On the Feasibility of Predicting Users' Privacy Concerns using Contextual Labels and Personal Preferences

Yaqing Yang
Tsinghua University
Beijing, China
yang-yq19@mails.tsinghua.edu.cn

Tony W Li
University of California San Diego
La Jolla, California, USA
toli@ucsd.edu

Haojian Jin
University of California San Diego
La Jolla, California, USA
haojian@ucsd.edu

ABSTRACT

Predicting users' privacy concerns is challenging due to privacy's subjective and complex nature. Previous research demonstrated that generic attitudes, such as those captured by Westin's Privacy Segmentation Index, are inadequate predictors of context-specific attitudes. We introduce ContextLabel, a method enabling practitioners to capture users' privacy profiles across domains and predict their privacy concerns towards unseen data practices. ContextLabel's key innovations are (1) using non-mutually exclusive labels to capture more nuances of data practices, and (2) capturing users' privacy profiles by asking them to express privacy concerns to a few data practices. To explore the feasibility of ContextLabel, we asked 38 participants to express their thoughts in free text towards 13 distinct data practices across five days. Our mixed-methods analysis shows that a preliminary version of ContextLabel can predict users' privacy concerns towards unseen data practices with an accuracy (73%) surpassing Privacy Segmentation Index (56%) and methods using categorical factors (59%).

CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**; **Human and societal aspects of security and privacy**;

KEYWORDS

Privacy, Empirical study that tells us about people

ACM Reference Format:
Yaqing Yang, Tony W Li, and Haojian Jin. 2024. On the Feasibility of Predicting Users' Privacy Concerns using Contextual Labels and Personal Preferences. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3613904.3642500

1 INTRODUCTION

Predicting users' privacy concerns towards unseen data practices can significantly improve today's privacy ecosystem. First, businesses can use these predictions to understand whether their data collection and usage approaches are in accordance with individuals' privacy expectations [30, 33]. Second, developers and HCI researchers may leverage these predictions to empower users with personalized privacy management tools, enabling them to have default privacy settings tailored to their specific attitudes [40, 51]. Third, policymakers can leverage these predictions to better understand the privacy protection needs of the population and design more effective privacy regulations [1, 53].

However, modeling users' privacy concerns is challenging. Typical modeling processes can be decoupled into two common steps: capturing users' privacy profiles, then predicting their privacy concerns using the captured profiles [27, 34, 37, 48]. For example, Alan Westin developed Privacy Segmentation Indexes consisting of a few questions and a set of rules to group participants into three categories based on their responses (fundamentalists, pragmatists, and unconcerned) [27]. However, many have found that Westin's categories have limited effectiveness in predicting people's privacy attitudes towards specific data practices [4, 17, 27, 35, 52].

Recently, many studies have found that contextual factors within a data practice significantly impact users' privacy attitudes [20, 42, 51, 52, 54]. Researchers have also started to use factorial vignette experiments to capture users' domain-specific privacy profiles [5, 10, 11, 28, 33, 37, 42, 54], suggesting the impact of contextual factors on users' privacy concerns or privacy norms. For example, Emami-Naeini et al. analyzed how a few domain-specific categorical factors (e.g., data types, purposes) affect users' comfort levels in the context of IoT data collection [37]. However, these models and captured factors can hardly be generalized to data practices in a different domain, such as applying an IoT model to forecast user comfort levels in targeted advertising [32, 41].

This paper introduces ContextLabel, a new method that enables practitioners to capture users' privacy profiles across domains and predict their privacy concerns towards unseen data practices. ContextLabel has two key ideas. First, rather than simplifying a real-world data practice to a limited set of exclusive categorical factors, ContextLabel asks practitioners to annotate a data practice with multiple **non-exclusive labels** (e.g., 'Price Discrimination', 'Absence of Consent'). We hypothesize that the incorporation of a wider array of labels, capturing the nuances of the specific data practice under consideration, will lead to a more accurate modeling of users' privacy concerns across domains. Second, ContextLabel captures users' privacy profiles by asking them to express their privacy concerns for a few free-text data practices, then predicts their attitudes to unseen data practice using associated labels. We hypothesize that users' discomfort towards data practices may correlate with different contextual labels and personal preferences [20, 42, 52, 54], and we want to leverage this correlation to predict users' privacy concerns towards an unseen data practice. For example, if a user strongly opposes data-driven price discrimination in airline bookings, she may also hold negative views towards another data practice involving price discrimination.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642500

We conducted an online survey involving 38 participants over five days to explore the feasibility of ContextLabel. We prompted participants daily to express their thoughts in free text regarding three different data practices and evaluate their overall comfort level quantitatively. Overall, each participant examined 13 distinct data practices. To gauge the consistency of users' privacy perceptions regarding a specific data practice, we asked participants to evaluate one repetitive data practice and respond to questions from the Privacy Segmentation Questions [27] on the first, third, and fifth days of the survey.

To predict users' privacy attitudes towards an unseen data practice, we assume that users' perspectives on privacy for a specific data practice show a certain level of logical reasoning. We investigated this by assessing whether individuals maintained consistency in rating the same scenario three times at different intervals, and by analyzing their free-text explanations for these ratings. We found that participants' privacy attitudes are highly consistent in the span of 5 days. 80% of participants expressed the same level of privacy concern in the generic privacy segmentation questions, while the remaining 20%, who were less consistent, mostly had some borderline perceptions. Further, we found that each participant mentioned nearly identical concerns in their free-text responses regarding the repeated privacy scenario across three days.

The authors then developed a codebook with 18 distinct labels and used it to annotate all 13 data practices. We then identified contextual labels that correlated to users' privacy concerns across domains. Our results show that some users have heightened sensitivity to particular contextual labels, leading to specific privacy concerns. Using these insights, we captured users' inclinations concerning various contextual labels by assessing their responses on several data practices and then utilized the captured preferences to predict their concerns on previously unseen data practices. Our data reveals that our predictive model improved prediction accuracy (73%) compared to the Privacy Segmentation Index (56%) and categorical contextual factors (59%).

Scope and Limitation: Our paper is an exploratory work that studies the feasibility of predicting users' privacy concerns across domains. Note the 18 labels provided are an initial set. We anticipate that subsequent research may introduce more labels by extending the codebook. A key advantage of ContextLabel is that researchers can re-label user feedback using the codebook without having to discard prior survey responses.

In this paper, our main contributions are as follows:

- We empirically demonstrated the feasibility of predicting users' privacy concerns across domains.
- We introduced a preliminary method to model users' privacy concerns using non-mutually exclusive labels and users' preferences. This method can predict users' privacy concerns towards unseen data practices with an accuracy that surpasses Privacy Segmentation Index and methods using categorical factors.
- We conducted a systematic study of users' privacy preferences concerning data practices and general questions over a brief period, specifically 5 days, utilizing a two-stage data collection method. We contribute new evidence that users' preferences are largely consistent and exhibit some levels of rationality.

2 RELATED WORK & RESEARCH QUESTIONS

The main objective of this paper is to predict users' privacy attitudes using contextual labels and personal preferences, which originates from three fundamental premises: (1) users' privacy concerns for a specific data practice show a certain level of logical reasoning, which can be approximated as a function of users' individual privacy preferences and the contexts of the data practice; (2) by incorporating a broader selection of non-mutually exclusive labels, we can capture more nuances of a data practice; and (3) instead of running factorial vignette experiments, it's feasible to capture users' privacy profiles by analyzing their open-ended feedback on selected data practices. We have organized related studies around these assumptions.

Modeling users' privacy concerns. Many studies have investigated whether users are rational in context-specific, privacy-related behaviors, either actual or intended [2, 18, 52]. For example, Privacy Calculus [29] assumes that users are rational beings whose decisions and actions are propelled by their intent to optimize their benefits. When the anticipated benefits of data sharing surpass the costs, users are generally expected to willingly share their data. In contrast, numerous studies on consumer decision behavior have also shown that the decision-making process is influenced by various cognitive biases and heuristics [3, 23, 50], such as availability bias [43], the framing effect [49], and confirmation bias [39]. Flender and Müller put forth a contrasting proposition [16], suggesting that a decision's outcome is not settled until the moment the decision is actually made [24], and two distinct decisions cannot be deemed interchangeable within the context of decision-making [16].

We hypothesize that users' privacy concerns are consistent over a short period and exhibit some levels of rationality. We asked participants to repeatedly express their privacy concerns about a data practice and respond to Westin's Index questions to validate the hypothesis. We then analyzed if their numerical ratings were consistent across multiple responses. Further, we analyzed participants' open-ended responses to infer whether these responses demonstrate rationality.

RQ1: How much rationality we can observe from users' privacy attitudes and concerns toward a specific data practice?

