Identifying Privacy Personas

Olena Hrynenko
Idiap Research Institute
École Polytechnique Fédérale de Lausanne
olena.hrynenko@idiap.ch

Andrea Cavallaro
Idiap Research Institute
École Polytechnique Fédérale de Lausanne
a.cavallaro@idiap.ch

ABSTRACT

Privacy personas capture the differences in user segments with respect to one's knowledge, behavioural patterns, level of self-efficacy, and perception of the importance of privacy protection. Modelling these differences is essential for appropriately choosing personalised communication about privacy (e.g. to increase literacy) and for defining suitable choices for privacy enhancing technologies (PETs). While various privacy personas have been derived in the literature, they group together people who differ from each other in terms of important attributes such as perceived or desired level of control, and motivation to use PET. To address this lack of granularity and comprehensiveness in describing personas, we propose eight personas that we derive by combining qualitative and quantitative analysis of the responses to an interactive educational questionnaire. We design an analysis pipeline that uses divisive hierarchical clustering and Boschloo's statistical test of homogeneity of proportions to ensure that the elicited clusters differ from each other based on a statistical measure. Additionally, we propose a new measure for calculating distances between questionnaire responses, that accounts for the type of the question (closed- vs open-ended) used to derive traits. We show that the proposed privacy personas statistically differ from each other. We statistically validate the proposed personas and also compare them with personas in the literature, showing that they provide a more granular and comprehensive understanding of user segments, which will allow to better assist users with their privacy needs.

KEYWORDS

privacy, personas, open coding, qualitative and quantitative analysis

1 INTRODUCTION

People differ in their attitudes, perceptions, and expectations towards privacy [40, 41]. Being able to model these differences is a key step for providing personalised support [22]. Personas allow to model user segments [19] based on various attributes such as goals, needs [19], preferences [8], concerns [25], knowledge [2, 28], behaviours [2, 11], or financial wealth [32]. Personas should be as discriminative and realistic as possible [31]. Privacy personas should describe user segments that help design privacy and security tools [11].

Common attributes defining privacy personas are level of concern about privacy ([12, 28]), level of knowledge ([2, 11]), one's behaviours ([2, 11, 30]) and attitudes ([11, 30]) towards privacy. While these attributes for privacy personas are commonly explored, there are a number of limitations. The attributes are often considered in isolation, leading to a loss of information for privacy personas [2, 30]. Examples include using only *level of concern* by Westin [28], or only behaviour and knowledge by Biselli et al. [2].

To overcome these limitations, we propose a new way of modelling privacy personas from the responses to an interactive questionnaire [13]. For modelling our privacy personas we use a mix of qualitative and quantitative approaches to facilitate a comprehensive understanding of the data [10]. Starting from questionnaire responses [13], we perform coding, code manipulation [11] and annotation to form participants' feature vectors. To account for differences in nature between answers to closed- and open-ended questions in a questionnaire, we propose a new dissimilarity measure for evaluating differences between the participants. Besides self-efficacy, PET's efficacy and willingness to use PET, we use a wide set of persona attributes (e.g., privacy protection importance perception, level of knowledge), which allow for a more comprehensive understanding of a persona. We propose a pipeline based on hierarchical clustering that ensures that the elicited personas are statistically different to each other (i.e., Boschloo's test of homogeneity of proportions [3]), rather than ad-hoc approaches previously used in the literature [11, 12, 28].

In summary, our main contributions are:

- The identification of eight privacy personas elicited with an interactive questionnaire [13], which incorporates responses about one's self-efficacy, willingness to use PET, alongside other privacy persona attributes, allowing for a comprehensive persona definition.
- The proposal of a dissimilarity measure that takes the nature of closed and open-ended questions into account for calculating differences between the responses of participants.
- A pipeline for validation of the cluster splits in a hierarchical structure based on statistical significance tests [3].
- The comparison of the identified personas with the privacy personas by Westin [28], Biselli et al. [2] and Dupree et al. [11], showing that our personas have a higher level of granularity, resulting in a more accurate representation of the population, which translates into better-aligned support for personas.

The paper is organised as follows. Section 2 covers previous methods used to define privacy personas. Section 3 describes the dataset we used for personas elicitation. Section 4 presents how we form personas' traits with open coding. Section 5 explains the new dissimilarity measure and how we elicit personas. Section 6 describes our personas, which are then validated in Section 7. Section 9 discusses the limitations of our approach, whereas Section 8 maps our personas with the personas in the literature. Finally, in Section 10 we draw conclusions and describe future work.

2 RELATED WORK

In this section, we discuss qualitative, quantitative, and hybrid methods for identifying privacy personas.

Ref.	Method	Privacy personas	Generation set		Validation set		
			size	demogr.	SD	size	demogr.
[28]	QN	Fundamentalist, Pragmatist, Unconcerned	N/A	N/A	N/A	N/A	N/A
[12]	QL	Fundamentalist, Intense Pragmatist, Relaxed Pragmatist, Marginally Concerned, Cynical Expert	40	65-91 yo British Canadians	N/A	N/A	N/A
[2]	QN	Fundamentalist, Pragmatist, Unconcerned	332	18-75 yo German	no	324	18-75 yo German
[11]	HY	Fundamentalist, Technician, Amateur, Marginally Concerned, Lazy Expert	32	22-35 yo North American	yes	200	18-65 yo American
[30]	QN	Guardian, Pragmatist, Cynic	337	above 14 yo German	N/A	N/A	N/A
Ours	HY	Knowledgeable Optimist, In-control Adopter, In-control Sceptic, Knowledgeable Pessimist, Helpless Protector, Occasional Protector, Adopting Protector, Unconcerned	130	25-35, British	yes	50	25-35, British

Table 1: Related work, methods that were used for eliciting privacy personas, and privacy personas. Key – QN: quantitative, QL: qualitative, HY: hybrid, SD: separate dataset, yo: years old.

Qualitative methods generally extract codes from interview transcripts, and group the codes into themes that represent the main ideas from the data [4, 9]. The goodness measure is *saturation*, namely "the point in data collection when no additional issues or insights emerge from data and all relevant conceptual categories have been identified, explored, and exhausted" [18]. Qualitative methods allow to discover new concepts and build the basis of the persona features, accommodating the emerging concepts. *Quantitative* methods follow a pipeline that includes question construction, data collection, feature vector construction, and personas elicitation. *Questions construction* generates a list of questions to ask the participants. This step is based on the findings of qualitative analysis. The list of questions is refined based on feedback from experts or focus groups (e.g., removing unclear or leading questions), or on statistical methods conducted on preliminary data. Note that the outcome of the questions construction step is a fixed set of closed-ended questions. The *data collection* step consists of reaching out to participants and collecting data. *Feature vector construction* translates the collected responses into the numerical feature vectors. *Personas elicitation* involves using quantitative methods for grouping participants into personas. With *hybrid* methods, the outcome of *questions construction* is a mixture of closed- and open-ended questions, which could later on be enriched with additional questions (e.g., in the case of semi-structured interviews). Including open-ended questions means that to convert a participant's response into a feature vector, further qualitative analysis is required (e.g., code extraction). Methods and corresponding privacy personas are summarised in Tab. 1 and described below.

Privacy personas by Westin [28] are often used as a baseline for the privacy personas, despite being criticised for the lack of predictive power for the behaviours and knowledge [21]. Westin [28] conducted more than 120 studies to explore the privacy concerns of people [35] and identified three privacy personas [28], namely the Fundamentalist, the Unconcerned, and the Pragmatist. The Fundamentalist feels very strongly about privacy matters, and supports new laws that allow for privacy regulations. This persona does not trust companies that ask for personal information, and chooses privacy protection over potential benefits. The Fundamentalist is worried about the use of their information, pessimistic about the future of privacy protection, and feel they have lost a lot of their privacy. The Unconcerned does not feel that their privacy is violated. This persona feels less anxious about how others use information about them. The Unconcerned trusts organisations collecting personal information about them, and the benefits they may receive by revealing their information carry higher weight than the potential privacy harm they might face. The Pragmatist feels strongly about privacy and seeks fair information practices, weighing the benefits against the level of intrusiveness. This persona wants to have the freedom of choice to opt-out. For the Pragmatist, business organisations should earn the trust, rather than have it unconditionally. We will compare our personas with Westin's personas in Section 8.1.

Elueze and Quan-Haase [12] investigated the differences in attitudes and concerns about online privacy in the lives of older adults. They conducted in-person interviews, coded the data and used the categorisation by Westin [28] as a basis for their personas. Based on the key responses to privacy questions, people were assigned to the Fundamentalist, the Pragmatist or the Unconcerned classes. Then based on further analysis of the codes and through analysis of negative cases, the Pragmatist persona was divided into two groups, namely the Intense Pragmatist and the Relaxed Pragmatist. An additional group, the Cynical Expert, was established.

Biselli et al. [2] identified privacy personas using a set of closed-ended questions that measure the knowledge and behaviour of participants. They showed a strong positive correlation between knowledge and privacy behaviour: the more people are aware of the threat, the more cautious they are. Biselli et al. [2] segmented people into three personas, adopting the naming convention of Westin [28]. However, as the authors acknowledged, no formal validation on linking their personas to Westin's has been done. We will compare our personas with Biselli's in Section 8.2.

Dupree et al. [11] measured people's attitudes and behaviour towards security practices. They used closed-ended questions for a

questionnaire and open-ended questions for the semi-structured interviews for their *data collection* step. Because Dupree et al. [11] used the hybrid method, their *feature vector construction* step involved manual coding and annotation. The personas were generated by clustering participants' most discriminative *traits* (participants' features): a *trait* shared by multiple clusters was removed, and the participants were re-clustered. Once the clusters were established, the unique cluster *traits* were used to describe the personas. To position personas in a motivation-knowledge space, they manually annotated the participants' responses with respect to their knowledge (low, medium, high), and motivation[1] (low, medium, high) after the clusters were formed. Dupree et al. [11] then calculated the overall knowledge and motivation per cluster. Dupree's Fundamentalist has high motivation and high knowledge, the Marginally Concerned has low motivation and low knowledge, the Lazy Expert has high knowledge and low motivation, the Struggling Amateur has medium motivation and medium knowledge, and the Technician has high motivation and medium knowledge.

While the traits per cluster were considered unique, there is no information on how common they were within the cluster, e.g., how many participants shared a similar opinion. Hence, we claim that Dupree's knowledge-motivation persona assignment has a higher validity than the clusters' description which comes from the unique traits only. Differently to Dupree et al. [11], we use an idea of statistically significant differences between the clusters and introduce a new dissimilarity measure which treats closed- and open-ended questions differently. We will compare our personas with Dupree's personas in Section 8.3.

Schomakers et al. [30] conducted the interviews to select a set of questions for their follow-up quantitative study for privacy personas generation. They clustered participants based on their level of concern and protective behaviours[2], and identified the Privacy Guardian (the highest level of concern, the highest level of proactive behaviour), the Privacy Cynic (the lowest level of behaviour, moderately high level of concern), and the Privacy Pragmatist (lowest level of concern, moderately high level of protective behaviour).

A key difference with our method is that we use a tightly intertwined mixture of qualitative and quantitative analysis, not separating one from another, whereas Schomakers et al. [30] use the findings of the interviews for developing a closed-ended questionnaire. Since privacy personas trends could change over time [12], our personas are more suitable for allowing to model these changes.

Moreover, we consider the relationship not only between the behaviour and concern attributes but also between a wider set of attributes, which allows for a more granular persona understanding. We will compare our personas with personas by Schomakers et al. [30] in Section 8.4.

[1]Motivation is defined as: "the effort they [people] expend to protect their privacy or security" [11].

[2]The privacy behaviour (PB) and privacy concern (PC) questions were [30]: *I use every option that I know to protect my online privacy (e.g., deleting cookies, anti-virus software)* (PB1); *I specifically search for more options to protect my online privacy* (PB2); *I use the default settings of my devices and applications without changing them. (rev)* (PB3); *I use the default settings of my devices and applications without installing additional software to protect my privacy. (rev)* (PB4); *In general, I am concerned about my privacy when I am using the internet.* (PC1); *With some types of information collected on the internet I do not feel comfortable* (PC2); and *I do not see risks when providing data on the internet. (rev)* (PC3).

3 QUESTIONNAIRE

We use a dataset provided by Ferrarello et al. [13] with privacy stimuli and an educational element. This dataset measured one's change in predisposition towards privacy once they learned more from the interaction. An answer to a question may be a sentence (e.g., "Privacy is the ability to share things online yet keep unwanted attention away.") or a phrase (e.g., "Moderately easy").

3.1 Questions and Domain of Answers

The questionnaire is composed of questions q_j, where $j = 1, \ldots, Q = 19$. Through the questionnaire, participants learnt about the existence of inferences on the images they share online. They were firstly asked a set of questions about privacy protection importance, their preferences, expectations about privacy, and their awareness about inferences on their images ($q_1 - q_{10}$). Then participants were asked to upload the most recent image that they have shared online with their audience. Participants were then informed of the existence of a different party (i.e., a company) having access to data inferred from the images. They were shown a demo of how much information can be extracted from the image. We refer to this process as the first privacy stimulus. After the first privacy stimulus participants were asked how this information extraction makes them feel, if they understand what is happening to their data, and their perceived self-efficacy ($q_{11} - q_{14}$). Then the participants interacted with an image filter that acted as a Privacy Enhancing Technology (PET), a second privacy stimulus. Finally, participants' willingness to use this PET and perceived efficacy of the PET were measured, and the updated view on privacy protection was recorded ($q_{15} - q_{19}$). The educational, interactive element allows to align participants with respect to an otherwise vague concept of privacy [15, 30] as well as to measure participants' self-efficacy, perception of the response efficacy, and motivation to use PET.

The questions (q_*) and the domain of answers to closed-ended questions (a_*) are provided below.

