount of noise). The purpose of distorting the common lip expression is to assure users' privacy by limiting the effects of each individual's selfie on the analysis results, i.e. the common lip expression. Therefore, the mechanism guarantees that the likelihood of privacy harm users may face by being identified as a result of sharing their selfies and having their selfies analysed together with those of other users is limited and insignificant.

This figure is not a precise representation of the underlying mathematical mechanism that perturbs the results of data analysis; it is just a simplified example of what perturbation means.

Note that although the health company deliberately perturbs the results of data analysis, in this case the common lip expression derived, the distorted results of the analysis can still be useful for the receivers, for example, the health researchers."

Description of metaphor in Figure 2d: "Now imagine your health company wants to create a model that can recognize a user's emotion from his/her facial expression. Again, note that you can think of a model as an artificial brain that learns from its inputs. In other words, a model can be trained based on the characteristics of its inputs to do a special thing. The health company requests its users, including Alex, to share their selfies and then uses the selfies to train a model so the model can recognize emotions based on facial expressions. For example, the model can predict if a user is very happy, sad, confused, stressed, furious, etc. In this figure, you see a trained model based on users' selfies. Now if the trained model receives a user's selfie as its input, it can predict the user's emotion.

As mentioned previously, the health company protects the users' privacy by using a differentially private mechanism to train the model, as depicted in this figure. Before sharing its locally trained model with the analyser, the health company perturbs (carefully distorts) the trained model based on the underlying mechanism. This means that the health company distorts the information the model has learned from selfies randomly but in a controlled way, for example, using the medium level of distortion. The purpose of distorting the trained model is to assure users' privacy by limiting the effects of each individual's selfie on what the model has learned from the selfies. Therefore, the mechanism guarantees that the likelihood of privacy harm users may face by being identified as a result of uploading their selfies and having their selfies used with other selfies to train the model is limited and insignificant.

This figure is not a precise representation of the underlying mechanism that distorts a trained model; it is only a simple example of what distortion means.

Although each health company deliberately distorts its

trained model, the final model is better than each of the locally trained models at recognizing the emotions. The final model made by the analyser is also a distorted model that protects users' privacy."

D Interview guide

This is our interview guide for the first scenario (depicted in Figure 3a). As mentioned in Section 4.2, half of the interviewees assigned to this scenario were first exposed to the spinner metaphor (Figure 2c) and then the metaphor of the noisy picture (Figure 2a). The other half were exposed to the same metaphors but in the reverse order. The interview questions for other scenarios were adjusted to fit the context.

Welcome and introduction. Participants are welcomed and instructed about the setup of Zoom and turning off their video. An introduction to the study, the goal and the different parts of the interview are provided. The consent form is given to the interviewees, and once they agree the session starts and the recording commences.

Scenario introduction and expectations discussion. The moderator first describes the persona and then describes one of the scenarios to the interviewee by showing the related figure (see Figure 3) and reading the related description provided in C. The following questions are asked.

Q1. Have you heard about any privacy protection techniques (techniques to guarantee users' privacy and to improve it)? Have you ever heard about differential privacy?

Q2. In what context did you hear about it?

Q3. Do you know what differential privacy is? Can you explain it in your own words?

Participants are then told to pretend they are Alex, who is using a wearable device, and have received the notification in the scenario while answering the following questions.

Q4. Would you agree to share your data to be analysed in the way described? What factors did play a role in the decision for Alex?

Q5. How did the differential privacy mechanism play a role in your decision? Would it matter if another mechanism were used to protect your privacy instead of differential privacy?

Q6. What should have been different so you would agree?

Q7. What do you want to know about the mechanism applied (the differentially private mechanism) to protect your privacy? What information would you like to be added to the scenario?

Q8. What would be the benefits for you if you agreed? What would be the risks for you?

Q9. In this scenario, from whom do you expect your actual stress-related data to be hidden? (follow-up: could your health app see your actual stress-related data? What about your health company?)

Q10. In this scenario, it was mentioned that your privacy is protected against potential privacy risks using a specific mechanism. What factors do play a role for you to trust this mechanism to protect your data?