Non-mutually exclusive contextual labels v.s. categorical contextual factors. One of the most widely used theoretical frameworks in modeling users' privacy decisions is Nissenbaum's contextual integrity [38]. This theory posits that privacy choices are guided by specific information norms tied to particular contexts. Traditionally, such contexts can be delineated using certain categorical factors, as highlighted in [10, 14], including actors, attributes, and transmission principles. While many studies have found that these contextual factors within a data practice significantly impact users' privacy attitudes [6, 8, 9, 14, 31, 37, 38, 46], it remains hard to use these factors to predict users' privacy concerns across domain.

We hypothesize that simplifying a real-world data practice to a small set of exclusive categorical factors might overlook its intricate nuances. Instead, we want to explore a new method that annotates

a data practice with multiple non-exclusive labels to improve the prediction models.

RQ2: How are different context labels and categorical factors correlated with users' privacy attitudes and concerns across domains?

Capturing users' preferences and predicting users' attitudes. Many studies seek to model users' privacy attitudes across domains, using demographic [12] information like education, gender, age, and ethnicity [17, 35, 36, 52], or personality traits [19, 52]. However, few studies have indicated the effectiveness of demographic predictors [52]. Other studies have used the widely adopted general questions from Westin Privacy index [27] to categorize participants into three groups with different privacy attitudes. However, no evidence showed that either the individual questions or the derived categories are predictive of participants' reactions to specific scenarios [52].

More recently, researchers have started to use vignette factorial surveys to profile users' privacy decisions/attitudes [7, 10, 31, 33, 35, 37, 47]. Researchers often identify a few common factors (e.g., data types) in a specific task setting (e.g., IoT, mobile permissions) and leverage the category factors to generate or control numerous tested scenarios. For instance, Emami-Naeini et al. conducted a 1,007-participant vignette study to capture privacy expectations of users in 380 IoT use-case scenarios [37]. Liu et al. [33] analyzed privacy and security decisions of smartphone users who were asked to choose between "granting", "denying" or "requesting to be dynamically prompted" for 12 permissions of the apps they downloaded. Schechter et al. [42] conducted a study examining users' reactions to a modified version of the Facebook Emotion Contagion Experiment [26]. Serramia et al. [44] selected factors like data types, recipients and transmission principles to generate smart devices scenarios, and leveraged a collaborative filtering approach to predict user preferences. Similarly, Abdi et al. [1] implemented data mining to find which contexts in the Smart Home Personal Assistants ecosystem shared attributes and had the same acceptability. These studies were able to investigate users' preferences [1, 26, 37, 44] or identify meaningful user profiles [33]. However, it is challenging to adjust the prediction model for a new domain [10], as the tested scenarios stem from domain-specific factors and researchers need to collect data again for the new domain.

Instead, we aim to capture users' preferences regarding different contextual factors (e.g., 'Price Discrimination', 'Absence of Consent') across diverse domains. We then use these preferences to predict their attitudes towards other unseen data practices.

RQ3: How can contextual factors and personal preferences be effectively captured, and to what extent can they predict users' privacy concerns towards unseen data practices?

3 METHOD

We conducted a five-day online study session on Amazon's Mechanical Turk (AMT) to collect participants' privacy attitudes towards 13 selected scenarios. We worked to deliver two main outputs: (1) using non-mutually exclusive labels to capture more nuances of a data practice, and (2) capturing users' privacy profiles by asking them to express their privacy concerns to a few data practices. The outputs showed participants' privacy rationale (see section 4.1), the correlation between ContextLabel and privacy attitudes, and concern categories (see section 4.2). ContextLabel predicts privacy concerns with promising accuracy (73%) (see Section 4.3).

3.1 Survey Sessions

Survey Design. Figure 1 shows the overall survey structure. Each participant was required to engage in each day's survey for five consecutive days. This methodology allowed us to collect privacy attitudes and concerns from the same participant across various scenarios and evaluate their consistency. Spreading the workload over five days instead of one also ensured that participants maintained their engagement and focus throughout the study. In order to obtain accurate assessments regarding participants' privacy attitudes, we split the chosen scenarios into different data actions (Figure 2, left), namely *data collection*, *data processing*, *data sharing*, and *data usage*. Next, we generated five different sets of questionnaire surveys for each day using scenarios described in the next section. Each set comprised three distinct scenarios from 13 cases in Table 1 and corresponding data actions are listed in Appendix C.

The main part of our surveys contained a consent page, tutorial examples, and scenario evaluation. On the tutorial page, participants were shown example answers for a scenario where an insurance company shares costumers' health data to third parties [52]. Then, respondents rated their comfort level towards data actions in each scenario using a five-item Likert scale (1 = Extremely uncomfortable, 5 = Extremely comfortable). After rating, participants were asked to express their concerns and reasons in free-form text (Figure 2, middle). To conduct efficient quantitative analysis, we forwarded the collected responses to other crowd workers for annotation of 14 predefined privacy concern categories [20]. Crowd workers were asked to review the privacy concerns listed and either select the relevant options or provide additional information to indicate which concerns were expressed in the free-form text responses (Figure 2, right).

To investigate whether participants' attitudes towards the same scenario remained consistent within a short time span and to unveil any rational reasoning behind their responses, we designed a consistency test whose analysis is further discussed in section 4.1. Specifically, we chose one scenario (Table 1, scenario 6) and integrated it into surveys of the first, third, and fifth days. At the beginning of those three surveys, we also included three frequently used questions from Westin's Privacy Segmentation Index [27, 52], asking participants to rate generic privacy related questions in the following manner: *For each of the following statements, how strongly do you agree or disagree? [1: Strongly Disagree, 2: Somewhat Disagree, 3: Somewhat Agree, 4: Strongly Agree]*:

(1) *Consumers have lost all control over how personal information is collected and used by companies.*
(2) *Most businesses handle the personal information they collect about consumers in a proper and confidential way.*
(3) *Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.*

Although Westin's index has been deemed inefficient in capturing participants' privacy attitudes in previous works [4, 17, 27, 35, 52], to our knowledge, no study has examined respondents' consistency in answering questions of Westin's index within a short time span. Therefore, we repeated these three questions together

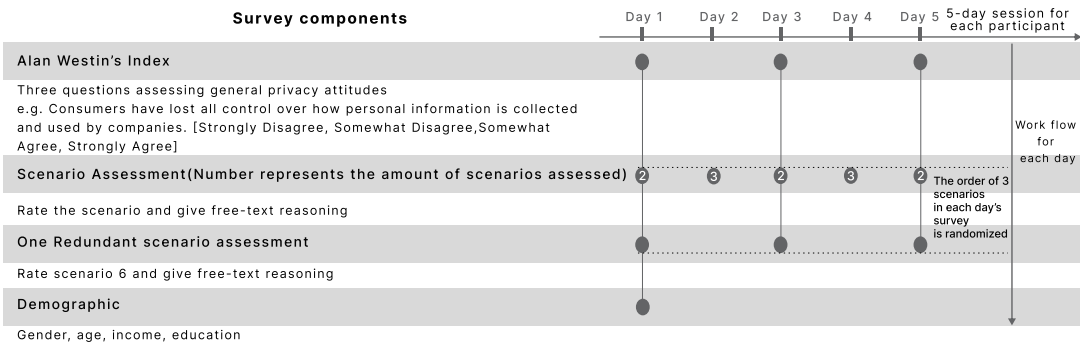

Figure 1: An overview of the five-day survey protocol for capturing users' privacy profiles by asking them to rate and express their concerns to data practices. Surveys 2 and 4 include 3 distinct stories. Surveys 1, 3, and 5 include the three questions from Westin's Privacy Segmentation Index and three stories, one of which is the repeated redundant one. The repeated scenario and questions are designed for the consistency test.

Table 1: Scenarios used in surveys to gauge privacy attitudes. The split data actions are listed in Appendix C. Case 6 is the scenario used as the redundant scenario in the consistency test.

#	Scenario Name	Description
1	Search engine clickthrough data	A company records users' clickthrough behavior in an A/B test experiment in a nonanonymous way.
2	Loyalty card in a retail store	A retail store collects users' data through a loyalty card and uses the data for insurance and coupon personalization.
3	Checkout-free retail store	An e-commerce company opens a checkout-free retail store by installing various sensors inside a physical store.
4	Game chat log	An online game company uses its chat logs to identify potential problems in the workplace
5	Pregnancy intimate data	A pregnancy app shares users' intimate body data with their employers
6	Social network	A social networking service company analyzes users' posts through sentiment analysis and uses insights in different ways.
7	Data science experiments in a dating app	An online dating app conducts several experiments to understand the nature of romance.
8	Email contacts for social network bootstrapping	A technology company appropriates users' email data to bootstrap a new social network service.
9	Fitness tracking	A wearable technology company collects users' intimate behavior data and makes them public by default.
10	Retail store pregnancy	A retail store predicts users' pregnancy status by analyzing their purchase history.
11	Insurer employs AI	An insurance company uses facial-recognition technology to identify untrustworthy and unprofitable customers.
12	Travel service dynamic pricing	A travel agency collects users' device data to adjust the service price dynamically.
13	Ride-share dynamic pricing	A ride-sharing app collects users' device battery data to adjust the service price dynamically.

with the redundant scenario for our consistency test. The responses also served as our baseline for predicting participants' attitudes and concerns in subsequent analysis. Additionally, unlike typical surveys with attention-check questions, we leveraged the free-text field [22] and the consistency test to prevent random responses.