- q_1 First of all, how important do you think it is to protect private information when sharing images online?
  - a_1 very important | important | neutral | not that important | unimportant
- q_2 Image 1: cat (What personal information is it revealing?)
- q_3 Image 2: van (What personal information is it revealing?)
- q_4 Image 3: rollercoaster (What personal information is it revealing?)
- q_5 What could you do to protect the personal information in these images?
- q_6 Now look at the last 10 images you shared online. What pieces of personal information do you find in those images?
- q_7 How easy do you think it is to protect the personal information you identified with what is currently available to you?
  - a_7 extremely easy | mostly easy | moderately easy | slightly easy | not at all easy
- q_8 How often would you like to be able to have control over the personal information that you have identified?
  - a_8 always | often | sometimes | rarely | never
- q_9 Did you know that social media platforms check the images you share for sensitive content?
  - a_9 yes, familiar with how this is done | yes, but don't understand how | no
- q_{10} But they may not always stop there. Did you know that, behind every picture that you share, pieces of personal information can be

found and used to generate a personal profile that is used to control what you see?
a_{10} yes, familiar with how this is done | yes, but don't understand how | no

Privacy stimulus 1

q_{11} How does this make you feel?
q_{12} Do you feel like you understand what happens to your personal information when you share an image online?
a_{12} fully | mostly | moderately | slightly | don't understand
q_{13} What could you do to protect the personal information you selected?
q_{14} Do you think you have control of your personal information when sharing images online?
a_{14} can fully | mostly | moderately | only little | cannot control

Privacy stimulus 2

q_{15} Do you feel like the filter would allow you to protect your privacy?
a_{15} yes | unsure | no
q_{16} If these new filters were available, would you use them?
a_{16} would always | often | sometimes | rarely | never use these filters
q_{17} Why?
q_{18} Think back on everything you learnt about sharing your images online. How important do you think it is to protect the privacy of the pictures you share online?
a_{18} very important | important | neutral | not that important | unimportant
q_{19} Finally, how would you now define the term privacy when related to online images?

3.2 Responses

The dataset by Ferrarello et al. [13] consists of 200 responses to an interactive questionnaire given by participants from the UK and was collected using Prolific [29]. The recruited participants were users of one of the following social media: Facebook, Twitter, Instagram, and Snapchat. The participants had at most 20 submissions on Prolific, and had their income in the range of £10,000-£50,000. The dataset is gender balanced, and the age of the participants ranged from 25 to 35 (mean: 30.2, standard deviation: 2.8). Participants were paid 7.10 GBP for taking part in a questionnaire. The study by Ferrarello et al. [13] was approved by the ethics board.

We represent a participant i, P_i, through their answers to the questionnaire. We used all questionnaire responses available, except four, because they had missing answers. We used 64 responses for coding and 180 for annotation. For generating personas we used $G = 130$ responses that were not previously used for coding. We call these responses the generation set, $\mathscr{G}$. For the validation set, $\mathscr{V}$, we re-used $V = 50$ questionnaire responses that were previously used for coding.

3.3 Biases and Priming

The order of the questions of the questionnaire [13] is important. Questions asked after the first *privacy stimulus* aimed at measuring the effect of being exposed to a personal information inference demo. Similarly, questions after the second *privacy stimulus* aimed at measuring the willingness of a participant to use the introduced PET. Finally, there were questions that measured the same attribute (i.e., privacy protection importance) at the beginning and at the end of the questionnaire to evaluate the effect on a participant of the educational interactive element (i.e. being introduced to privacy-related threats, being offered the PET to protect personal information).

The existence of an interactive element also introduced the bias of the personalised element. One of the biases of the questionnaire is that we only observe the self-reported behaviours and future decisions, which might not correspond to the actual behaviour. For example, the decision to often use filter (q_{16}) might be different to the actual observed behavioural patterns [2].

4 FEATURE VECTOR CONSTRUCTION

The dataset includes information from closed-ended and open-ended questions. For a feature vector construction, we use open coding through Grounded Theory [16] to extract codes, which are sentences or phrases that propagate one idea [4]. We then group codes into traits, which are low-level groupings of codes that are summarised with a phrase or a sentence [11]. To prepare for annotation, we form an affinity diagram [24] of traits. Finally, we perform annotation. Fig. 1 shows the process of code extraction, and trait and affinity diagram formation.

4.1 Code Extraction

We randomly sample 64 questionnaire responses for code extraction from the open-ended questions. This is twice the number of participants on which Dupree et al. [11] have identified their personas. We adopt a random sampling approach to avoid any human bias in the sampling process.

Codes from open-ended questions. The process of extracting codes for open-ended questions consists of independent code extraction followed by code cleaning.

Coders $C1$ and $C2$ extract the codes from the questionnaire responses without consulting each other to avoid introducing bias at this stage. The process involves reading the questionnaire responses and extracting atomic pieces of meaning – phrases that cannot be subdivided into smaller ideas. Here we follow advice on extracting codes by answering the W's and H's questions [16]: who, why, when, what and how, how much, and how long? For example, one of the answers to q_{11} was "*Very uneasy, not entirely surprised. Confirmed why I don't post often on social media*". This answer generated multiple codes: "feels uncomfortable", "feels not surprised", and "claims does not post often".

We then create a list of codes, where each code is unique. The codes provided by each coder were concatenated and further processed. $C1$ removed duplicates of the codes which were caused by slight variability in annotation. For example, codes "high level of knowledge" and "highly knowledgeable" were mapped to the same code. Additionally, due to inevitable human error in the initial coding, the codes that encapsulated two or more distinct ideas were separated into multiple codes. For example, the code "filters serve a dual purpose: nice image + protection" was split into "filters increase the quality of the image" and "filters allow to protect me".

Codes from closed-ended questions. Each answer to closed-ended questions is a code once it is paraphrased. We paraphrase the answer for it to be understood without reading the corresponding question ("I am neutral" answer to q_1 is paraphrased to "initially does not have an opinion on privacy protection"). More examples of the closed-ended codes are in Tab. 2. Coding and paraphrasing

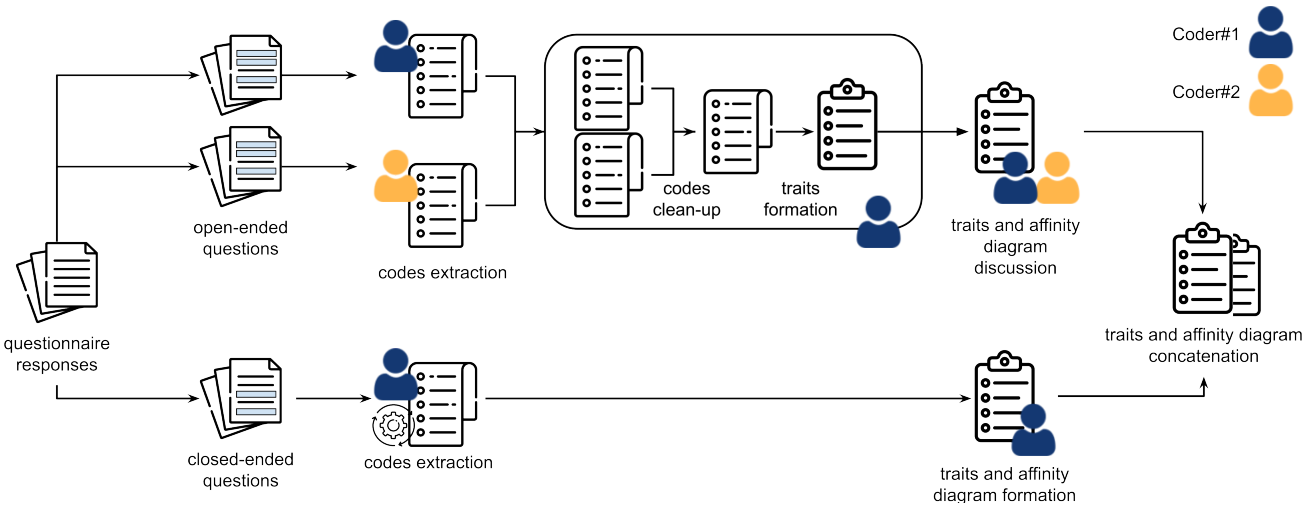

Figure 1: Process of extracting codes and traits formation. We follow different processes for open and closed-ended questions. Coding and traits extraction for open-ended questions consists of code extraction (done by two coders independently); code clean-up, traits formation (done by one coder); traits and affinity diagram discussion (done by two coders). For closed-ended questions the Likert-scale answers were the codes themselves (however coder paraphrased them to full sentences), traits and affinity diagram formation. An affinity diagram helps to organise data which initially seems unstructured [24], allowing to find categories of the traits [11]. At the end of the process, there is a concatenation of the affinity diagram and the extracted traits.

codes from open-ended questions	codes from closed-ended questions
would use filters to avoid relevant info being sold	initially expressed a desire to always be in control of their personal information
I would use filters to avoid being objectified	initially does not have an opinion on privacy protection
would use filters not to share things with strangers	illustrated no awareness of social media being to extract personal information from an image and generate a personal profile on a user
filters "throw off any algorithm"	after being exposed to how companies can extract data from their images they feel that they have moderate control over personal information when sharing images online
filters deteriorate colours	significant mismatch between the perceived and desired level of control over personal information after being exposed to companies being able to extract personal information
filters protect information about me	they would sometimes use the proposed privacy enhancing filter when sharing images online

Table 2: Examples of the codes extracted from closed-ended and open-ended questions.

of closed-ended responses enable Boschloo's test [3] and ensure that codes are understood without additional context.

Additionally, we generate codes that capture the change of privacy protection importance (q_1 and q_{18}), and the desired and perceived level of control mismatches (q_8 and q_{14}) to identify the mismatch between perceived and desired levels of control.

4.2 Traits

Following the approach proposed by Dupree et al. [11], after code extraction, we form traits and build an affinity diagram [24]. We form traits to ensure that when we describe participants, we refer to an idea (i.e., *trait*) rather than an instance of an idea (i.e., *code*). For example, in q_{17} one of the participants expressed that they would not use the proposed PET. They explained their decision by saying that the "keywords were not erased". This answer generated a code "expressed an opinion that keywords can be erased". Here the participant was not aware that a machine learning pipeline that is trained to extract concepts from an image would always produce an output, it is the quality of the output that could be altered. So code "expressed an opinion that keywords can be erased" was mapped to a trait "lacking knowledge in some aspects". Other codes that were mapped to the same trait are "thought that it is possible to change the automatically generated keywords by hand", and "expressed an opinion that setting profile to private would help with preventing the inference on their images".

For the codes elicited from the open-ended codes, $C1$ manually performed low-level clustering on codes, forming the traits. Then, the traits were discussed and agreed on by $C1$ and $C2$, following the approach by Dupree et al. [11]. As a result, 75 traits were generated from answers to open-ended questions. Since closed-ended codes are different to each other by design (i.e., different response anchors of the Likert scale), we adopt a one-to-one mapping between the codes and the traits. We derive 58 traits from answers to closed-ended questions. A set of all traits $\mathcal{T}$ consisted of $T = 133$ traits.

After the traits were formed, $C1$ grouped them with respect to higher-level categories [11] by creating an affinity diagram. Forming an affinity diagram allows us to mimic the axial coding step in Grounded theory. The aim of the affinity diagram is to reduce the cognitive load for annotators later in the process. Identified higher-level categories include the participant's level of knowledge, the emotions after the first stimulus, and categories of personal or private data. Finally, $C1$ and $C2$ discussed the coherence of both the elicited traits, and the affinity diagram, and resolved disagreements.

4.3 Annotation

By considering the traits in $\mathcal{T}$, we form a basis of the feature space for participants. Each participant P_i can be represented in the form of $\boldsymbol{p_i} \in \{0,1\}^T$. The n'th value of $\boldsymbol{p_i}$ is equal to one if P_i propagated trait t_n in their questionnaire response, and is zero otherwise. We use this participant representation later on for creating persona descriptors.

To map a questionnaire response of P_i into a feature vector $\boldsymbol{p}_i$, we need to perform annotation. For traits elicited from closed-ended codes, we did not require the annotators, since the participant's answer to a closed-ended question(s) could be directly mapped to a trait. However, for traits elicited from open-ended codes, we asked annotators to mark the absence/presence of a trait.

Nine annotators took part in the annotation process of 180 questionnaire responses. Concepts derived with an affinity diagram provided annotators with a clearer structure of the task and built connections between the traits, minimising annotators' fatigue. The annotation workload was split into eight parts: with respect to questionnaire responses, and the type of traits to be annotated. Since the generation set $\mathcal{G}$ consisted of a high number of responses, we partitioned it into three subsets. We did not split the validation set $\mathcal{V}$. We used the concepts derived from an affinity diagram to split the traits to be annotated. Annotators $A1$, $A2$, $A3$, $A4$ worked with traits related to the following affinity diagram concepts: beliefs about advertisement, monetisation, inference on participant's data, types of protective behaviour (prior to the first stimulus, after the first stimulus), beliefs about what the PET does, attitude to the PET, and beliefs about privacy online. Annotators $A5$, $A6$, $A7$, $A8$ worked with traits related to concepts about the participants' level of knowledge, emotions after the first stimulus, categories of personal or private data, use of social media platforms, sensitivity towards possible privacy risks, and beliefs about images. Annotators $A1$, ..., $A8$ were assigned one of the eight parts of the annotation workload. Annotator $A9$ annotated all eight parts of the annotation workload. Each trait per participant was annotated by two annotators. The disagreements between the annotators were discussed during follow-up meetings and resolved in unanimous agreement. Based on the understanding of the annotation process in qualitative studies [4, 11, 12, 31, 38], we do not compute Cohen's kappa [7] for the inter-annotator agreement.