Metaphor introduction and perceptions gauging. The moderator shows a specific metaphor depending on the data analysis scenario for the interview and reads the description of the metaphor to the interviewee. The descriptions of the metaphors are provided in Appendix C. Participants are told to consider the description of differential privacy and the scenario when answering the following questions.

Q12. Would you change the decision you made on behalf of Alex in the previous step after receiving more information about differential privacy? Why?

Q13. In general, do you think that receiving information about the underlying privacy techniques a system uses would help you decide to use a system? How (in what way) could it be helpful?

Q14. Is the description of differential privacy understandable and easy to grasp for you? What is not clearly described or is missed in the description? How can the description be improved?

Q15. Is there any information that is surprising to you—you did not expect? Please elaborate.

Q16. Would you like to know more about the technical and mathematical details of the underlying differentially private mechanism? Why?

Q17. The mechanism perturbs (changes) your data in a controlled way. Can you explain in your own words what it means to perturb your responses? How does data perturbation protect your privacy?

Q18. How would your privacy be better protected—using a spinner with a bigger area for YES or a smaller area for YES? What happens if the area for YES is zero and is 100% for NO?

Q19. Which of the spinners do you prefer to be used to perturb your data? Why?

Q20. Can you explain whether there is a trade-off between the accuracy of the data analysis results and the privacy of your data?

Q21. How would you as Alex be affected if the data analysis results are not accurate? Would you rely on the recommendations the app gives you to cope with stress? Why?

Q22. Imagine that as Alex you agreed to share your stress-related information in the scenario. Which of the following entities would be able to see your (Alex's) actual stress-related data? (Why do you think so?): a. Hackers who access the database of the health company. b. People who know how the differentially private mechanisms work if they access the perturbed data.

Q23. Would the health company be able to prove that a YES answer is your true answer? Would a close friend (if she/he gets access to your perturbed answers) be able to prove that a YES answer is your true answer?

Q24. Imagine you agreed and that the health company analysed the proportion of users who said YES or NO to

each question, based on the perturbed responses it received. If you did not agree to share your responses, how would it affect the proportions of YES/NO responses to a question that is calculated by the health company based on the perturbed responses it receives? (follow-up: Do you think the proportion of users who said YES/NO to each question greatly depend on your decision to share your responses? What if the responses were not perturbed?) (Follow-up (SC2): Imagine you as Alex have a feature that no other user has. For example, you have a dark spot on your lip. Therefore, the common lip expression derived will include a dark spot as well. How would it change if your selfie was not included? Now imagine that we distort the common lip expression so that the dark spot is not shown. How would the distorted lip expression change if you did not agree to share your selfie?)

Q25. How would you describe the likelihood of remaining privacy risks? Would you accept the remaining risks? Would more information about the remaining risks be of your use in deciding to share your data or not?

Q26. Now that you know more about differential privacy, would you trust this method in general to protect your privacy? Why? (If the answer is NO:) What are your concerns in this regard?

Q27. How would you describe differential privacy to someone who does not know about it? Can you think of any alternative description/example of perturbation (data changes) other than the one we used to describe the concept?

Second metaphor introduction and discussion. The moderator reads the second metaphor and then asks:

Q28. Can you describe how your privacy (as Alex) will be protected by perturbing (distorting) your data? How would your privacy be better protected, by adding more noise or by reducing the noise?

Q29. How can data distortion affect the accuracy of data the health company receives?

Q30. Which of these two description better conveys the trade-off between accuracy and privacy? Why?

Q31. Which of these two descriptions is easier for you to understand? Why?

Q32. Which one do you prefer to be exposed to when you want to decide to use a differentially private system? Why?

Q33. Which of these two descriptions better helps you to relate data perturbation (changing your data or distorting your data) to privacy protection? (follow-up: What are the shortcomings of this description? How it can be improved?)

Feedback and thanks. The moderator asks the participants if they have any comments and/or questions. Then the moderator thanks the participants for their participation.