Story Selection. To collect participants' privacy concerns across diverse domains, we selected 13 data practices (Table 1) from Jin et al. [20]. These practices covered areas such as Internet of things (IoT), e-commerce, social networks, advertising, computational psychology, data science experiments, and scenarios involving vulnerable populations. Initially, the practices involved multiple information applications, resulting in diverse outcomes. We assigned a single data application to each practice, enabling broader response collection across domains while mitigating participant fatigue. To

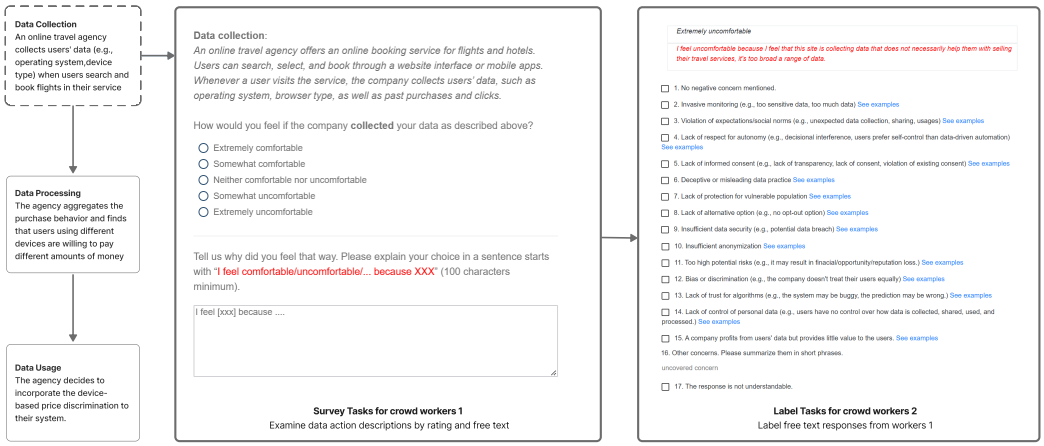

Figure 2: A survey example. Privacy scenario (left) is split into data actions and organized in a information flow. Crowd workers were involved in two different tasks: we first asked crowd workers to examine data action descriptions by rating and writing free text (middle), then we forwarded the collected free-text responses to another group of workers for privacy concerns annotation (right).

prevent potential biases, we distributed scenarios that might give rise to similar concerns across separate surveys. Therefore, we analyzed general concerns in each scenarios using results from Jin et al. [20]. Following their work, we categorized privacy concerns into three high-level classes: respect for persons, beneficence, and justice. Then we identified 12 data applications where concerns from one high-level category were more prevalent than from the other two. Each application was assigned to its corresponding category, and applications representing distinct categories were included in each day's survey.

Among the 13 practices analyzed, we found that in one particular scenario (case 6 in Table 1), the distribution of concerns across the three categories was notably even. This suggests that users have a more diverse range of concerns in this context, potentially demonstrating varied rationales. We observed that users often exhibit specific types of concerns in certain scenarios. For example, concerns about price discrimination are common in situations with evident unfairness in pricing (case 12 and 13 in Table 1). However, the consistency of concerns in such special cases may not extend to more diverse domains. Therefore, we selected case 6 to test for consistency, as illustrated in Figure 1.

Recruitment and Demographics. We conducted the experiment on AMT from May to July 2023. To ensure the quality of survey responses, we only recruited participants with a HIT Approval Rate greater or equal to 95% and Number of HITs Approved greater than or equal to 50, who are aged 18 or above. We also carefully designed our surveys with two pilot studies with 5 participants who were excluded from our official experiment. Considering the amount of data needed, we scheduled to recruit 35 to 40 participants. We ended up recruiting 42 participants located in the United States, and 4 workers were removed from consideration due to failing to finish all five surveys or giving overly uniform answers to a large number of questions in a row.

On average, participants spent 30 minutes on each day's survey and received 30 USD as compensation for the five-day session. We collected the demographic information in our first day's survey to ensure the diversity of our participants. Among all the 38 workers providing valid answers, 24 (63%) identified as female and 14 (37%) identified as male. Participants' age buckets ranged from 25-34 to 65-74, with most reporting to be 36-44 years old (39%). There was also a wide range of reported educational degrees, with most reporting a 4-year Bachelor's degree as their highest degree obtained (37%). Reported income ranged from less than $10,000 to $150,000 or more, with most reporting between $10,000 and $50,000 (58%).

Ethical Considerations. Our project was approved by the IRB at our institute. Participants read and signed an informed consent document before filling out the surveys. We instructed participants to focus on their own experiences and opinions and to not reveal private or sensitive information throughout the surveys. Collected data was stored in a secure location accessed only by the research team. We only collected participants' contact emails for compensating them for their time, did not connect these emails to the rest of the study data, and deleted them after the study completion.

3.2 ContextLabel Codebook

Two authors annotated 43 data actions from 13 scenarios using multiple labels, creating a codebook to capture contextual nuances. The synthesized labels are related to privacy concerns from previous works and applicable across various scenarios, transcending specific domains.

Previous studies have utilized Contextual Integrity [38] with five category factors to model information flow. While not exhaustive, we annotated scenarios following this framework to cover our tested information flows; the factors used are listed in Table 3. In section 4.2 and 4.3, we compared the scenario-specific category factors with non-exclusive labels.

Table 2: 18 selected non-exclusive labels and their definitions used in the annotation and analysis. The labels were synthesized from previous works [7, 15, 19–21, 35, 42, 54] and were derived from the information flow process and its consequences.

Label	Definition
Absence of Consent	Lack of transparency or consent, or violation of existing consent
Algorithmic Assessment Imperfections	Imperfect implementation or adoption of algorithm for assessing personal data
Automated Data-Driven	Loss of initiative due to data-driven automation
Behavioral Data Collection	Users divulge their behavioral data in the scene, which include metadata (e.g. browse history, message history), activity records (e.g. purchase record) and so on
Bio Data Collection	Users divulge their physiology data related to medical, health, or intimacy information
Data Breach	Inadequate data protection measures or unexpected data sharing
Data Control Loss	Loss of control over personal data
Empathy for the Vulnerable	Potential harm for vulnerable populations
Financial Loss	Monetary harm or economic damage
High Risk Probability	The risk is very likely to happen
High Risk Significance	The outcome is severe
Opportunity Loss	Loss of potential opportunities (e.g. promotion, competitive advantage, etc.)
Personal Identifiable Data Collection	Users divulge their personal identifiable information (PII) in the scene (e.g. e-mail address, ID information, etc.)
Price Discrimination	Charging of different prices for the same or similar products or services to different groups of consumers
Reputation Loss	Deterioration of an individual's or an organization's standing or credibility in the eyes of others
Restricted Choices	Lack of an alternative choice, and no opt-out
Third Party Transfer	Data is transferred to third parties
Unexpected Use	Violation of social norms or of expected results

Table 3: Contextual Integrity Factors used in the scenario annotation. The subject element, which is "users" in all tested scenarios, is not listed. The recipient and transmission principle factors are categorized into two categories each. In Table 8, two binary variables (i.e. 'Third Party Transfer' and 'Absence of Consent') are used to represent those two factors.

Category	Value
Sender	Platform, Self, Iot
Attribute	Behavioral Data, Personal Identifiable Data, Bio Data Collection
Recipient	Third Party, Server
Transmission Principle	Absence of Consent, User permission

Instead of using Contextual Integrity factors exclusively, we selected parameters such as 'Third Party Transfer' and 'Bio Data Collection' that have demonstrated impacts on privacy concerns in certain contexts [35, 47]. We also identified a few labels from participants' reported concerns that are related to data processing but failed to be fully captured by contextual integrity labels, such as 'Algorithmic Assessment Imperfections' [21]. Besides the information flow process, studies have shown that in domains like IoT devices, the perceived benefits can significantly impact people's privacy attitudes [5]. We hypothesized that people's privacy concerns can be influenced by attributes associated with consequences of data actions, since they ultimately determine whether the data actions lead to tangible harm to individuals. Therefore, labels such as 'Financial Loss' and 'High Risk Probability' are included. These labels have also been frequently mentioned in privacy-related works [7, 19, 20, 35, 42, 54].