4.4 Likert-scale and Binary Variables

While performing manual annotation, we noticed differences in the nature of the traits: some of the traits were mutually exclusive (e.g., if multiple traits related to the same closed-ended question, only one of those traits could be non-zero), whereas others were not (e.g., a participant could have expressed multiple traits in the same category, such as emotions). To incorporate these traits' differences in a dissimilarity measure between the participants, we first introduce the Likert-scale and binary explanatory variables. The Likert-scale explanatory variables, l_j, allow to group:

(1) the open-ended traits that are mutually exclusive and measure the degree of certain activity/belief, e.g., "has shown to share little/moderate/a lot of personal information in the last 10 images";
(2) the closed-ended traits that come from the same closed-ended question, e.g., "they would never/.../always use the proposed PET";
(3) the closed-ended traits which are manually derived and measure the change in privacy protection importance, and control mismatch, e.g., "experienced a drastic decrease/ .../drastic increase in privacy protection importance".

As a result, we obtain $L = 2 + 10 + 2 = 14$ Likert-scale explanatory variables. The binary explanatory variables, b_i, are propagated from the remaining open-ended traits, $B = 67$. We hence identify $E = L + B = 81$ explanatory variables based on the traits in $\mathcal{T}$. The full list of explanatory variables, their indices and corresponding traits is provided in Appendix B.

We convert $\boldsymbol{p_i} \in \{0,1\}^T$ to its alternative form:

$$\boldsymbol{p'_i} = \boldsymbol{l_i} \oplus \boldsymbol{b_i}, \tag{1}$$

where $\boldsymbol{l_i} \in \mathbb{R}^L$ is a vector of Likert-scale explanatory variables, $\boldsymbol{b_i}$ is the vector for binary explanatory variables, $\boldsymbol{b_i} \in \{0,1\}^B$; $\oplus$ is a vector concatenation operation.

The list of identified Likert-scale variables, $l_* \in [0,1]$, is provided below:

l_1 the amount of shared information, [a lot, few]
l_2 the level of sensitivity to personal data, [low, high]
l_3 the level of privacy protection importance at the beginning of the questionnaire, [unimportant, very important]
l_4 the perception of how easy it is to protect privacy, [very easy, very hard]
l_5 the desired level of control, [never, always]
l_6 the level of awareness about social media checking the content, [not ..., very ...] aware
l_7 the level of awareness about social media building profiles of users, [not ..., very ...] aware
l_8 the level of understanding of what is happening to their data online, [no ..., full ...] understanding
l_9 the perceived level of control over their privacy, [no ..., full ...] control
l_{10} the perception of PET's usefulness, [not useful, useful]
l_{11} how often they would use the proposed PET, [never, always]
l_{12} the level of privacy protection importance at the end of the questionnaire, [unimportant, very important]
l_{13} the change in the privacy protection importance, [drastic decrease, drastic increase]
l_{14} the mismatch between desired and perceived level of control over privacy, [extremely less ..., extremely more ...] control than wanted

5 PERSONAS ELICITATION

To elicit the personas we build on top of divisive hierarchical clustering. We propose a new dissimilarity measure, which incorporates the nature of open- and closed-ended questions by considering their corresponding explanatory variables. Additionally, we propose a two-step dendrogram pruning process, which uses Boschloo's statistical test [3] to validate differences between the clusters.

5.1 Dissimilarity Measure

The closed-ended questions that are linked to Likert-scale explanatory variables require choosing between the set of possible predefined answers. For this kind of data norms like L_1 or L_2 are more suitable than a cosine similarity measure or a dot product[3].

Binary explanatory variables extracted from open-ended questions could not be predicted before the coding process – the participant could not choose between the set of possible answers. Zero value for such a binary explanatory variable does not mean that a participant did not share the views of the trait formulation, it means that they did not express any opinion, rather than disagree with it. For example, the value of one for the trait "thinks that privacy online is deceptive" means that the participant has vocalised this view. However, a value of zero for this trait does not mean that the participant believes that privacy protection is transparent (non-deceptive). This means that using a metric like a dot product (like it was done in the work of Dupree et al. [11]) is preferable over using L_1 metric for measuring dissimilarity, since L_1 penalises for each disagreement, while dot product focuses on the agreements only. However, sometimes an open-ended question can be associated with a Likert-scale variable (e.g., traits that describe the amount of information that a participant is sharing – a lot, moderate, few). In such cases, L_1 or L_2 norms should be used.

We merge the approaches[4] used by Dupree et al. [11] and Biselli et al. [2] by introducing a new dissimilarity measure, which takes into account the nature of the explanatory variable. To achieve this, we consider the Likert-scale part of a feature vector, $\boldsymbol{l}$, and the binary, $\boldsymbol{b}$, separately, since they are different in nature. To this end, we propose a new dissimilarity measure:

$$d(\boldsymbol{p'_i}, \boldsymbol{p'_j}) = \max\left(0,\ \frac{L_1(\boldsymbol{l_i}, \boldsymbol{l_j})}{\sum_{k=1}^{L} r(l_k)} - \frac{\boldsymbol{b_i} \cdot \boldsymbol{b_j}}{B}\right), \tag{2}$$

where L_1 is Manhattan distance, $r(\cdot)$ is a function for calculating the range of the k-th Likert explanatory variable, l_k. The maximum possible distance between the participants is defined by their Likert-scale responses normalised by the maximum distance in L_1. We adjust the distance obtained on $\boldsymbol{l}$ by a normalised dot product on $\boldsymbol{b}$. In Eq. (2), $\max(0,\ \cdot)$ was added to avoid negative dissimilarity scores, which reduced the range of $d(\boldsymbol{p'_i}, \boldsymbol{p'_j})$ to $[0, 1]$.

[3]Dot product or cosine similarity metric should not be used in the case of Likert-scale questions. Treating Likert-scale questions as an unordered set of traits ignores the distances between the Likert items. If the Likert-scale variables are represented as our vector $\boldsymbol{l}$, cosine similarity between $\boldsymbol{l}_i$ and $\boldsymbol{l}_j$ can lead to participants with opposite views (e.g., $\boldsymbol{l}_i = (1, ..., 1)$, $\boldsymbol{l}_j = (5, ..., 5)$) being considered identical.

[4]Dupree et al. [11] have coded the survey responses which were of a closed-ended nature and performed hierarchical agglomerative clustering on a binary feature vector using the dot product. Biselli et al. [2] have clustered participants using the hierarchical agglomerative clustering using Ward's linkage and L_2-norm.

5.2 Dendrogram Construction

We use the dissimilarity measure to perform divisive hierarchical clustering. Hierarchical clustering offers a number of clusters at different levels of abstraction, starting with only one cluster for the whole population and ending with as many clusters as there are data points. We use divisive clustering and separate the most dissimilar instances at each level of a dendrogram [20]. We denote a cluster k at a level of granularity v as $\mathcal{U}_k^v$. We represent a cluster $\mathcal{U}_k^v$ by using a cluster descriptor $\boldsymbol{u}_k^v$. The t-th entry of the descriptor is the frequency of appearance of t-th trait in a cluster, and is calculated as follows:

$$\boldsymbol{u}_k^{v(t)} = \frac{\sum_{\boldsymbol{p_i} \in \mathcal{U}_k^v} \boldsymbol{p_i}^{(t)}}{|\mathcal{U}_k^v|}. \tag{3}$$

We now define a notation for the statistical similarity between clusters. By statistical similarity, we refer to clusters' descriptors having no differences based on a statistical measure. For this, we check if the frequency of appearance of the t-th trait is different between clusters using Boschloo's test [3], which is a test that is suitable for small datasets.

Next, we perform discriminative feature selection. Once the initial dendrogram is built, we identify the most discriminative traits for the final dendrogram construction by reducing the number of traits elicited on the open-ended questions. We perform pairwise comparisons of the clusters in the first 15 levels of the dendrogram. We retain the trait t_i if for at least one comparison the trait had $p < 0.001$. If a trait is associated with a binary explanatory variable and has a $p \geq 0.001$, it is removed. For Likert variables that consist of traits from open-ended questions, if at least one trait is significant, all of the traits of the corresponding variable are retained. This has reduced the number of traits that are significant from T to $S = 72$. All other traits were masked out (set to zero) from the feature vectors $\boldsymbol{p_i}$ and $\boldsymbol{p'_i}$. We built a final dendrogram using only the discriminative features.

5.3 Pruning

Once we build the dendrogram with discriminative features, we prune it to avoid clusters being split into sub-clusters that are statistically similar. To compare clusters $\mathcal{U}_k^v$ and $\mathcal{U}_m^n$ we run Boschloo's test [3] on the discriminative traits, running S statistical tests per comparison of two clusters. We consider clusters $\mathcal{U}_k^v$ and $\mathcal{U}_m^n$ to be significantly different if there exists at least one trait for which the difference in proportions[5] is $p < 0.05$.

We now follow a two-step process for pruning the dendrogram:

(1) each cluster $\mathcal{U}_k^v$ is split into sub-clusters $\mathcal{U}_m^{v+1}$ and $\mathcal{U}_n^{v+1}$ if and only if there exists at least one trait of significance $p < 0.05$; otherwise we say that cluster $\mathcal{U}_k^v$ is non-divisible.
(2) leaves of the dendrogram are merged into their parent cluster if they are statistically similar to other leaves. For each leaf, we perform a comparison with other leaves and count the number of comparisons for which the leaves are not statistically different. We select a cluster with the highest number of insignificant comparisons and go one step back in the dendrogram, merging the leaf with the highest number of insignificant comparisons and its sibling(s) into their parent

[5]We scale 0.05 by S, $S-1$, $S-2$... to allow for Holm-Bonferroni correction.

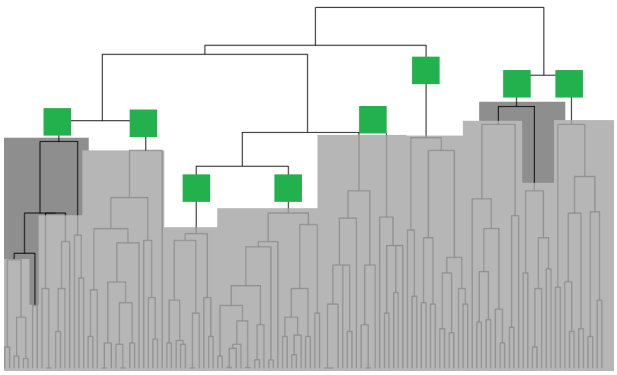

Figure 2: Our two-step method for discovering privacy personas. In step one ▢, the dendrogram is pruned if a parent cluster is split into two sub-clusters that are statistically similar to each other, meaning there are no traits that make these clusters different based on Boschloo's test [3]. In step two ■, the dendrogram is further pruned if there exists at least one leaf that is statistically similar to other leaves. The final personas are in green ■.

cluster. We iteratively repeat this process until for pairwise comparisons of all of the leaves in a tree there exists at least one significantly different trait.

After following Step 1 we obtain twelve leaves, which after Step 2 got reduced to eight leaves (see Fig. 2). We call the obtained leaves the privacy personas. This process allows us to elicit privacy personas that are different to each other based on a statistical measure. Additionally, we compared the confidence intervals (CI) on the traits of the personas, checking if 95% CI for the traits would overlap [1]. For all comparisons, there exists at least one trait for which confidence intervals are not overlapping, supporting the differences between the personas.

6 PRIVACY PERSONAS

The hierarchical structure identified in Section 5 allows us to define eight privacy personas as leaf clusters (see Fig. 3), which are described below.

Unconcerned. This persona shares the most personal information online out of all personas. They have the lowest desired level of control and have no opinion on the importance of privacy protection. They have a significant level of mismatch between the perceived and desired level of control, and would rarely use PET.

In-control Adopter. This persona has a high level of perceived control and they have a strong faith in the usefulness of PET. This persona has only a slight mismatch between the desired level of control and the perceived one. They would almost always use the proposed PET.

In-control Sceptic. This persona is the most sceptical about the proposed PET while having one of the highest levels of perceived control.

Helpless Protector. This persona has only a slight understanding of what is happening to their data. They think that the PET should somehow help to protect their data, but would only occasionally use it.

Knowledgeable Pessimist. For this persona, privacy protection is very hard and they believe that the PET should somehow help to protect their data, and they would use it.

Occasional Protector. This persona sees privacy protection as a very important task. While this persona always wants to be in control over their data, they would only sometimes use the PET.

Adopting Protector. This persona wants to always be in control of their data and sees privacy protection as a very important, but hard task. They are convinced that the PET would allow them to protect their privacy and would always use it.

Knowledgeable Optimist. This is the only persona who strongly believes that privacy protection is mostly easy. This is the second most knowledgeable persona. They see privacy as very important, and would always want to be in control of their data. Because of their knowledge, they see privacy as a mostly easy task. They have a strong faith in the PET and would almost always use it.

The persona cluster sizes are: *Unconcerned* – 14, *In-control Adopter* – 18, *In-control Sceptic* – 11, *Helpless Protector* – 17, *Knowledgeable Pessimist* – 18, *Occasional Protector* – 18, *Adopting Protector* – 11, and *Knowledgeable Optimist* – 23.

We discuss our personas in more detail in Section 8, where we analyse them with respect to their perception of privacy protection importance (comparing to Westin [28]), their knowledge and behaviour (comparing to Biselli et al. [2]), their motivation and knowledge (comparing to Dupree et al. [11]), and their behaviour and privacy protection importance perception (comparing to Schomakers et al. [30]). More detailed descriptors of the personas are provided in Appendices C, D, and E.

7 PRIVACY PERSONAS VALIDATION

In this section, we test the sensitivity of the dendrogram to small perturbations in the dataset (when a subset of participants from a generation set is removed), and participants' saturation (a measure of similarity of participants in the validation and generation sets).

7.1 Sensitivity Analysis

We evaluate the sensitivity of the final clustering structure to the changes in the dataset. To this end, we use the Fowlkes-Mallows (FM) Index [14], which allows us to compare two hierarchical clustering results by incorporating information about their topology and label assignment [14]. FM takes cluster assignment of a set of the same datapoints as an input. Hence, for computing the FM Index (the higher the value, the better the clustering), we compare the dendrogram obtained on a full generation set $\mathcal{G}$ ($n = G$ datapoints) and a dendrogram obtained on a reduced dataset ($n - r$ datapoints): (1) we consider both dendrograms at a level of granularity v; (2) remove the corresponding r datapoints from the labels list obtained on a full generation set $\mathcal{G}$; (3) compute the FM Index.