E Themes

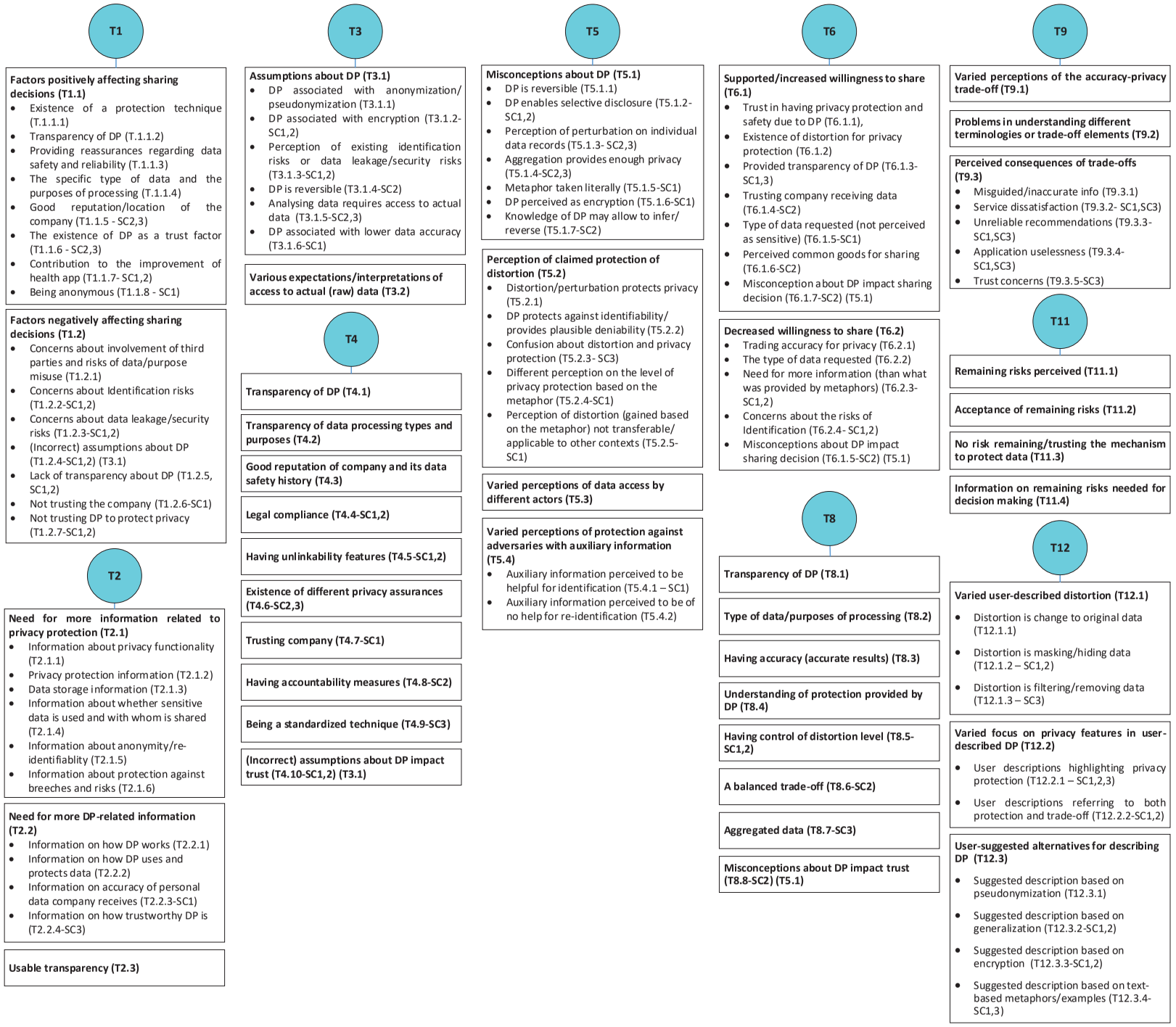


Figure 4: Themes with the sub-themes (note that the themes without sub-themes are not listed.)