Out of all the identified labels, we selected 18 labels to evaluate their correlation with participants' privacy attitudes and concerns. The selected labels are listed in Table 2.

4 RESULTS

We collected 1,862 valid ratings and free-text responses for 43 distinct data actions in our surveys. The aggregated results for each

scenario can be found in Appendix B. The distribution of participants' comfort or concern levels varies across different context labels, as depicted in Figure 3. Participants' concern categories also display variations, as shown in Figure 4. Notably, some participants (lower portion of Figure 4) exhibited an overall lower level of concern but displayed sensitivity to specific concern categories (i.e. cells on the lower portion of Figure 4 but with warm colors). This highlights the nuanced privacy profiles that may not be captured by generic indices like Westin's index.

Privacy concern refers to an expression of worry towards a specific privacy-related situation [13]. In the following sections, 'privacy attitude' reflects participants' numerical concern ratings for data actions, while 'concern categories' represent the specific types of concerns expressed by participants, as indicated in Figure 7.

4.1 RQ1: Consistency and Rationality

We examined the consistency of participants' own responses of the repeated three data actions and Westin's Index questions. While the consistency test aimed to rule out the possibility of entirely random privacy attitudes among participants, we took a further step by examining the correlation between participants' privacy concerns towards all the tested scenarios. Therefore, we validate our hypothesis that in general, people's privacy attitudes result from their own logical reasoning. Our results suggest that participants' attitudes and concerns toward privacy scenarios exhibit consistency and rationality.

Method. We designed nine questions to test participants' consistency including the repeated scenario assessment (case 6 in Table 1) and Westin's Index [27] across three surveys (see Figure 1). Since the scenario selected for the consistency test comprises three distinct data actions: data collection, processing, and usage, to prevent participant fatigue, we avoided including other cases as redundant scenarios across different surveys.

In each survey, we used each participant's average rating of three data actions in the redundant scenario to represent their own overall attitudes. To gauge the consistency of each participant's attitude, we computed the intraclass correlation coefficient (ICC) for their ratings across the three surveys and calculated Pearson correlation coefficients between all pairwise combinations of the surveys. We employed a similar approach to assess the consistency of their responses to Westin's Index. We also investigated whether participants expressed the same categories of concern regarding scenario 6 across three surveys to assess the consistency of their reasons for discomfort. See Appendix A.1 for additional evaluation details.

For correlation between privacy concern and attitude, we conducted a linear regression analysis on participants' average rating of data actions in the 13 scenarios using their concern categories expressed in the corresponding scenario. The value of each concern category for each participant and scenario is calculated as the sum of the corresponding concern labels in the participant's response to data actions within the specific scenario. Then we looked into the data action level by testing if the concern categories had predictive effects on participants' privacy attitudes. To differentiate participants' positive and negative attitudes, we split the 5-scale comfort rating for each data action into scores below 3 ('somewhat uncomfortable' and 'extremely uncomfortable') as negative and all other scores as positive. We constructed four classification models evaluated on 10-fold cross validation.

Results. For both Westin's questions and ratings for the tested scenario, our findings indicate a strong alignment between users' own ratings across three tests, and their privacy concerns also remained consistent. In the broader analysis for all the scenarios, concern category showed strong correlation with and predictive effects on attitudes, suggesting the participants' rationality exhibited privacy attitudes.

Privacy attitude consistency. The Pearson coefficients indicate a strong correlation across three cases within each measurement for each participant (see Table 4). The average ICC value (Table 5) for the average Likert scores is above 0.75 for case 6 and at least 0.67 for Westin's three questions, suggesting good reliability for all four tests [25]. The majority of participants' general privacy attitudes remained consistent across the three surveys. To identify outliers in the consistency test, we categorized participants' attitudes as either negative or positive, depending on whether their comfort ratings were below 3. Only eight participants exhibited varying attitudes across the three surveys, with five of them displaying a relatively neutral stance, as their ratings fell between 2 and 4. Additionally, one of the three participants with higher variation in rating only assigned negative ratings in the second survey, but the free-text reasoning only presented positive feedback for the news-filtering system in case 6, such as "weeding out anything [they don't] want to see" with similar responses from the other two surveys, so we consider a miss-rating for this case. The analyses of the responses from the only two exceptions are in the following section.

Privacy concern and reasoning consistency. Among all participants, 80.26% of labeled concern categories in scenario 6 remained consistent across all three surveys, suggesting the majority of participants' privacy concerns are consistent over the three surveys. Since concerns expressed towards the scenario varied among individuals, we chose one outlier mentioned in last section and another participant with more consistent concern patterns and visualized their concern categories in Figure 5. The majority of participants demonstrated consistent reasoning across the three surveys, regardless of whether they had negative or positive privacy attitudes. For those who expressed concerns, the most frequently reported issues centered around the lack of trust in algorithms and the lack of control over personal data. The Venn diagram on the right side of Figure 5 presents one typical concern pattern of the majority. On the other hand, participants with more positive perceptions of the scenario constantly referred to it as a "common practice" or "providing benefits for users".

Participants with neutral attitudes exhibited more complex considerations. Two participants (see Figure 5, left) with inconsistent attitudes varied in their benefit assessment across the surveys. Both participants treated the scenario as common practice for Internet companies and mentioned "improving browsing experience" in one survey but expressed the desire for more initiative in another survey. However, for those outliers, their own detailed reasoning often covered consistent themes across three surveys. For instance, despite expressing different attitudes in two surveys, in the third

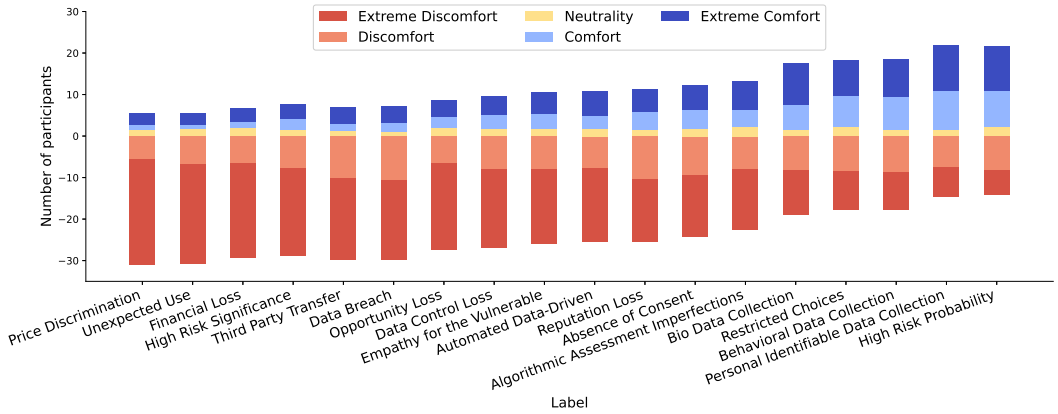

Figure 3: Average participant distribution across different comfort levels for each label. From left to right, labels are ranked in increasing order of average comfort score. A lower bar represents a lower proportion of participants expressing a positive attitude towards data actions with the label.

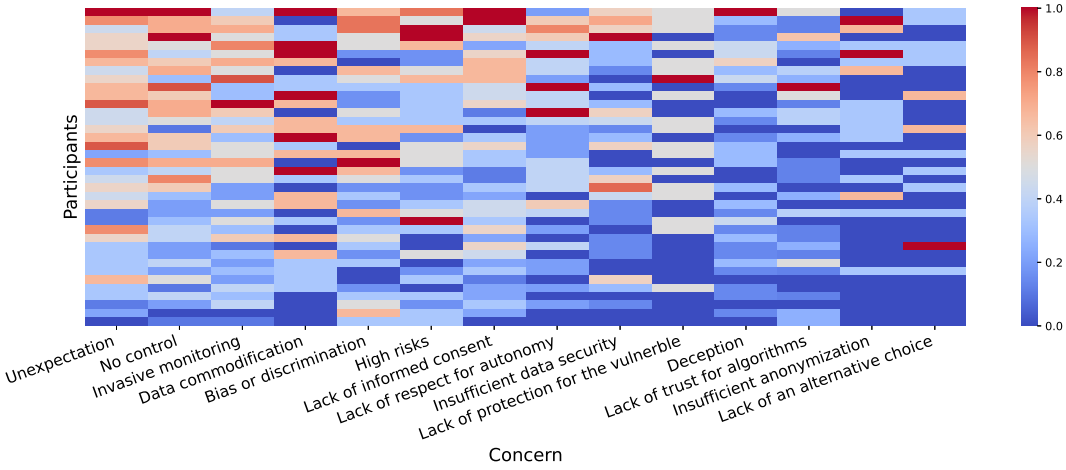

Figure 4: Concerns aggregated from participants. A higher value in the map means the participant expressed the concern category more frequently in surveys. Concerns are ordered left to right by the frequency in all participants' responses, in decreasing order. Participants are ordered top to bottom by the numbers of expressed concerns across all scenarios, in decreasing order.