We compute the FM Index for $r = 1, 2, ..., 6$, selecting half of the minimal persona cluster size as a threshold for r. To ensure the validity of the index, we compute it 500 times, each time randomly sampling $n - r$ datapoints from $\mathcal{G}$. See Fig. 4 for the mean of the distributions for different r and Fig. 7 in Appendix A for more detailed results. For all r when $v = 3$ the $\overline{FM} \geq 0.8$, which means

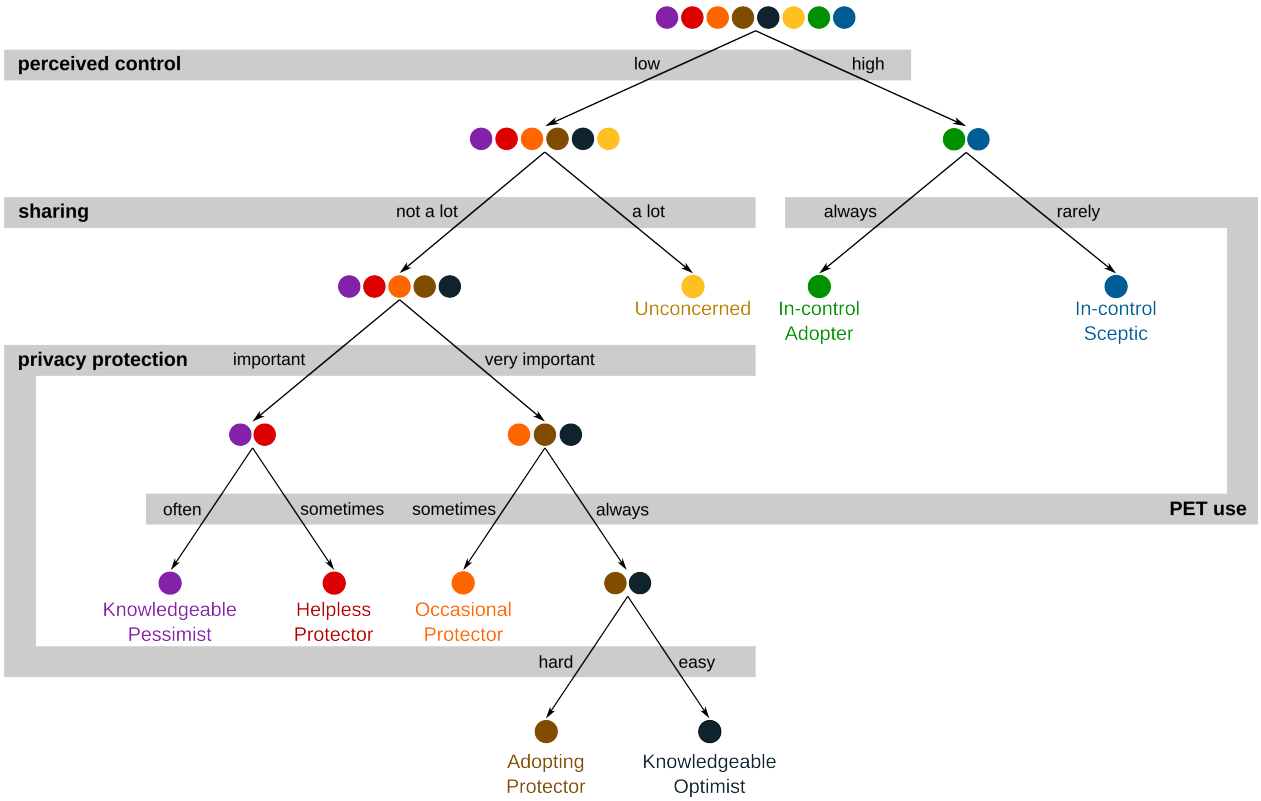

Figure 3: The discriminative features of our identified privacy personas. Text in bold defines an attribute on which the personas differ from each other, the values next to the arrows are the values of a corresponding attribute.

that at least to a level of granularity 3 our results are stable. For all r we see a trend of $\overline{FM}$ gradually decreasing, where the $\overline{FM}$ starts to drop significantly for all r at $v = 10$. The splits in a dendrogram at $v = 2, 3$ are the most stable ones, whereas after $v = 10$ the dendrogram becomes more sensitive to noise. We attribute the noise to the rejected cluster splits, rejecting the first cluster split at a level of granularity $v = 5$. The last accepted split of a parent cluster was a level $v = 15$, which is just above the 0.6 threshold for most r values. Even with the rejected clusters the FM Index stays high, supporting the stability of our clustering.

7.2 Validation of Participants' Saturation

To validate that we have reached participants' saturation, we compare the annotated questionnaires from the $\mathcal{G}$ and $\mathcal{V}$, which were annotated by $A1$, ..., $A7$, $A9$ and $A4$, $A8$, $A9$, respectively. We pose the following question: are there any participants from $\mathcal{V}$ that are considered to be outliers with respect to the distribution of participants in $\mathcal{G}$?

Both coders $C1$, and $C2$ stated that saturation was reached after both of them coded a corresponding set of questionnaire responses. Based on this, we make the following assumption: within our generation set $\mathcal{G}$, each of the participants has at least one person to whom they are similar.

Since the validation set $\mathcal{V}$ is smaller than the generation set $\mathcal{G}$, we do not aim to have a direct comparison of the distributions of participants in $\mathcal{G}$ and $\mathcal{V}$. Instead, we first obtain the distribution of the distances to the closest neighbour for all $P_i \in \mathcal{G}$. Then for each $P_j \in \mathcal{V}$, we find the nearest neighbour from $\mathcal{G}$. We check if the corresponding distance is an outlier in the distribution of the distances to the closest neighbour for all $P_i \in \mathcal{G}$.

To achieve this we compute two distance matrices:

- $\boldsymbol{D_1} = [d_{ij}]$, where $i, j = 1, 2, ..., 130$; $d_{ij} = d(\boldsymbol{p'_i}, \boldsymbol{p'_j})$ if $i \neq j$, $d_{ij} = 1$ otherwise (to avoid self-similarities); $\boldsymbol{p'_i}, \boldsymbol{p'_j} \in \mathcal{G}$.
- $\boldsymbol{D_2} = [d_{pq}]$, where $p = 1, 2, ..., 130$, $q = 1, 2, ..., 50$; $d_{pq} = d(\boldsymbol{p'_p}, \boldsymbol{p'_q})$ $\forall p, q$; $\boldsymbol{p'_p} \in \mathcal{G}$ and $\boldsymbol{p'_q} \in \mathcal{V}$.

We then calculate two vectors, $\boldsymbol{d_1}$ and $\boldsymbol{d_2}$:

$$\boldsymbol{d_1} = [d_{\cdot j}] = [\min_{i=1}^{130} d_{ij}] \quad \text{and} \quad \boldsymbol{d_2} = [d_{\cdot q}] = [\min_{p=1}^{130} d_{pq}], \tag{4}$$

where $\boldsymbol{d_1}$ contains the distances to the closest neighbour in $\mathcal{G}$ for each $P_i \in \mathcal{G}$ and $\boldsymbol{d_2}$ contains the distances to the closest neighbour in $\mathcal{G}$ for each $P_j \in \mathcal{V}$. Based on Tukey's method [33], none of the elements in $\boldsymbol{d_2}$ are outliers, since all of the z-scores for elements in $\boldsymbol{d_2}$ are within $[-1.25, 2.05]$ range. This means that for each participant in $\mathcal{V}$, there exists at least one participant in $\mathcal{G}$ to which they are similar.

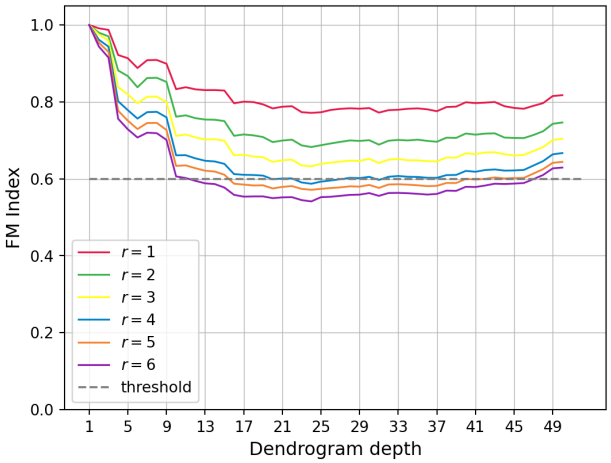

Figure 4: Sensitivity analysis of the dendrogram obtained on the generation set $\mathcal{G}$. We randomly remove r participants from the generation set $\mathcal{G}$, form a new dendrogram, and compute the Fowlkes-Mallows (FM) Index [14] between the newly obtained dendrogram and a dendrogram obtained on the initial generation set $\mathcal{G}$. We repeat the sampling procedure 500 times for each value of r, and report the mean value. We notice that the highest drop in performance takes place at a dendrogram depth equal to 10 for all r. For more detailed plots of the sensitivity analysis, see Appendix A.

8 RELATIONSHIP WITH OTHER PERSONAS

In this section, we compare our personas and the privacy personas identified by Westin [28], Biselli et al. [2], Dupree et al. [11] and Schomakers et al. [30]. For comparison with the related work, we selected the explanatory variables that best reflect the underlying attribute used to define personas.

8.1 Westin

The main classification criteria used by Westin [28] was the level of participants' concern about privacy, which can be mapped to privacy protection importance. By considering traits that capture only privacy protection importance, we can map our personas into Westin's (see Fig. 5). Our personas refine Westins' by considering additional privacy attributes. Our personas support the findings of King [21] that Westin's personas classification does not help to predict "knowledge, behaviours, or an alternative measure of attitudes" [21]. The reason for this is that questions about privacy protection importance are not rich enough to capture the space of privacy personas, losing variability that corresponds to knowledge and behaviour. The level of privacy protection importance can only be informative when defining the *Unconcerned* persona.

Urban [34] considered Westin's classification in a different light. They proposed to consider Westin's Fundamentalist as a privacy-resilient persona, and the Pragmatist and the Unconcerned as privacy-vulnerable personas. They hypothesised that the privacy-resilient persona would be more willing to protect their privacy, whereas privacy-vulnerable ones would be more hesitant to use self-help technologies. We use our personas as a proxy for validating this hypothesis.

We observe that Westin's Unconcerned, which has a one-to-one mapping with our Unconcerned, is the least willing persona to use the proposed PET. However, when it comes to Westin's Pragmatist, we cannot predict how willing one would be to use the PET. We partitioned Westin's Pragmatist with four of our personas, namely *In-Control Adopter*, *Knowledgeable Pessimist* (both are willing to use the PET), *Helpless Protector* (somehow willing to the use PET), and *In-Control Sceptic* (not willing to use PET). Similarly, we partitioned Westin's Fundamentalist with three of our personas, namely *Adopting Protector* and *Knowledgeable Optimist* (both are willing to use the PET), and *Occasional Protector* (who is somehow willing to use the proposed PET). While the least willing personas come from Westin's Unconcerned and Pragmatist clusters, some of Westin's Pragmatists are willing to use the PET. Similarly, Westin's Fundamentalists could have a different degree of willingness to use PET. This means that it is not sufficient to use Westin's categorisation to predict if one would use the PET.

8.2 Biselli

Biselli et al. [2] elicited personas using only two attributes, *behaviour* and *knowledge*, and showed a strong positive correlation between these. They elicited three personas, adopting Westin's [28] naming convention: the Fundamentalist (high knowledge, high privacy-enhancing behaviour), the Pragmatist (moderate knowledge, moderate privacy-enhancing behaviour), and the Unconcerned (low knowledge, low privacy-enhancing behaviour). To compare our personas and Biselli's, we first project our identified personas into the traits with the attributes used by Biselli et al. [2]: *knowledge* and *behaviour* (see Fig. 6).

We observe that some of our personas obey the positive correlation discovered by Biselli et al. [2]. For example, *Knowledgeable Pessimist* and *Knowledgeable Optimist* both fall under the category of high knowledge and high privacy-preserving behaviour. *Occasional Protector*, *Adopting Protector* and *In-Control Sceptic* fall under the category of medium knowledge and medium privacy-preserving behaviour. The only persona that shows the lowest level of protective behaviour is our *Unconcerned*, however, this persona does not show the lowest level of knowledge. We therefore say that our *Unconcerned* can only be weakly mapped to Biselli's. Similarly, our *In-control Adopter* can be weakly mapped to both Biselli's Fundamentalist and Pragmatist, since they both have a high level of privacy-protecting behaviour and medium knowledge. We also identify a new type of persona: *Helpless Protector*, which has the lowest level of knowledge, but a high level of protective behaviour.

While some of our personas follow similar patterns as personas elicited by Biselli et al. [2], they allow for more granularity. For example, by considering only *knowledge* and *behaviour*, the difference between our *Knowledgeable Optimist* and *Knowledgeable Pessimist* would have been lost (e.g., the perception of how easy it is to protect privacy).

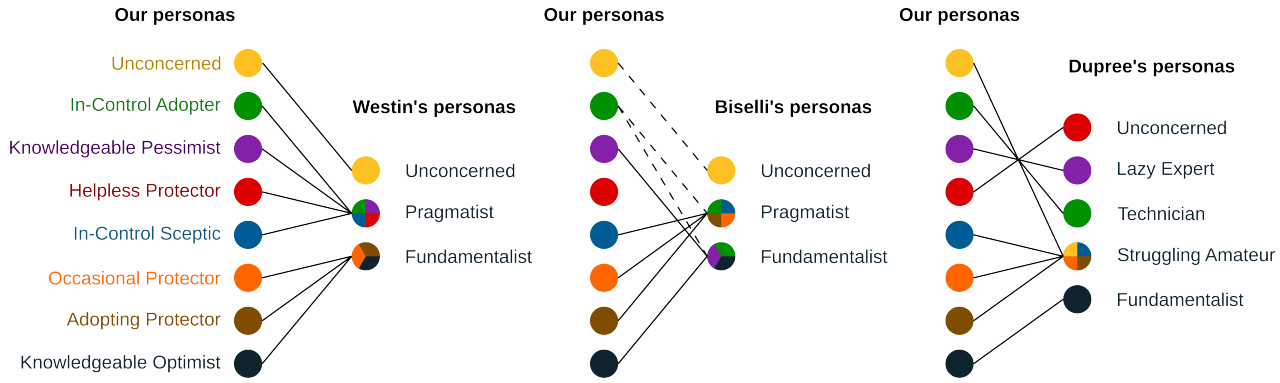

Figure 5: Mapping of our personas into Westin's [28] (left), Biselli's [2] (centre), and Dupree's [11] (right). Our personas are mapped into Westin's by considering the privacy protection importance attribute: high, moderate and low privacy protection importance is mapped into Westin's Fundamentalist, Pragmatist and Unconcerned personas accordingly. Our personas are mapped into Biselli's by considering knowledge and behaviour attributes: high, moderate and low knowledge and privacy-protective behaviour are mapped into Biselli's Fundamentalist, Pragmatist and Unconcerned personas accordingly). Our personas are mapped into Dupree's by considering primarily the knowledge attribute.

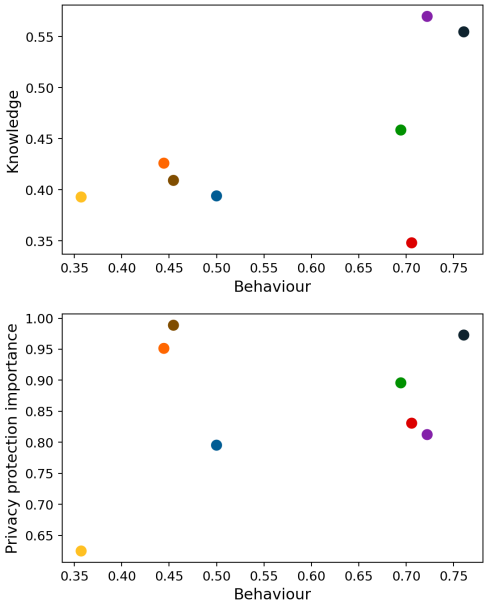

Figure 6: Projection of our personas into 2D spaces: behaviour-knowledge space (top), behaviour-privacy protection importance (bottom). Zero/one for behaviour corresponds to sharing few/a lot of personal elements in their images (maps to l_1 explanatory variable). Zero/one for knowledge is low/high level of knowledge (maps to $(l_6 + l_7 + l_8)/3$). Zero/one for privacy protection importance is low/high level of importance (maps to $(l_3 + l_{12})/2$). Colour-coding: Unconcerned (●), In-control Adopter (●), In-Control Sceptic (●), Knowledgeable Pessimist (●), Helpless Protector (●), Occasional Protector (●), Adopting Protector (●), Knowledgeable Optimist (●).

8.3 Dupree

Dupree et al. [11] elicited five personas using a mixture of qualitative and quantitative analysis. They mapped their Fundamentalist to Westin's Fundamentalist, their Marginally Concerned to Westin's Unconcerned, and the remaining personas (Lazy Expert, Struggling Amateur, and Technician) were segmenting Westin's Pragmatic majority. Dupree et al. [11] positioned their personas in the knowledge-motivation space. Since Ferrarello et al. [13] were measuring knowledge, but not motivation, we use the level of knowledge, as the primary factor for mapping. Dupree et al. [11] identified four levels of knowledge among their personas: low (Moderately Concerned), medium-low (Struggling Amateur), medium-high (Technician), and high (Fundamentalist and Lazy Expert). We map our personas into these categories based on their knowledge, using the four levels of knowledge for this assignment.

Our *Knowledgeable Optimist* and *Knowledgeable Pessimist* have the highest level of knowledge and can be mapped to either Dupree's Fundamentalist or Lazy Expert (see Fig. 5). To disentangle this mapping, one would need to record the future behaviour of these personas. However, given that *Knowledgeable Pessimist* wants to be in control over their data often (and not always), this could mean that they are closer to a Dupree's Lazy Expert than to a Fundamentalist. Similarly, based on the level of knowledge only, our *In-control Adopter* could be mapped to Dupree's Technician. Then, Dupree's Struggling Amateurs can be represented by four of our personas: *Unconcerned, In-control Sceptic, Occasional Protector*, and *Adopting Protector*, given that these personas have similar levels of knowledge. Finally, our *Helpless Protector* has the lowest level of knowledge, making them Dupree's Unconcerned.

While we approximate the mapping by the level of personas' knowledge, we do not have a direct measurement of a motivation attribute. The closest of our attributes are motivation to use a PET, perception of privacy importance, and desired level of control, however, none of these has a direct link with Dupree's motivation.

8.4 Schomakers

Once Schomakers et al. [30] established the clusters, they were analysed with respect to participants' level of awareness, experience, perceived privacy, need for privacy, trust in online companies, and privacy self-efficacy. Notably, these attributes were not used during the clustering process. This means that even if there was a level of variability among these attributes, it was not captured. Values for these attributes could be averaging across multiple hidden subpersonas, hence should not be used for a comparison. This reduces the list of attributes for comparison to concern and behaviour. The closest attributes in our personas are the behaviour and the privacy protection importance (see Fig. 6). Based on this, we can conclude that our *Knowledgeable Optimist* can be mapped to Schomakers' Guardian, having a high level of concern and protective behaviour. Similarly, *Occasional Protector* and *Adopting Protector*, could be mapped to Schomakers' Cynic, having a high level of privacy protection importance, but a lower level of demonstrated protective behaviour. Our *In-Control Adopter*, *Helpless Protector, Knowledgeable Pessimist*, and *In-Control Sceptic* can be mapped to the Pragmatists majority. Finally, our method also provides an additional persona – *Unconcerned*, which is not represented by Schomakers et al. [30].

9 LIMITATIONS

The dataset [13] showcased the responses from the participants from the UK, however, the cultural differences also influence privacy perception [37]. There are conflicting opinions if age influences [2, 39], or does not influence [17] privacy perception, however, it could be one of the influencing factors in privacy personas. We perform validation on a subset of the same dataset, and not a new dataset. All of this means that while these elicited personas are valid for this dataset, additional validation with people from other countries and other age groups is beneficial. People who do not use social media were not represented in the study by Ferrarello et al. [13], and they might represent other personas (i.e. highly secure ones [11]). We plan to collect more data and use the approach by Caines et al. [5] for automating the annotation.

We did not record the user interactions with an application, as it is done in behavioural data studies [26, 36]. Additionally, measuring user preferences towards the protection of a specific type of content [23, 25] was out of the scope of our paper. Bridging the gap between modelling users by considering their privacy attitudes and perceptions (via surveys), factual behaviours (via user's interactions with applications), the type of content to be protected and its use is a part of our future work.

10 CONCLUSION

We proposed a method that combines qualitative and quantitative approaches for discovering statistically different privacy personas. We used open coding and Boschloo's test [3], and accounted for the nature of the people's responses to questions (closed-ended vs open-ended). We built a dendrogram structure of clusters of participants, performed a two-step pruning process to ensure that the elicited personas were different to each other, and introduced a new measure for calculating this difference. We discovered eight personas: *Knowledgeable Optimist, In-control Adopter, In-Control Sceptic, Knowledgeable Pessimist, Helpless Protector, Occasional Protector, Adopting Protector, and Unconcerned.* We provided personas' descriptors that allow one to understand the persona's behaviour, attitudes and other attributes, and have a numerical comparison between the personas. We have shown that it is important to consider the attitude towards privacy protection importance, participants' behaviours, decisions, and knowledge when discovering personas. Considering these attributes together when eliciting personas allows one to capture more complex patterns, which leads to better persona understanding, and hence more fine-grained privacy support.

Relevant use cases of our personas include facilitating persona-tailored privacy support, accounting for the privacy needs of each persona according to their descriptor, modelling user interactions, and identifying privacy threats that can be posed in an application. Finally, privacy personas can be used for recruiting people for privacy studies based on their persona type, and not demographics only. In future work, we will create a larger dataset and explore automatic trait detection to allow for scalability and its applicability to other domains. We will also minimise the number of questions that could be used to elicit personas to have a short and reliable questionnaire by validating the scales using reliability and validity measures [6, 27]. Finally, we aim to use the identified privacy personas to support the design of online privacy settings [2, 11, 31].

ACKNOWLEDGMENTS

We thank the anonymous reviewers and the revision editor for their detailed comments, which helped improve our work. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] Alan Agresti and Brent A. Coull. 1998. Approximate Is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician* 52, 2 (May 1998), 119–126. https://doi.org/10.2307/2685469

[2] Tom Biselli, Enno Steinbrink, Franziska Herbert, Gina M. Schmidbauer-Wolf, and Christian Reuter. 2022. On the Challenges of Developing a Concise Questionnaire to Identify Privacy Personas. *Proceedings on Privacy Enhancing Technologies* 4 (Oct. 2022), 645–669. https://doi.org/10.56553/popets-2022-0126

[3] Ronald D. Boschloo. 1970. Raised Conditional Level of Significance for the 2×2-Table When Testing the Equality of Two Probabilities. *Statistica Neerlandica* 24, 1 (March 1970). https://doi.org/10.1111/j.1467-9574.1970.tb00104.x

[4] Virginia Braun and Victoria Clarke. 2022. *Thematic Analysis: a Practical Guide.* SAGE Publications, London, UK; Thousand Oaks, US.

[5] Andrew Caines, Sergio Pastrana, Alice Hutchings, and Paula J. Buttery. 2018. Automatically Identifying the Function and Intent of Posts in Underground Forums. *Crime Science* 7, 1 (Dec. 2018). https://doi.org/10.1186/s40163-018-0094-4

[6] Eunseong Cho. 2016. Making Reliability Reliable: a Systematic Approach to Reliability Coefficients. *Organizational Research Methods* 19, 4 (Oct. 2016), 651–682. https://doi.org/10.1177/1094428116656239

[7] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 1 (April 1960), 37–46. https://doi.org/10.1177/001316446002000104

[8] Jessica Colnago, Lorrie F. Cranor, Alessandro Acquisti, and Kate H. Jain. 2022. Is It a Concern or a Preference? An Investigation into the Ability of Privacy Scales to Capture and Distinguish Granular Privacy Constructs. In *Proceedings of the Eighteenth USENIX Conference on Usable Privacy and Security (SOUPS 2022).* USENIX Association, Boston, US.

[9] Juliet M. Corbin and Anselm Strauss. 1990. Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qualitative Sociology* 13, 1 (March 1990). https://doi.org/10.1007/BF00988593

[10] John W. Creswell and Vicki L. Plano Clark. 2007. *Designing and Conducting Mixed Methods Research.* SAGE Publications, Thousand Oaks, US.

[11] Janna L. Dupree, Richard Devries, Daniel M. Berry, and Edward Lank. 2016. Privacy Personas: Clustering Users via Attitudes and Behaviors toward Security

Practices. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI 2016)*. ACM, San Jose, US, 5228–5239. https://doi.org/10.1145/2858036.2858214

[12] Isioma Elueze and Anabel Quan-Haase. 2018. Privacy Attitudes and Concerns in the Digital Lives of Older Adults: Westin's Privacy Attitude Typology Revisited. *American Behavioral Scientist* 62, 10 (Sept. 2018), 1372–1391. https://doi.org/10.1177/0002764218787026

[13] Laura Ferrarello, Rute Fiadeiro, Riccardo Mazzon, and Andrea Cavallaro. 2022. Reframing the Narrative of Privacy Through System-Thinking Design. In *Proceedings of the Design Research Society (DRS 2022)*. Design Research Society, Bilbao, SPN. https://doi.org/10.21606/drs.2022.620

[14] Edward B. Fowlkes and Colin L. Mallows. 1983. A Method for Comparing Two Hierarchical Clusterings. *J. Amer. Statist. Assoc.* 78, 383 (Sept. 1983), 553–569. https://doi.org/10.1080/01621459.1983.10478008

[15] Andrea Gallardo, Chris Choy, Jaideep Juneja, Efe Bozkir, Camille Cobb, Lujo Bauer, and Lorrie F. Cranor. 2023. Speculative Privacy Concerns About AR Glasses Data Collection. *Proceedings on Privacy Enhancing Technologies* 4 (Oct. 2023), 416–435. https://doi.org/10.56553/popets-2023-0117

[16] Graham R. Gibbs. 2018. *Analyzing Qualitative Data*. SAGE Publications, London, UK. https://doi.org/10.4135/9781526441867

[17] David Goyeneche, Stephen Singaraju, and Luis Arango. 2024. Linked by Age: A Study on Social Media Privacy Concerns Among Younger and Older Adults. *Industrial Management & Data Systems* 124, 2 (Jan. 2024), 640–665. https://doi.org/10.1108/IMDS-07-2023-0462

[18] Monique M. Hennink, Bonnie N. Kaiser, and Vincent C. Marconi. 2017. Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? *Qualitative Health Research* 27, 4 (March 2017), 591–608. https://doi.org/10.1177/1049732316665344

[19] Bernard J. Jansen, Joni O. Salminen, and Soon-Gyo Jung. 2020. Data-Driven Personas for Enhanced User Understanding: Combining Empathy with Rationality for Better Insights to Analytics. *Data and Information Management* 4, 1 (March 2020). https://doi.org/10.2478/dim-2020-0005

[20] Leonard Kaufman and Peter J. Rousseeuw. 2005. *Finding Groups in Data: an Introduction to Cluster Analysis*. Wiley, Hoboken, US.