Table 4: Pearson's correlation of average ratings of the redundant scenario (Case 6) and three questions from Westin's Index (WQ) between surveys from all participants. All correlations are significant at the 0.01 level.

Case 6				WQ1			
	survey 1	survey 3	survey 5		survey 1	survey 3	survey 5
survey 1	1.00			survey 1	1.00		
survey 3	0.70	1.00		survey 3	0.78	1.00	
survey 5	0.86	0.84	1.00	survey 5	0.61	0.65	1.00

WQ2				WQ3			
	survey 1	survey 3	survey 5		survey 1	survey 3	survey 5
survey 1	1.00			survey 1	1.00		
survey 3	0.81	1.00		survey 3	0.65	1.00	
survey 5	0.77	0.83	1.00	survey 5	0.73	0.80	1.00

Table 5: Intraclass Correlation Coefficient (ICC) of users' average ratings of the redundant scenario (Case 6) and three questions from Westin's Index (WQ). Results are all significant with p-values at 0.001 level.

	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0		
		Lower Bound	Upper Bound	Value	df1	df2
Case 6	0.80	0.69	0.88	12.98	37	74
WQ1	0.67	0.52	0.80	7.42	37	74
WQ2	0.81	0.70	0.89	13.28	37	74
WQ3	0.73	0.59	0.84	8.95	37	74

survey, participant P1 described news filtering as "able to save time" but "I don't like a third party hiding content from people or businesses that I am willingly following, which I decided to follow because I want to see those updates". This suggests similar factors to assess the same privacy context within a short time frame. This participant's Venn diagram (Figure 5, left) also shows a certain degree of consistency, as there is only one concern category in the non-intersecting area.

Privacy concern and privacy attitude rationale. Apart from the shared concerns among participants with negative attitudes towards scenario 6, Figure 5 illustrates the average rating tends to decrease as participants express more concern categories in each survey, corroborating our regression findings in Table 6. Most concern categories negatively affect participants' ratings across all scenarios, indicating their role in explaining participants' attitudes towards specific contexts. This correlation is further confirmed by the promising predictive effect (with prediction accuracy of 87%) of concerns on privacy attitudes, as demonstrated in Table 7. We also introduced data action types (data collection, processing, sharing, and usage) as new category variables into the prediction model, but it did not significantly alter the model's performance compared to the models in Table 7. This suggests that concern categories' predictive effect is broad and not limited to specific contextual actions.

Additionally, the regression coefficient in Table 6 shows that concern categories exhibit varying degrees of influence, with categories like 'Bias or Discrimination' having notably higher coefficients, indicating their stronger impact on participants' attitudes.

4.2 RQ2: Correlation between ContextLabel and Privacy Concerns

To validate if ContextLabel can effectively capture the essence of privacy contexts, we analyzed labels' correlation with participants' privacy attitudes and concern categories. Our results suggest ContextLabel exhibits stronger correlations with participants' comfort ratings and concern categories compared to generic privacy index and category factors.

Method. To measure the correlation between ContextLabel and comfort rating for data actions, we calculated their Pearson correlation coefficient and Kendall rank correlation in all scenarios as shown in Table 8. The Pearson correlation is calculated as Point-biserial correlation, a special case of the Pearson Correlation to measure the relationship between continuous variables and dichotomous variables. For comparison, we annotated the labels associated with Contextual Integrity elements and added the sender parameter. We also computed the Pearson correlation and Kendall rank correlation using responses to Westin's three questions as shown in Table 9. To address individual differences, we assessed the frequency of a ContextLabel appearing among the top five labels with the highest Kendall rank correlation to individual comfort ratings, as depicted in Figure 6.

For the correlation between ContextLabel and concern categories, we defined a concern score C_l to gauge the extent to which labels contribute to general levels of privacy concern. See Appendix A.2 for more evaluation details. Figure 7 shows the overall results. We also calculated the odds ratio between each label and the expressed concern categories for each user, indicating the strength of correlation in an individual's profile. Sets of concern-label pairs with odds ratios greater than 10 and significance levels below 0.05 were identified. The occurrence of each set across all user cases was then counted, reflecting the transferability of specific concerns across scenarios with that label. Figure 8 displays the top 20 sets with the highest occurrences.

In section 4.2 and 4.3, for the redundant scenario 6 included in three surveys, we only used the responses from the third day's survey since section 4.1 already showed the responses to be consistent across the surveys.

Results. Overall, non-exclusive labels like 'High Risk Significance', 'Price Discrimination', and 'Financial Loss' demonstrated stronger correlations with participants' comfort ratings and concern categories compared to exclusive categories factors and the Privacy Segmentation Index. This underscores the effectiveness of ContextLabel in capturing crucial aspects of diverse privacy contexts and representing individuals' perceptions.

Correlations between ContextLabel & Comfort Score. People exhibit varying sensitivities to different labels, but some non-exclusive ContextLabels have a noticeable impact on the majority's privacy attitudes.

The correlation between ratings and Westin's question (Table 9) is notably weaker compared to that of context labels (Table 8). However, many Pearson coefficient values for the labels do not indicate a strong correlation with ratings. Compared with the predictive effect of concern categories on individual's attitudes (Table 7), individual variance in concern categories towards the same ContextLabel could explain this. For instance, the Pearson correlation between 'Price Discrimination' and comfort rating ranges from -0.651 to 0.007 among participants, with Kendall correlations ranging from -0.595 to 0.015. Despite the individual variation, non-exclusive labels like 'Unexpected Use', 'High Risk Significance', 'Price Discrimination', and 'Financial Loss' demonstrated stronger correlations with

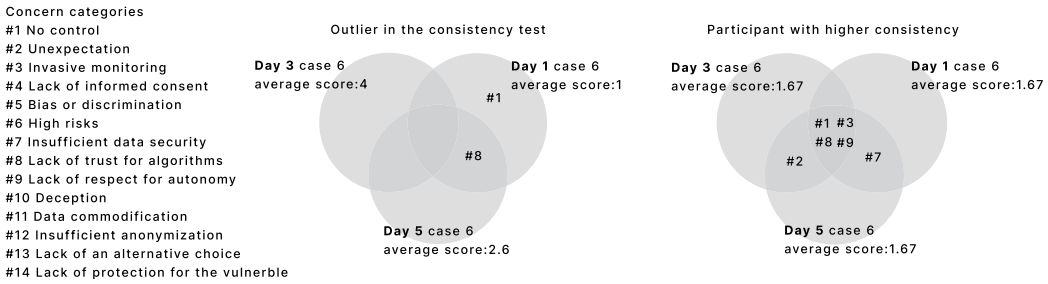

Figure 5: Venn diagram of examples of two participants with low or high concern consistency. Concern categories absent in the circles represent the ones unexpressed by the corresponding participant across all three surveys.

Table 6: Linear regression on the average scores of scenarios. To represent participants' attitudes toward each scenario, the evaluation pertains to data at the complete scenario level rather than split data actions. The value of each concern category is calculated as the sum of the corresponding concern labels in each participant's response to data actions within the specific scenario, and used as continuous variable. The reported coefficients are unstandardized.* indicates < .05 statistical significance and ** indicates < .001.

Concern Category	Coefficients	Std. Error
Lack of trust for algorithms	-.311**	.084
Lack of an alternative choice	-.186	.210
Insufficient anonymization	-.199	.176
Lack of respect for autonomy	-.262**	.107
Bias or discrimination	-.645**	.073
Insufficient data security	-.202*	.099
Deception	-.203	.116
Lack of informed consent	-.243**	.071
Invasive monitoring	-.443**	.063
Data commodification	-.536**	.127
No control	-.204*	.068
High risks	-.175*	.081
Unexpectation	-.218**	.067
Lack of protection for the vulnerable	-.613*	.212

Table 7: Accuracy, F1 score, and recall of Logistic regression (Logistic), Support Vector Classification (SVC), AdaBoost classifier (AdaBoost) and k-nearest neighbors classifier (KNN) on attitude prediction using concern categories (binary variables). The average results of 10 folds are reported with standard deviation. The evaluation pertains to data at the data action level.

Model	Accuracy(Std)	F1 Score(Std)	Recall(Std)
Logistic	0.86(0.02)	0.85(0.03)	0.89(0.03)
SVC	0.87(0.02)	0.86(0.03)	0.88(0.03)
AdaBoost	0.86(0.03)	0.85(0.03)	0.91(0.03)
KNN	0.8(0.13)	0.69(0.33)	0.72(0.35)

participants' comfort ratings than category factors from Contextual Integrity. These labels were also frequently influential to individuals' comfort levels (see Figure 6), and they received lower average ratings (Figure 3). Labels with notable impacts are discussed in the following section.