[21] Jennifer King. 2014. Taken Out of Context: an Empirical Analysis of Westin's Privacy Scale. In *Proceedings of the Workshop on Privacy Personas and Segmentation (SOUPS 2014)*. USENIX Association, Menlo Park, US. https://api.semanticscholar.org/CorpusID:481189

[22] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. 2020. Towards a Taxonomy of Content Sensitivity and Sharing Preferences for Photos. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu, US. https://doi.org/10.1145/3313831.3376498

[23] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I. Hong. 2014. Modeling Users' Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings. In *Proceedings of the Tenth USENIX Conference on Usable Privacy and Security (SOUPS 2014)*. USENIX Association, Menlo Park, US, 199–212.

[24] Andrés Lucero. 2015. Using Affinity Diagrams to Evaluate Interactive Prototypes. In *Proceedings of the Human-Computer Interaction (INTERACT 2015)*. Springer International Publishing, Bamberg, GER, 231–248. https://doi.org/10.1007/978-3-319-22668-2_19

[25] George R. Milne, George Pettinico, Fatima Hajjat, and Ereni Markos. 2017. Information Sensitivity Typology: Mapping the Degree and Type of Risk Consumers Perceive in Personal Data Sharing. *Journal of Consumer Affairs* 51, 1 (March 2017), 133–161. https://doi.org/10.1111/joca.12111

[26] Jonathan Mugan, Tarun Sharma, and Norman Sadeh. 2011. *Understandable Learning of Privacy Preferences Through Default Personas and Suggestions*. Technical Report. Pittsburgh, US.

[27] Melvin R. Novick and Charles Lewis. 1967. Coefficient Alpha and the Reliability of Composite Measurements. *Psychometrika* 32, 1 (March 1967). https://doi.org/10.1007/BF02289400

[28] Ponnurangam Kumaraguru and Lorrie F. Cranor. 2005. *Privacy indexes: a survey of Westin's studies*. ISRI CMU-ISRI-05-138. Institute for Software Research International, Carnegie Mellon University, Pittsburgh, US. http://reports-archive.adm.cs.cmu.edu/anon/anon/home/ftp/usr0/ftp/isri2005/CMU-ISRI-05-138.pdf

[29] Prolific. 2024. *Prolific*. London, UK. Retrieved Mar. 1, 2024 from https://www.prolific.co

[30] Eva-Maria Schomakers, Chantal Lidynia, and Martina Ziefle. 2019. A Typology of Online Privacy Personalities: Exploring and Segmenting Users' Diverse Privacy Attitudes and Behaviors. *Journal of Grid Computing* 17, 4 (Dec. 2019), 727–747. https://doi.org/10.1007/s10723-019-09500-3

[31] Alina Stöver, Nina Gerber, Henning Pridöhl, Max Maass, Sebastian Bretthauer, Indra Spiecker gen. Döhmann, Matthias Hollick, and Dominik Herrmann. 2023. How Website Owners Face Privacy Issues: Thematic Analysis of Responses from a Covert Notification Study Reveals Diverse Circumstances and Challenges. *Proceedings on Privacy Enhancing Technologies* 2 (April 2023), 251–264. https://doi.org/10.56553/popets-2023-0051

[32] Alina Stöver, Sara Hahn, Felix Kretschmer, and Nina Gerber. 2023. Investigating how Users Imagine their Personal Privacy Assistant. *Proceedings on Privacy Enhancing Technologies* 2 (April 2023), 384–402. https://doi.org/10.56553/popets-2023-0059

[33] John W. Tukey. 1977. *Exploratory Data Analysis*. Addison-Wesley, Reading, US; London, UK.

[34] Jennifer Urban. 2014. The Privacy Pragmatic as Privacy Vulnerable. In *Proceedings of the Workshop on Privacy Personas and Segmentation (SOUPS 2014)*. Menlo Park, US. https://doi.org/10.31235/osf.io/yh8nj

[35] U.S. Government Publishing Office. 2001. *Opinion Surveys: What Consumers Have to Say About Information Privacy*. Retrieved Mar. 1, 2024 from https://www.govinfo.gov/content/pkg/CHRG-107hhrg72825/html/CHRG-107hhrg72825.htm

[36] Pamela Wisniewski, Bart P. Knijnenburg, and Heather R. Lipford. 2014. Profiling Facebook Users Privacy Behaviors. In *Proceedings of the Workshop on Privacy Personas and Segmentation (SOUPS 2014)*. USENIX Association, Menlo Park, US.

[37] Anran Xu, Zhongyi Zhou, Kakeru Miyazaki, Ryo Yoshikawa, Simo Hosio, and Koji Yatani. 2023. DIPA2: An Image Dataset with Cross-Cultural Privacy Perception Annotations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (Dec. 2023). https://doi.org/10.1145/3631439

[38] Yu Xu and Michael J. Lee. 2020. Identifying Personas in Online Shopping Communities. *Multimodal Technologies and Interaction* 4, 2 (May 2020). https://doi.org/10.3390/mti4020019

[39] Eva-Maria Zeissig, Chantal Lidynia, Luisa Vervier, Andera Gadeib, and Martina Ziefle. 2017. Online Privacy Perceptions of Older Adults. In *Proceedings of the Human Aspects of IT for the Aged Population (ITAP 2017)*. Springer International Publishing, Vancouver, CAN, 181–200. https://doi.org/10.1007/978-3-319-58536-9_16

[40] Sergej Zerr, Stefan Siersdorfer, and Jonathon Hare. 2012. PicAlert!: a System for Privacy-Aware Image Classification and Retrieval. In *Proceedings of the Twenty-First ACM international conference on Information and knowledge management*. ACM, Maui, US, 2710–2712. https://doi.org/10.1145/2396761.2398735

[41] Chenye Zhao, Jasmine Mangat, Sujay Koujalgi, Anna Squicciarini, and Cornelia Caragea. 2022. PrivacyAlert: a Dataset for Image Privacy Prediction. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. AAAI Press, Washington, US, 1352–1361. https://doi.org/10.1609/icwsm.v16i1.19387

APPENDICES

A SENSITIVITY ANALYSIS

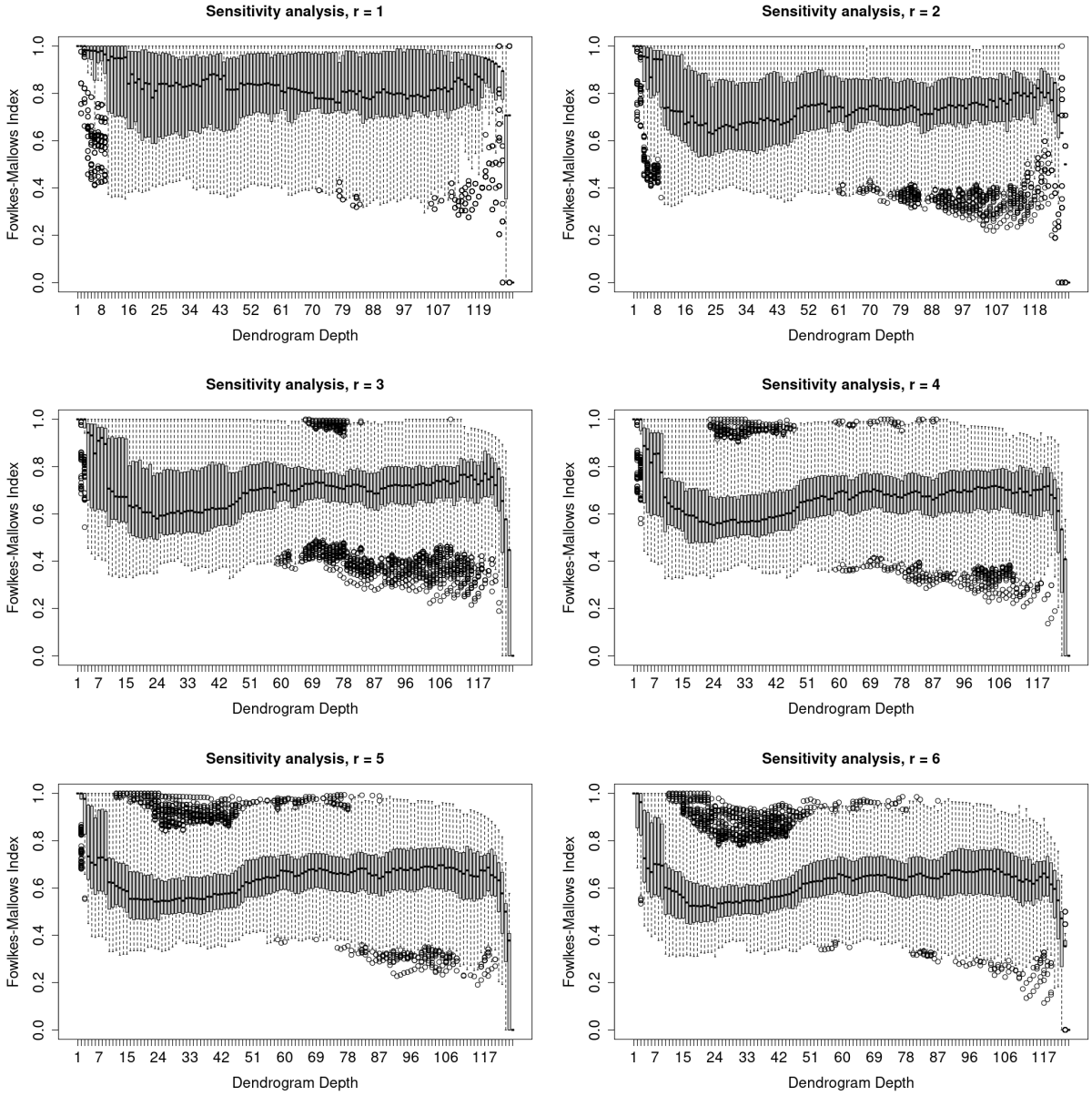

Figure 7: Extended sensitivity analysis of the dendrogram obtained on the generation set $\mathcal{G}$. We randomly remove r participants from the generation set $\mathcal{G}$, form a new dendrogram, and compute the Fowlkes-Mallows (FM) Index [14] between the newly obtained dendrogram and a dendrogram obtained on the initial generation set $\mathcal{G}$. We repeat the sampling procedure 500 times for each value of r, and report the distribution of the FM Index.