Correlations between ContextLabel & Concern Categories. Though privacy concerns towards the same ContextLabel vary among individuals, ContextLabel is able to capture more transferable concern categories than category factors and the Privacy Segmentation Index.

Figure 7 reveals variations in how participants associate labels with specific concern categories on average. Contrarily, Figure 8 illustrates the concern-label sets that show strong correlation in individual participants' data. Despite diverse individual concerns shown in Figure 4, Figure 8 shows approximately 50% of participants closely link 'Bias or Discrimination' to six ContextLabels. Labels

Table 8: Pearson correlation (calculated as Point-biserial correlation) and Kendall rank correlation between ContextLabel and comfort rating. Labels are treated as binary variables and ratings as continuous. Labels marked with CI are Contextual Integrity factors (see detailed definition in Table 3).

ContextLabel	Pearson Correlation	p value	Kendall Rank Correlation	p value
Empathy for the Vulnerable	-0.238	0.000	-0.223	0.000
High Risk Probability	0.336	0.000	0.309	0.000
Restricted Choices	0.15	0.000	0.139	0.000
Algorithmic Assessment Imperfections	-0.107	0.040	-0.102	0.000
Automated Data Driven	-0.216	0.000	-0.203	0.000
Data Breach	-0.201	0.000	-0.179	0.000
Data Control Loss	-0.271	0.000	-0.253	0.000
Financial Loss	-0.573	0.000	-0.46	0.000
Reputation Loss	-0.125	0.000	-0.108	0.000
High Risk Significance	-0.296	0.000	-0.277	0.000
Opportunity Loss	-0.23	0.000	-0.221	0.000
Price Discrimination	-0.457	0.000	-0.247	0.000
Unexpected Use	-0.412	0.000	-0.496	0.000
Absence of Consent(CI)	-0.226	0.000	-0.199	0.000
Behavioral Data Collection(CI)	0.174	0.000	0.16	0.000
Personal Identifiable Data Collection(CI)	0.197	0.000	0.176	0.000
Third Party Transfer(CI)	-0.17	0.000	-0.153	0.000
Bio Data Collection(CI)	0.069	0.005	0.062	0.006
Sender-Platform(CI)	0.023	0.356	0.03	0.182
Sender-Self(CI)	-0.141	0.030	-0.138	0.000
Sender-Iot(CI)	0.156	0.010	0.142	0.000

Table 9: Pearson correlation and Kendall rank correlation between participants' responses to three questions from Westin's Index (WQ) and their comfort rating towards data actions.

	Pearson Correlation	p value	Kendall Rank Correlation	p value
WQ1	-0.068	0.006	-0.059	0.003
WQ2	0.092	0	0.06	0.002
WQ3	0.084	0.001	0.056	0.004

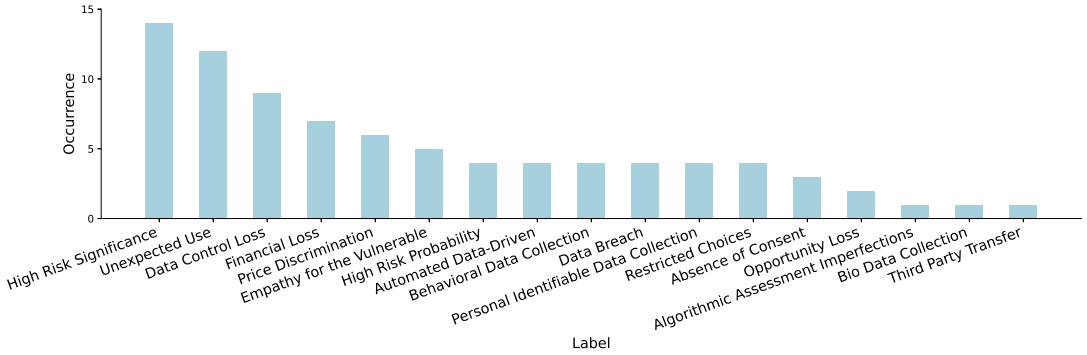

Figure 6: Influential labels for individuals' comfort level. When a label ranks in the top five based on Kendall rank correlation coefficients with individual user's comfort ratings among all labels, and the p value is less than 0.05, then it is considered influential. The y-axis represents the frequency of labels being labeled as influential across all users' cases.

like 'Third party transfer' and 'Data Breach' also connect to concern categories, aligning with Figure 7. This suggests ContextLabel's potential in pinpointing primary sources of concern. Moreover, Figure 8 underscores that category factors alone are inadequate for

capturing transferable concerns, since concern categories exhibit significant correlations with non-exclusive labels such as 'Financial Loss' and 'Unexpected Use', but not collected information type or sender type. Among those individuals who expressed 'Bias or Discrimination' concern in various scenarios, their responses to each question in Westin's index in our survey spanned the entire range of scales from 1 to 4 without any discernible specific patterns, indicating that the general criteria used in Westin's index failed to adequately capture the concerns of participants within specific contextual scenarios.

Influential labels. While Figure 8 only shows the top 20 of 252 concern-label sets, we identified a total of 111 sets with high odds ratio in individuals' profiles, of which the 'High Risk Significance', 'Algorithmic Assessment Imperfections', 'Empathy for the Vulnerable', 'Financial Loss', 'Opportunity Loss', and 'Unexpected Use' labels show the highest frequency, each of them appearing in nine or ten sets. Among those labels, 'High Risk significance', 'Financial Loss', 'Price Discrimination', and 'Unexpected Use' are also closely related with participants' discomfort based on results from Table 8 and Figure 3. Notably, those are labels that determine whether the data actions lead to tangible harm or pose threats to individuals. This validates our hypothesis that these types of labels are very likely to arouse concern and therefore influence privacy attitudes. In addition, participants showed less concern for 'Opportunity Loss' and 'Reputation Loss,' compared to 'Financial Loss', though all lead to potential harm. This disparity suggests that individuals prioritize tangible harms, such as 'Financial Loss', over more abstract or latent consequences like reputation or opportunity loss. In contrast, most category factors, such as the attributes (i.e., collected data types) in Contextual Integrity frame, did not display significant correlations with privacy concerns. However, labels synthesized from Contextual Integrity, like 'Third Party Transfer,' exhibited correlations with specific concern categories. This implies that non-exclusive ContextLabels are more proficient at capturing the aspects of privacy contexts that genuinely concern people.

4.3 RQ3: Prediction Modeling

Participants' rationality for privacy attitudes (RQ1) and the effectiveness of ContextLabel in capturing dominant concern categories (RQ2) suggest ContextLabel's potential predictive effects on participants' privacy attitudes towards unseen data actions. The assumption is supported by the results in this section.

Method. In this study, we framed both predicting concern categories and privacy attitudes as classification tasks. We built the ContextLabel prediction model using 18 labels and examined the predictive effect of ContextLabel on individuals' concern and privacy attitudes. We adopted Contextual Integrity category factors (see Table 3) and Westin's Privacy Segmentation Index as our baselines. As Westin's index has been found to be ineffective in predicting users' privacy contextual attitudes or concerns [52], we used each participant's average scores in three tests for Westin's three questions to build the prediction model. For ContextLabel, we trained a Naive Bayes classifier which takes ContextLabels as input and predicts whether a particular user has certain concern categories. We evaluated the model predictions with leave-one-out cross validation (LOOCV). To explore privacy attitude prediction, we used the same threshold as Table 7 in RQ1 to differentiate positive and negative attitudes, thus making attitude prediction a binary classification. We built a neural network in the form of two-layer multi-layer perceptrons (MLP) to model participants' decision process. To test the prediction effect on novel scenarios (i.e. data actions with novel label combinations), we evaluated models using cross-validation where in each fold, data actions serving as the test set were excluded from the training set. All the models were built using the Scikit-learn package.

Results. Combined with personal preferences, ContextLabel shows overall better predictive effects on privacy attitude and concern categories than category factors and Segmentation Index.

Concern category prediction towards unseen data action. In our survey, we only considered prominent categories, requiring consensus from at least two of three label workers for each free-text response. This resulted in sparse individual-level concern distribution and overall high prediction accuracy of models. To gauge model performance, we emphasize recall scores in Table 10, focusing on the model's ability to identify existing concerns. Figure 9 displays model performance across all participants.

Notably, ContextLabel outperforms the other models, especially in categories such as 'Bias or discrimination,' 'Unexpectation,' and 'Invasive monitoring,' which are among the top 5 most expressed concerns. The low average recall score is attributed to sparse data arising from infrequent expressions or divergent decisions among crowd workers for specific concerns, such as 'Lack of Informed Consent,' 'Lack of Respect for Autonomy,' and 'No Control.' Notably, the 'No Control' category had 652 annotations, with the majority (64.5%) contributed by a single worker, leading to less than 40% of retained concern labels and consequently lower recall scores.