B EXPLANATORY VARIABLES

l_1 **sharing**
$t_{1,2,3}$ has shown to share a few / moderate / a lot pieces of personal information in the last 10 images
l_2 **sensitivity**
$t_{4\text{-}8}$ has a low / low medium / medium / medium high / high level of sensitivity regarding the private information
l_3 **privacy protection importance (initial)**
$t_{9\text{-}13}$ initially believes that privacy protection is unimportant / not that important / no opinion / moderately important / very important
l_4 **how easy it is to protect**
$t_{14\text{-}18}$ initially thinks that protecting personal information is extremely easy / mostly easy / moderately easy / hard / very hard with the tools that are available to them
l_5 **desired level of control**
$t_{19\text{-}23}$ initially expressed that they never / rarely / sometimes / often / always want to be in control of their personal information
l_6 **knowledge about checking for sensitive content**
$t_{24,25,26}$ illustrated no awareness / only awareness / familiarity of social media checking images for sensitive content
l_7 **knowledge about awareness for personal information extraction**
$t_{27,28,29}$ illustrated no awareness / only awareness / familiarity of social media being able to extract personal information from an image and generate a personal profile on a user
l_8 **knowledge about what is happening with the data**
$t_{30\text{-}34}$ claimed to have no understanding / little understanding / moderate understanding / to mostly understand / full understanding of what is happening to their images when uploading them online
l_9 **perceived level of control**
$t_{35\text{-}39}$ after being exposed to how companies can extract data from their images, they feel that they have no control / only little control / moderate control / can mostly control / full control over their personal information when sharing images online
l_{10} **usefulness of the PET**
$t_{40,41,42}$ they feel that the proposed privacy enhancing technology will not help them / not sure will help them / will help them to protect their privacy when sharing images online
l_{11} **would use the PET**
$t_{43\text{-}47}$ they would never / rarely / sometimes / often / always use the proposed privacy enhancing filter when sharing images online
l_{12} **privacy protection importance (end)**
$t_{48\text{-}52}$ after being exposed to companies being able to extract personal information and a proposed privacy enhancing tool, they believe that privacy protection is unimportant / not that important / no opinion / moderately important / very important
l_{13} **privacy protection importance (delta)**
$t_{53\text{-}59}$ drastic decrease / significant decrease / slight increase / no change / slight decrease / significant increase / drastic increase in privacy protection importance after being exposed to companies being able to extract personal information and a proposed privacy enhancing tool
l_{14} **control mismatch**
$t_{60\text{-}66}$ extreme (less control than wanted) / significant (less control than wanted) / slight (less control than wanted) / no / slight (more control than wanted) / significant (more control than wanted) / extreme (more control than wanted) mismatch between the perceived and desired level of control over personal information after being exposed to companies being able to extract personal information
b_1 t_{67}: demonstrated a high level of knowledge in some aspects
b_2 t_{68}: demonstrated a lack of knowledge in some aspects
b_3 t_{69}: appreciates that before experience, did not realise all the potential risks when sharing their image online
b_4 t_{70}: after being exposed to how companies can extract data from their images, the participant felt surprised
b_5 t_{71}: after being exposed to how companies can extract data from their images, the participant felt confused / lost
b_6 t_{72}: after being exposed to how companies can extract data from their images, the participant felt unimpressed (extracted information was generic)
b_7 t_{73}: after being exposed to how companies can extract data from their images, the participant felt uncomfortable / scared
b_8 t_{74}: after being exposed to how companies can extract data from their images, the participant felt angry / violated / offended
b_9 t_{75}: after being exposed to how companies can extract data from their images, the participant did not feel surprised (was aware of this happening)
b_{10} t_{76}: has expressed that physical security is an element of private information for them
b_{11} t_{77}: has expressed that financial status is an element of private information for them
b_{12} t_{78}: has expressed concern about protecting the loved ones/friends
b_{13} t_{79}: has expressed that sexual orientation is an element of private information for them
b_{14} t_{80}: has expressed concern about protecting a younger generation
b_{15} t_{81}: has expressed that a political view is an element of private information for them
b_{16} t_{82}: has expressed that health status is an element of private information for them
b_{17} t_{83}: has expressed that identity is an element of private information for them
b_{18} t_{84}: thinks that posts too much personal/private information
b_{19} t_{85}: claims that does not post personal/private information
b_{20} t_{86}: shares personal information in private messages or private account
b_{21} t_{87}: does not post often
b_{22} t_{88}: believes that all images are private
b_{23} t_{89}: believes that some images are more private than others (e.g., they would like to protect only some images or that only some images reveal personal information)
b_{24} t_{90}: believes that one image can tell a lot (e.g., from one image a lot of information can be inferred)
b_{25} t_{91}: believes that some photos reveal information that is already shared regardless for other purposes
b_{26} t_{92}: likes relevant ads / relevant ads are useful
b_{27} t_{93}: does not like relevant ads
b_{28} t_{94}: feels that irrelevant ads are useful (e.g., less distraction)
b_{29} t_{95}: feels that relevant ads are easy to spot (they believe they can easily recognise when something is advertised to them)
b_{30} t_{96}: feels that relevant ads create an isolation bubble or feel invasive
b_{31} t_{97}: feels that ads are annoying
b_{32} t_{98}: believes that monetisation could be ok as long certain conditions are met (e.g., anonymisation, giving consent)
b_{33} t_{99}: found the number of companies/bids surprising when were exposed to how companies can extract data from their images
b_{34} t_{100}: does not like that profile data is sold (e.g., makes them feel like a "commodity")
b_{35} t_{101}: believes that inference could be ok as long certain conditions are met (e.g., anonymisation, giving consent)
b_{36} t_{102}: has mentioned that it is ok for a company to extract / use the data from the images they post online
b_{37} t_{103}: has found the inference on images is very shallow (e.g., stereotypical, offensive) when exposed to how companies can extract data from their images
b_{38} t_{104}: has mentioned that they do not care about inference on what they do share
b_{39} t_{105}: has mentioned that they do not like inference on their images
b_{40} t_{106}: believes that it is too late for them to worry when sharing the images online or do not care about what they reveal
b_{41} t_{107}: has accepted inference performed by the companies on the images that they share online because this is how the world works
b_{42} t_{108}: initially, believes that not sharing is a safe option when it comes to protecting the personal/private information they share online
b_{43} t_{109}: initially, believes that image editing could help (e.g., blurring, cropping) when it comes to protecting the personal/private information they share online
b_{44} t_{110}: initially, believes that profile / post settings manipulations could help (e.g., manually changing the hashtags to the image, setting the account to private, using incorrect location when sharing photos or sharing with delay), when it comes to protecting the personal / private information they share online
b_{45} t_{111}: after being exposed to how companies can extract data from their images, has expressed that not sharing is a safe option when it comes to protecting the personal / private information they share online
b_{46} t_{112}: after being exposed to how companies can extract data from their images, has expressed that image editing could help (e.g., blurring, cropping) when it comes to protecting the personal / private information they share online
b_{47} t_{113}: after being exposed to how companies can extract data from their images, has expressed that profile / post settings manipulations could help (e.g., manually changing the hashtags to the image, setting the account to private, not including the name when sharing the image) when it comes to protecting the personal / private information they share online
b_{48} t_{114}: after being exposed to how companies can extract data from their images was not sure how to protect the personal / private information they share online
b_{49} t_{115}: believes that the proposed privacy enhancement filter messes up monetisation based on personal data
b_{50} t_{116}: believes that the proposed privacy enhancement filter messes up the gathering of personal data and removes objectification
b_{51} t_{117}: has expressed that the proposed privacy enhancement filter makes the image look better or closer to their taste from their perspective

b_{52} t_{118}: has expressed that the proposed privacy enhancement filter makes the image look not as good from their perspective
b_{53} t_{119}: has expressed that the proposed privacy enhancement filter provides general protection and makes them feel better ("adding just to be sure", "they allow to protect me")
b_{54} t_{120}: believes that the effect of the proposed privacy enhancement filter is very little or none (e.g., "Filters would not change my day-to-day life")
b_{55} t_{121}: expressed that they would be choosing between the aesthetics of the image and protection when if they were to use a proposed privacy enhancement technology filter
b_{56} t_{122}: is amazed by the proposed privacy enhancement filter
b_{57} t_{123}: does not trust the the proposed privacy enhancement filter (agree to use it only when they understand how the privacy enhancement technology works, or believe that the proposed technology is "too good to be true")
b_{58} t_{124}: believes that privacy is about who has access to image / data inferred from the image
b_{59} t_{125}: believes that privacy is about what or how much image / data extracted from an image reveals (identities, lifestyle, etc.)
b_{60} t_{126}: believes that privacy is about a conscious decision of what to share / hide
b_{61} t_{127}: believes that privacy is about what you might “unintentionally” share (e.g., hidden inference from data)
b_{62} t_{128}: believes that privacy is about controlling who owns your data
b_{63} t_{129}: believes that privacy is about to whom image/data is sold
b_{64} t_{130}: believes that privacy is about controlling what the image is used for (e.g., for example, not allowing it to be a part of the algorithm)
b_{65} t_{131}: believes that privacy online is deceptive (e.g., information is less private than you think)
b_{66} t_{132}: believes that privacy online does not exist
b_{67} t_{133}: believes that privacy is something we do not have full control of

C OUR PRIVACY PERSONAS' DESCRIPTORS

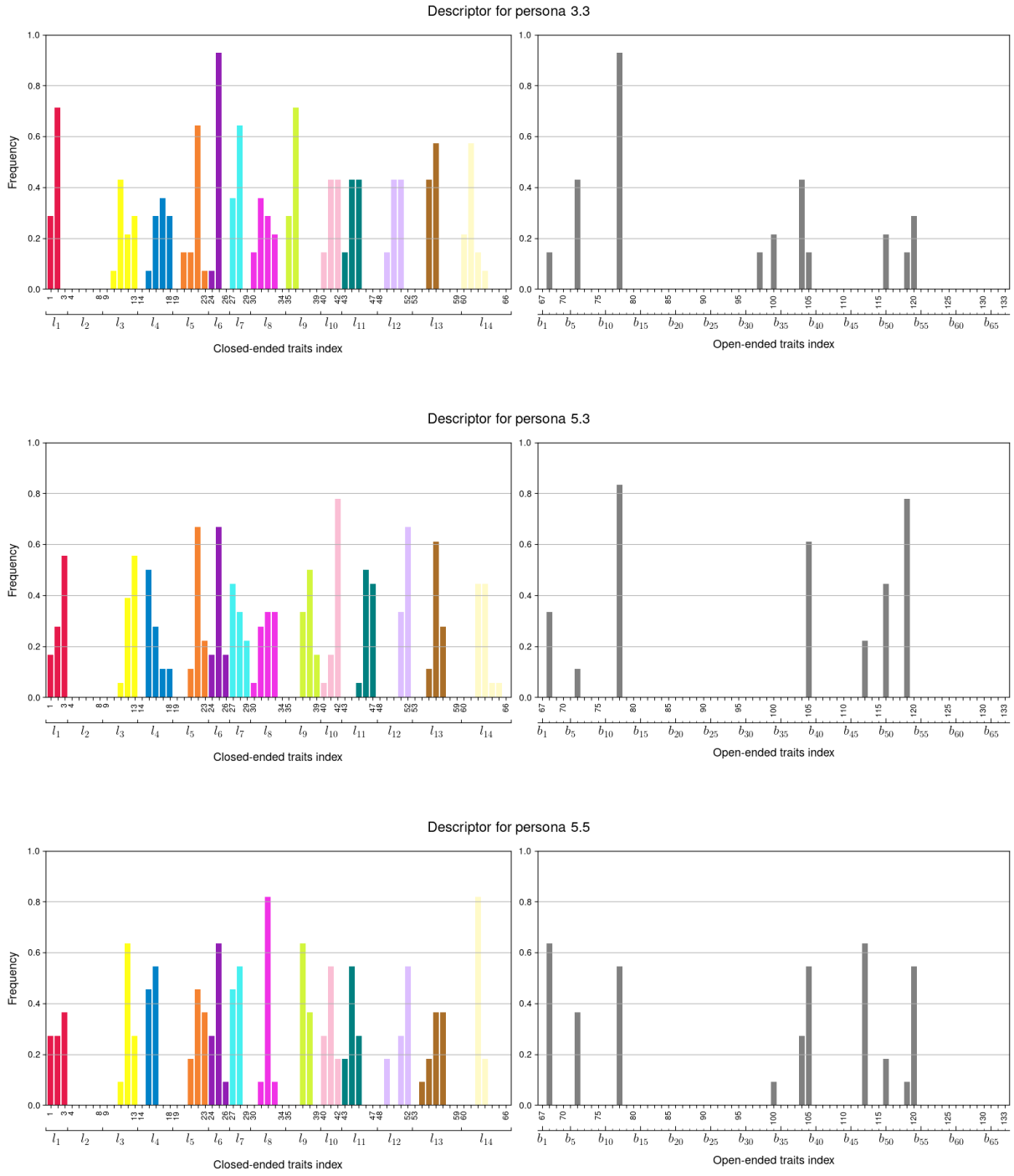

Fig. 8 continued on the next page...

Figure 8: Privacy personas' descriptors for Unconcerned (persona 3.3), In-Control Adopter (persona 5.3) and In-Control Sceptic (persona 5.5). Traits are grouped by the corresponding Likert-scale variables (left) and binary variables (right). The colours of the traits on the left indicate grouping with respect to the Likert-scale variables: l_1, ..., l_{14}. The upper horizontal axis corresponds to the trait ID, lower horizontal axis corresponds to the explanatory variable ID.

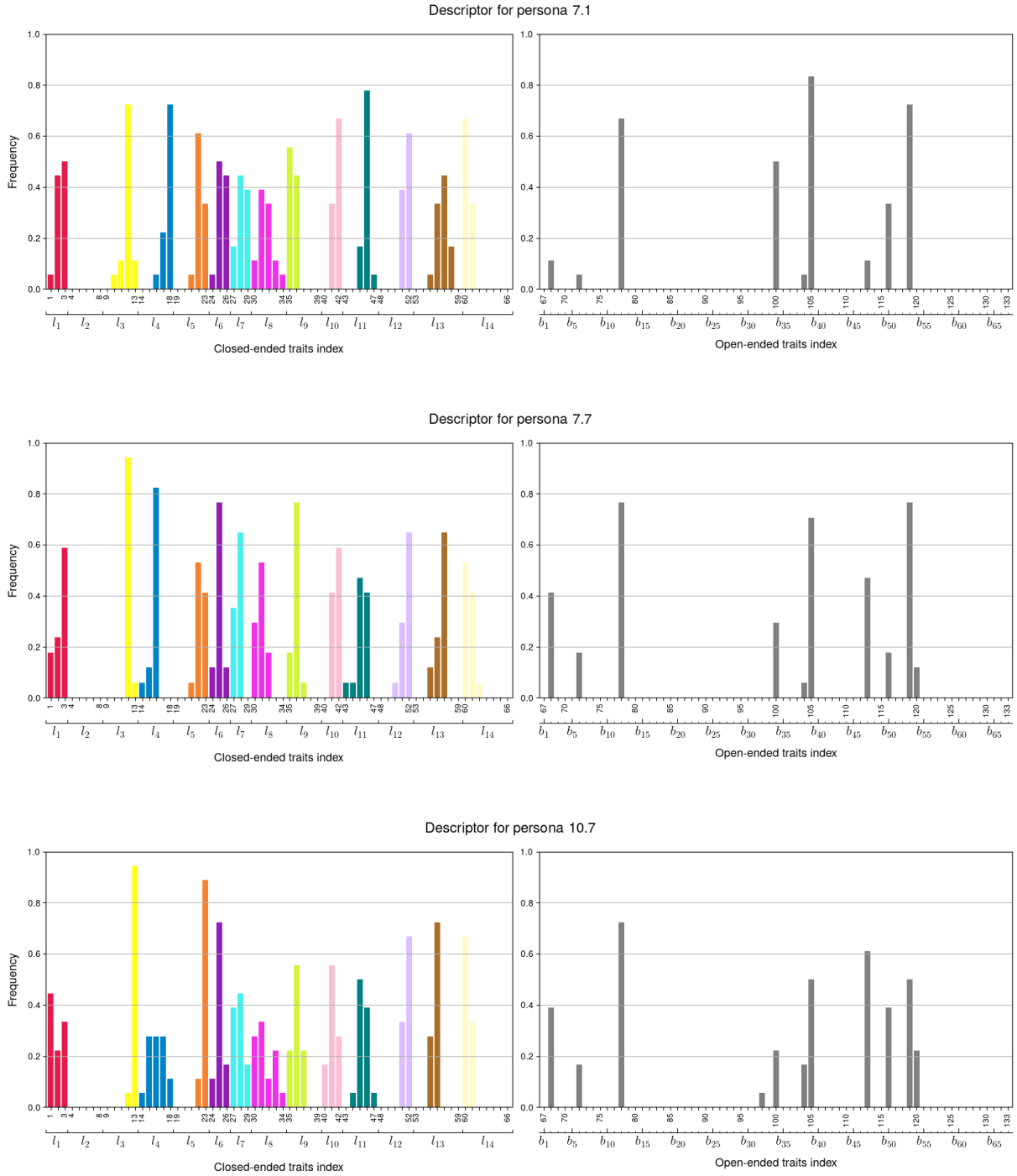

Fig. 8 continued on the next page...