Privacy attitude prediction towards unseen data practice. Table 11 displays the model performances. The results suggest that for predictions on the individual level, the ContextLabel model significantly improved overall attitude prediction accuracy (73%) than category factors (59%) and Privacy Segmentation Index (56%). Notably, there is also an increase in the recall and F1-score, which is around 20% higher than those achieved by the Contextual Integrity model trained without individual preference specification. This indicates a noteworthy predictive effect on people's privacy attitudes towards unfamiliar scenarios when combining both contextual information and personal preferences. For the ContextLabel model trained on individual data, the top three mispredicted data actions all belong to cases annotated with only four context labels, lacking influential labels such as 'Empathy for the Vulnerable', 'Financial Loss', 'High Risk Significance', 'Price Discrimination', and 'Unexpected Use', as illustrated in Figure 6. Furthermore, these actions are either categorized as data collection or processing, but they do not represent the final actions that directly lead to consequences. Participants tend to exhibit diverse attitudes toward these actions. For those three cases, approximately 50% express positive sentiments and the remaining 50% express negative ones.

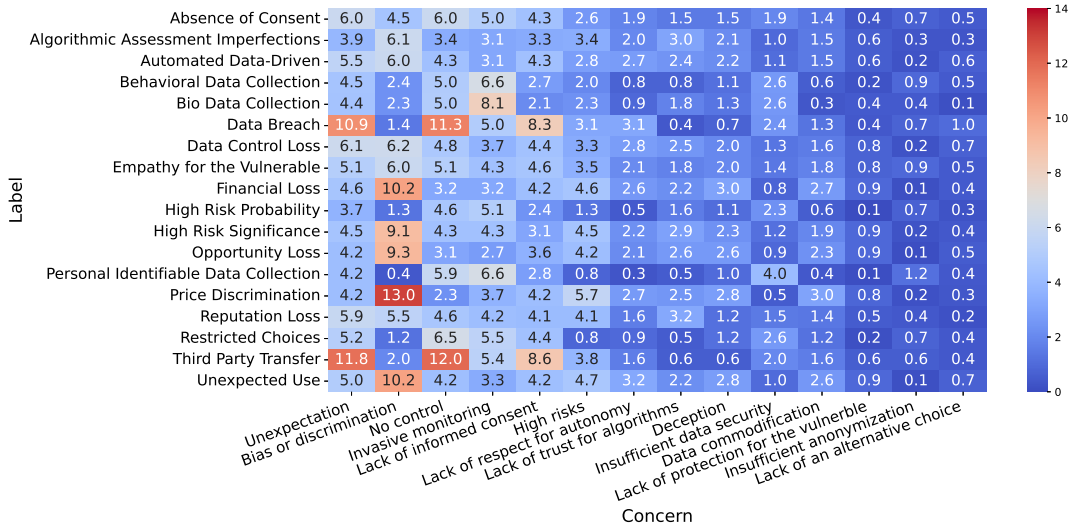

Figure 7: The average number of participants who expressed a specific concern category for each label. Concerns are ordered left to right by the sum of each column. We excluded the sender type from the Contextual Integrity factors as it did not exhibit high odds ratios with concern categories as labels in Figure 8 (odds ratio > 10 with p value < 0.05), suggesting weak correlation.

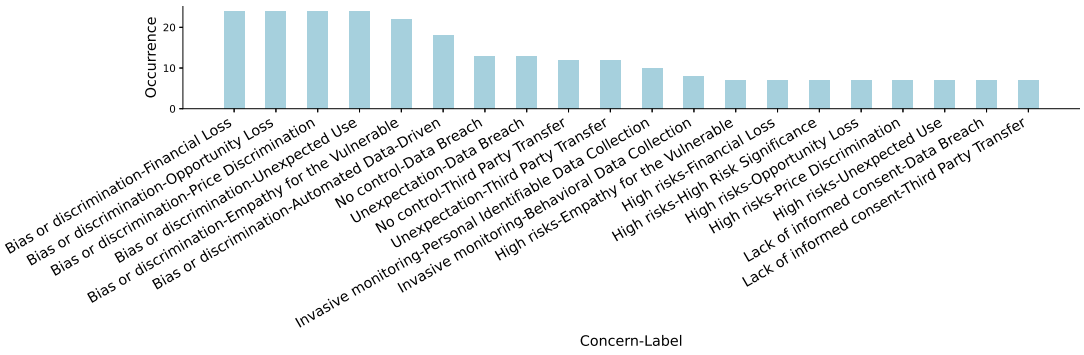

Figure 8: The occurrences of the top 20 concern-label sets with the highest impact (with odds ratios > 10 and p value < 0.05 in individual participants' data). From left to right, labels are ranked in descending order based on how frequently they appeared in various users' cases.

Table 10: Average Recall and Average Accuracy with standard devision in the parenthesis. The reported results are the average value across all participants' profile.

Model	Recall(Std.)	Accuracy
ContextLabel	0.46(0.12)	0.90(0.04)
Contextual integrity	0.26(0.21)	0.93(0.04)
Westin's Index	0.24(0.22)	0.93(0.04)

5 CROSS-CHECKING WITH EXISTING DATASET

Before carrying out our surveys, we analyzed the available survey results from Shvartzshnaider et al.'s study [47] to validate the predictability and similarity of users' privacy attitudes in similar scenarios. Table 12 presents the average accuracies of binary logistic models, SVM, and k-nearest neighbors classifiers using Leave-One-Out Cross-Validation (LOOCV) on the dataset from Shvartzshnaider et al.'s study [47].

The user models achieved an average accuracy of 71.04%. Despite variations in scenario descriptions, they share an educational context and limited contextual integrity labels, indicating their

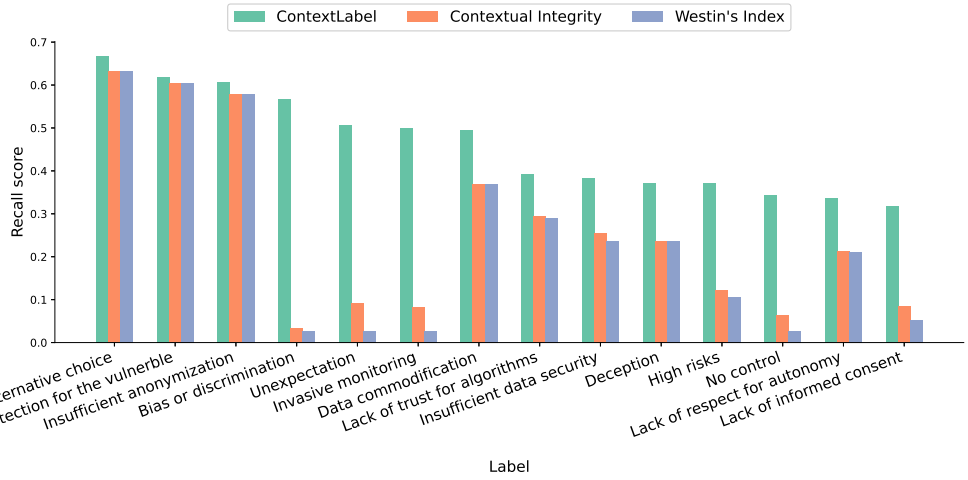

Figure 9: The average recall score of all participants' expressed concern on each concern category and categories are ordered left to right by the recall score, in decreasing order. ContextLabel shows significantly higher score in most concern categories.

Table 11: The average prediction accuracy, recall and F1-score of the MLP models, with the thresholds of 3 to differentiate the positive and negative attitude. Individual scope suggests the models are trained and evaluated on individual users' data. All users scope refers to the cross-validation where the models were trained on all users' data, and each story served as the test set in each fold.

Model	Training Scope	Accuracy	Recall	F1-score
ContextLabel	individual	**0.73**	**0.71**	**0.66**
Contextual Integrity	individual	0.59	0.56	0.52
Westin's Index	individual	0.56	0.37	0.27
ContextLabel	all users	0.64	0.64	0.59
Contextual Integrity	all users	0.51	0.52	0.42
Westin's Index	all users	0.59	0.42	0.47

similarity. The observed accuracies suggest that users' privacy preferences can be applied across comparable contexts and used to predict their attitudes in similar scenarios. These findings align with our observations that users' privacy attitudes are influenced by their rational reasoning and potential contextual predictors can be employed to model concerns and predict attitudes.