Figure 8: Privacy personas' descriptors for Knowledgeable Pessimist (persona 7.1), Helpless Protector (persona 7.7) and Occasional Protector (persona 10.7). Traits are grouped by the corresponding Likert-scale variables (left) and binary variables (right). The colours of the traits on the left indicate grouping with respect to the Likert-scale variables: l_1, ..., l_{14}. The upper horizontal axis corresponds to the trait ID, lower horizontal axis corresponds to the explanatory variable ID.

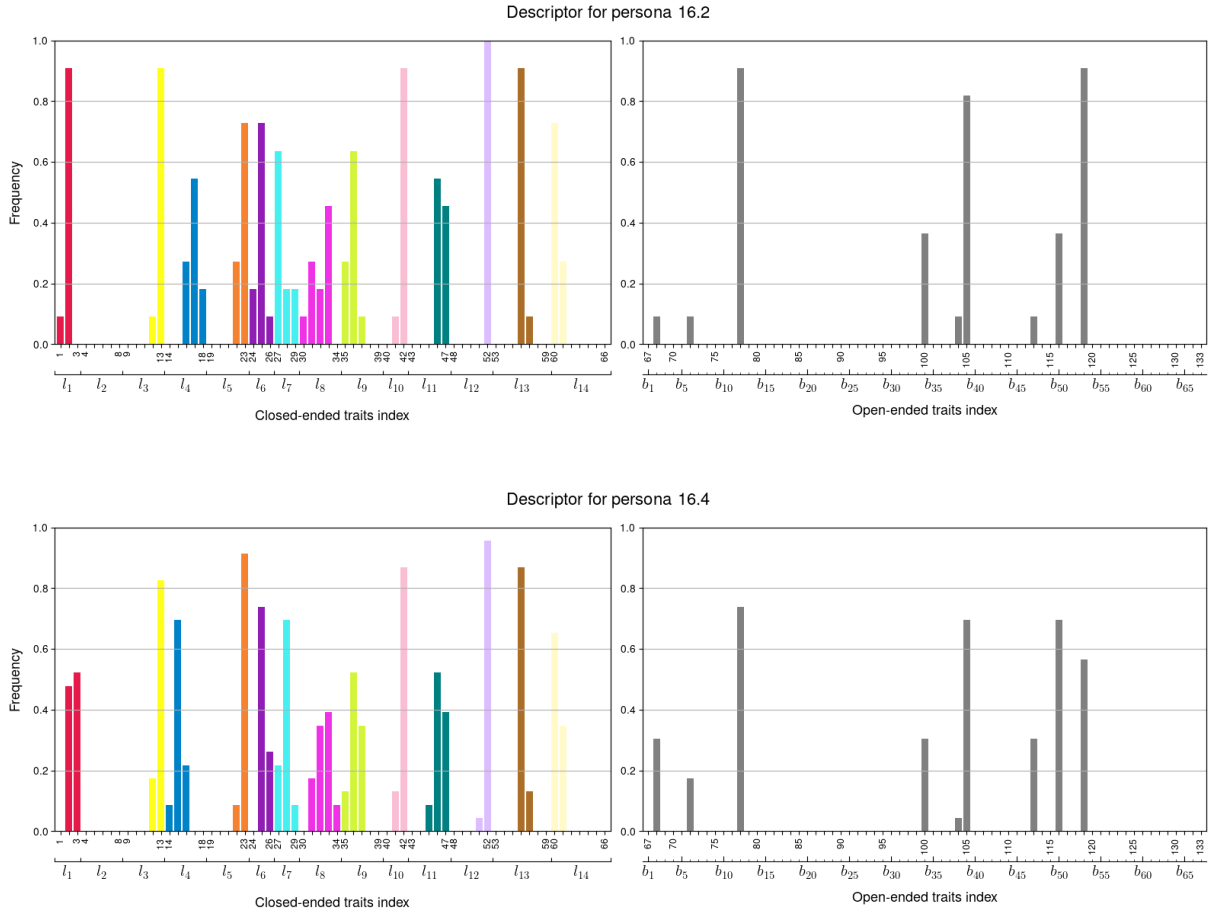

Figure 8: Privacy personas' descriptors for Adopting Protector (persona 16.2) and Knowledgeable Optimist (persona 16.4). Traits are grouped by the corresponding Likert-scale variables (left) and binary variables (right). The colours of the traits on the left indicate grouping with respect to the Likert-scale variables: l_1, ..., l_{14}. The upper horizontal axis corresponds to the trait ID, lower horizontal axis corresponds to the explanatory variable ID.

D COMPARISON OF LIKERT-SCALE EXPLANATORY VARIABLES FOR OUR PRIVACY PERSONAS

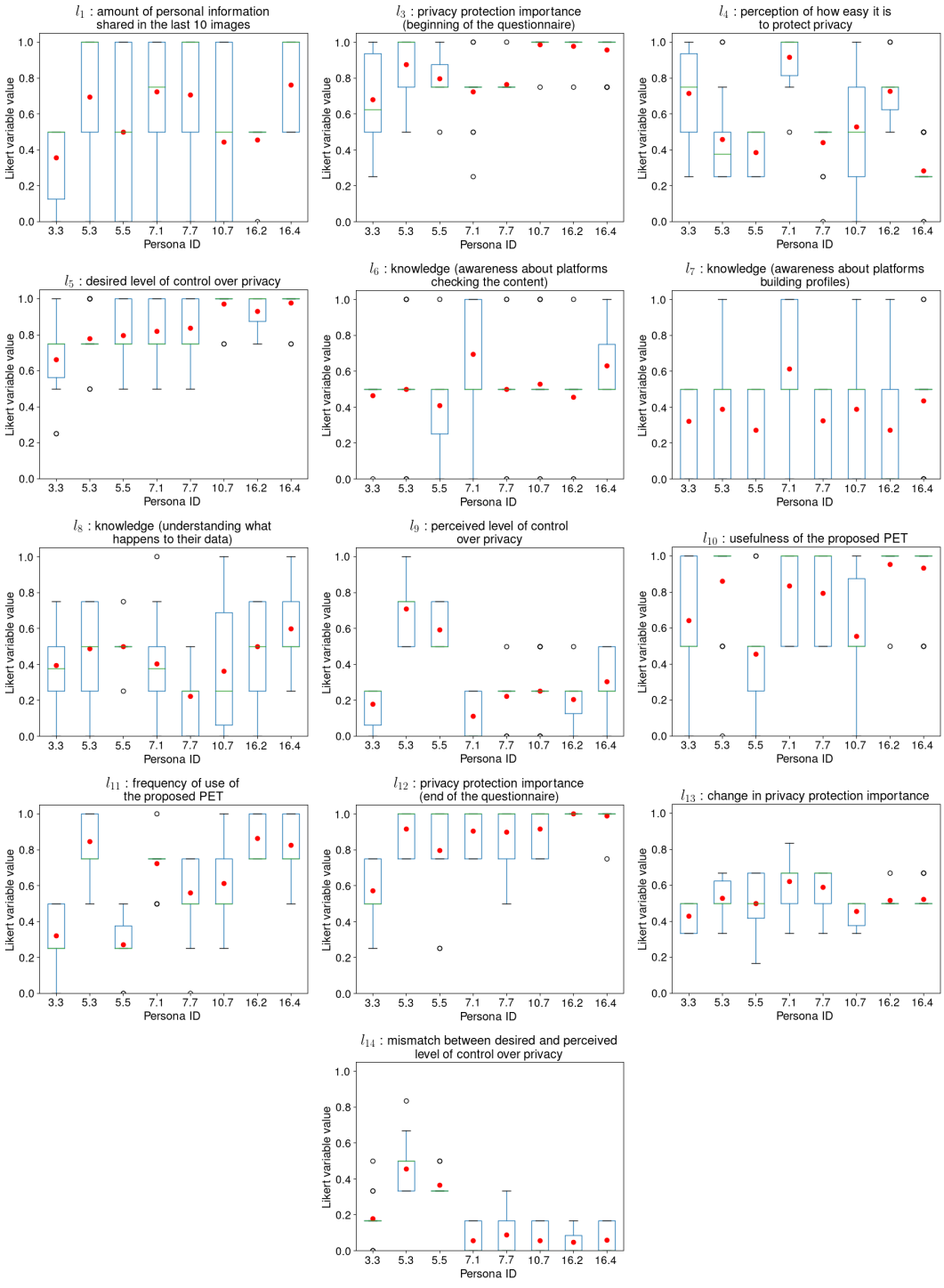

Figure 9: Distributions of the Likert-scale explanatory variables for personas comparison (the values were re-scaled to $[0, 1]$ range). The red square corresponds to the mean of the explanatory variable. Key – 3.3: Unconcerned, 5.3: In-Control Adopter, 5.5: In-Control Sceptic, 7.1: Knowledgeable Pessimist, 7.7: Helpless Protector, 10.7: Occasional Protector, 16.2: Adopting Protector, 16.4: Knowledgeable Optimist.

E OUR PRIVACY PERSONAS IN 2D SPACES

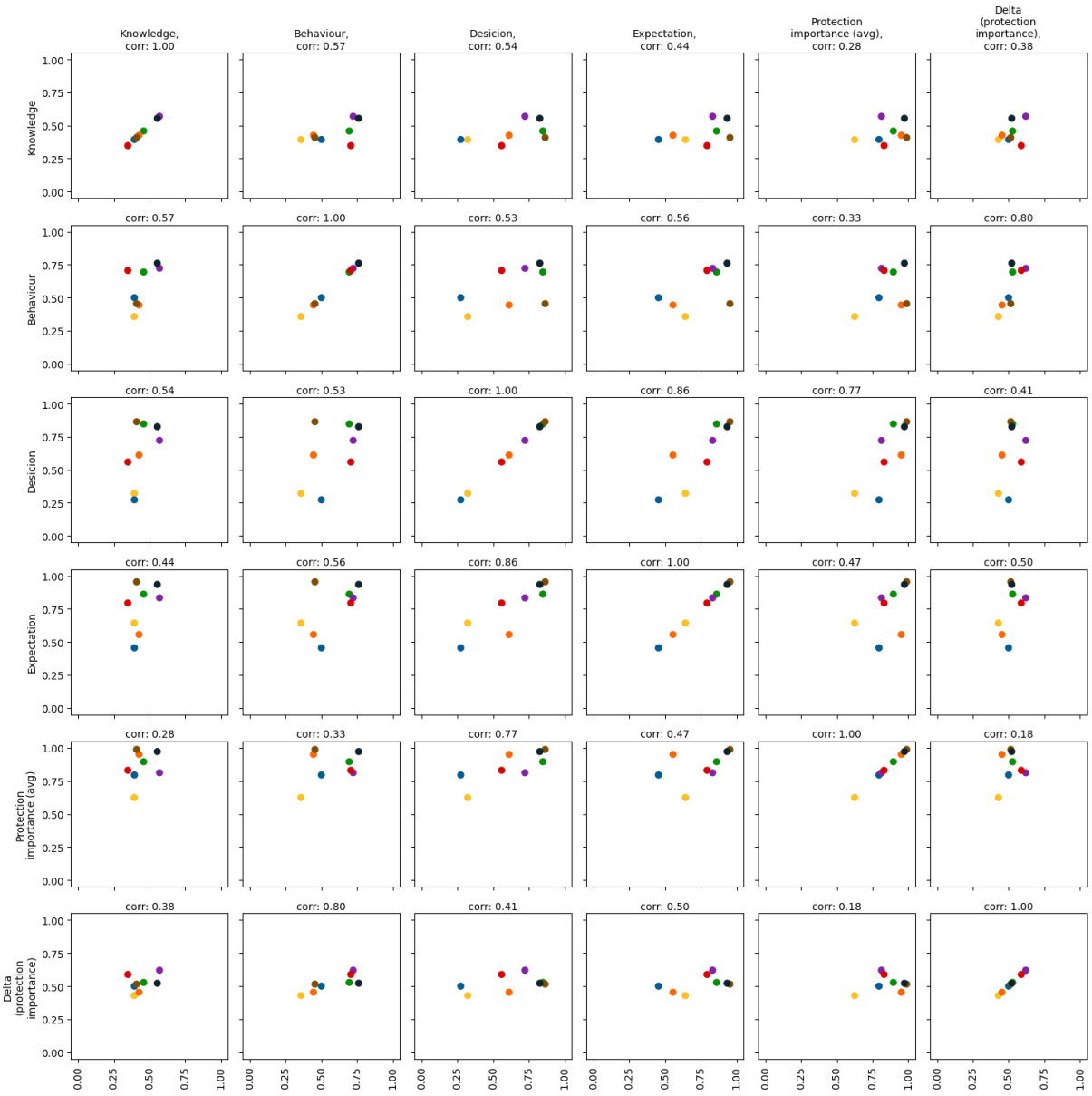

Figure 10: Projection of our privacy personas into 2D spaces, where the dimensions are: knowledge, behaviour, decision to use PET, expectation of PET's efficacy, privacy protection importance, and change in privacy protection importance. Zero/one for knowledge is low/high level of knowledge (maps to $(l_6 + l_7 + l_8)/3$). Zero/one for behaviour is low/high level of self-reported privacy-preserving behaviour (maps to l_1). Zero/one for decision to use PET is low/high level of willingness to use PET (maps to l_{11}). Zero/one for expectation of PET's efficacy is low/high level of perceived efficacy of the proposed PET tool (maps to l_{10}). Zero/one for privacy protection importance is low/high perceived level of privacy protection importance (maps to $(l_3 + l_{12})/2$). Zero/one for change in privacy protection importance is drastic decrease/increase in privacy protection importance throughout the questionnaire (maps to l_{13}). Colour-coding: Unconcerned (●), In-control Adopter (●), In-Control Sceptic (●), Knowledgeable Pessimist (●), Helpless Protector (●), Occasional Protector (●), Adopting Protector (●), Knowledgeable Optimist (●).