6 DISCUSSION

Our results illustrate the diversity of individuals' privacy attitudes and concern categories across various contexts. Nonetheless, individuals consistently apply their own logic and exhibit relatively stable reasoning for their expressed privacy attitudes. This provides opportunities to model their decision-making processes by identifying the factors that raise their awareness when they encounter privacy issues. We identified several non-exclusive ContextLabels and compared them with category factors and generic indices like Westin's Index. ContextLabel proves to have stronger correlations with individuals' privacy concerns and attitudes towards specific contexts. While not exhaustive, our non-exclusive labels effectively predict people's privacy attitudes, demonstrating their predictive feasibility in real-world scenarios.

User rationality behind privacy attitudes. Participants' consistency of their expressed attitudes toward the same contexts and the strong correlation between their comfort rating and concern categories demonstrate their rationality behind the privacy attitudes. Even for concern categories that do not typically appear in participants' responses (see the right portion of Figure 4), specific participants still pay special attention to them. For instance, one participant who reported 'Lack of Alternative Choice' the most expressed this as a reason for the discomfort in three specific scenarios. When explaining the reasons for the high comfort score in other scenarios, however, the same participant consistently used a similar expression, emphasizing the importance of "having the choice to opt out".

ContextLabel associated with risks and benefit assessment. Our results align with the assumption from previous works that users' decisions and actions are propelled by their intent to optimize their benefits [29]. We found that labels associated with the final outcomes and tangible harms (e.g., 'Financial Loss') rather than information collection or processing have stronger correlation with

Table 12: Average accuracy of three types of prediction models across all users in each survey from a separate study

Survey Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
KNN	0.69	0.64	0.66	0.71	0.66	0.73	0.67	0.68	0.79	0.73	0.70	0.78	0.68	0.77	0.68	0.73
Logistic	0.70	0.63	0.66	0.71	0.67	0.74	0.67	0.69	0.80	0.72	0.70	0.78	0.66	0.78	0.70	0.74
SVC	0.70	0.65	0.66	0.71	0.68	0.74	0.67	0.69	0.79	0.73	0.70	0.78	0.67	0.78	0.69	0.74

participants' privacy attitudes and concerns. Besides participants whose concerns are correlated with those labels as analyzed in section 4.2, the borderline also suggests a similar conclusion. For example, despite the general negative attitude towards the a scenario where iOS users may receive a special unadvertised discount and Chrome or IE users may be charged an extra $50 (i.e. case 12 in Table1), one participant provided positive feedback, as the participant mentioned they were an iOS user and would benefit from this. This finding aligns with recent work [5, 45] that individuals' privacy attitudes can be influenced by the perceived benefits.

Predictive effect of ContextLabel combining personal preferences. ContextLabel achieved significantly better predictive effect on participants' privacy attitudes and concerns towards unseen data practice than our baseline. The prediction effect is further improved when the model is built upon individuals' existing data, which combines the ContextLabel and participants' personal preferences. The results are in line with the cross-checking results on Shvartzshnaider et al.'s study [47], indicating the possibility of predicting people's privacy attitude leveraging the scenario similarity and personal preferences. Though both contextual integrity category factors and ContextLabel showed the potential of capturing contextual nuances, ContextLabel synthesized efficient factors, including those from the contextual integrity framework, achieving promising prediction results for privacy attitudes while simplifying the modeling process due to non-exclusive labels across various domains.

Protential Applications of ContextLabel. Our results regarding correlation and prediction accuracy show that the current labels have transitivity across diverse contexts. Developers can utilize these predictions to provide users with customized privacy management tools. This enables the creation of default privacy settings that align with individual user preferences. For instance, in the HCI domain, the ContextLabel's insights could guide the development of Contextual Privacy Policy (CPP) tools [51] by highlighting which factors users are most likely to be concerned about in particular situations while reducing unnecessary notifications. Furthermore, for in-depth research into privacy within specific contexts, researchers can enhance the detail by expanding the label set. They can also re-label user feedback for these specific contexts without negating the validity of previous results.

7 LIMITATIONS AND FUTURE WORK

While our experiments have shown participants' consistency in a short time frame, future research should consider testing over a longer duration to gain a more comprehensive understanding of people's rationality in privacy scenarios. Furthermore, many studies have been done on the privacy paradox [18] to investigate the inconsistency between behavior and intent. The insights into rationality from our study can assist in modeling people's privacy attitudes, but there is still a need for extensive research to fully grasp the intricacies of the privacy paradox. Users' decisions are influenced by various biases, so our predictions should be used with caution in certain policymaking situations.

We defined 18 labels to test their correlation with people's privacy attitude and concerns. A wider range of labels could be identified and used in a future study to provide a more comprehensive method to model users' perception and decision processes in privacy scenarios. Researchers could give labels more fine-grained attributes or scores to capture more nuances of contexts, leading to more accurate concern predictions. Additionally, though the annotation in our study was completed by two professional annotators, the heuristic-like annotation process could be further explored leveraging large language models.

Our paper is an exploratory work that studies the feasibility of predicting users' privacy concerns across domains. Future research could expand by incorporating additional scenarios to enhance the generalization performance of the prediction model. While our survey design, including features like free-text explanation and the consistency test, addresses concerns related to crowd workers' inattention on AMT, further validation of ContextLabelcould involve a more diverse selection of participants from various platforms in future research.

8 CONCLUSION

We presented ContextLabel, a novel method for capturing users' privacy profiles across domains and predicting their privacy attitudes towards unseen data practices. By incorporating non-exclusive labels and users' preferences, ContextLabel offers a more accurate modeling of privacy concerns compared to categorical factors and generic index. The results of our empirical study involving 38 participants over five days demonstrated the feasibility of predicting users' privacy concerns across domains. We observed consistent privacy attitudes among participants and identified contextual labels that correlated with users' privacy concerns. Leveraging these insights, we built a predictive model which achieved a higher accuracy (73%) compared to the Privacy Segmentation Index (56%) and categorical contextual factors (59%).

ACKNOWLEDGMENTS

We would like to thank our study participants for their kind participation, without whom this work would not have been possible. We are also grateful to our anonymous reviewers for their insightful feedback. In addition, we sincerely thank Zirui Cheng, Xiaomeng Xu and Jixuan He for their constant support.

REFERENCES

[1] Noura Abdi, Xiao Zhan, Kopo M Ramokapane, and Jose Such. 2021. Privacy norms for smart home personal assistants. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.

[2] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2020. Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology* 30, 4 (2020), 736–758.

[3] Alessandro Acquisti, Stefanos Gritzalis, Costos Lambrinoudakis, and Sabrina di Vimercati. 2007. *Digital privacy: theory, technologies, and practices*. CRC Press.

[4] Alessandro Acquisti and Jens Grossklags. 2005. Privacy and rationality in individual decision making. *IEEE security & privacy* 3, 1 (2005), 26–33.

[5] Ahmed Alhazmi, Ghassen Kilani, William Allen, and TJ OConnor. 2021. A Replication Study for IoT Privacy Preferences. In *2021 IEEE International Conference on Omni-Layer Intelligent Systems (COINS)*. 1–8. https://doi.org/10.1109/COINS51742.2021.9524236

[6] Ahmed Alhazmi, Ghassen Kilani, William Allen, and TJ OConnor. 2021. A replication study for iot privacy preferences. In *2021 IEEE International Conference on Omni-Layer Intelligent Systems (COINS)*. IEEE, 1–8.

[7] Ashwaq Alsoubai, Reza Ghaiumy Anaraky, Yao Li, Xinru Page, Bart Knijnenburg, and Pamela J Wisniewski. 2022. Permission vs. App Limiters: Profiling Smartphone Users to Understand Differing Strategies for Mobile Privacy Management. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.

[8] Mary Jean Amon, Rakibul Hasan, Kurt Hugenberg, Bennett I Bertenthal, and Apu Kapadia. 2020. Influencing photo sharing decisions on social media: A case of paradoxical findings. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 1350–1366.

[9] Mary Jean Amon, Aaron Necaise, Nika Kartvelishvili, Aneka Williams, Yan Solihin, and Apu Kapadia. 2023. Modeling User Characteristics Associated with Interdependent Privacy Perceptions on Social Media. *ACM Transactions on Computer-Human Interaction* (2023).

[10] Noah Apthorpe, Yan Shvartzshnaider, Arunesh Mathur, Dillon Reisman, and Nick Feamster. 2018. Discovering smart home internet of things privacy norms using contextual integrity. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 2, 2 (2018), 1–23.

[11] Natã Miccael Barbosa, Joon S Park, Yaxing Yao, and Yang Wang. 2019. " What if?" Predicting Individual Users' Smart Home Privacy Preferences and Their Changes. *Proc. Priv. Enhancing Technol.* 2019, 4 (2019), 211–231.

[12] Devasheesh P Bhave, Laurel H Teo, and Reeshad S Dalal. 2020. Privacy at work: A review and a research agenda for a contested terrain. *Journal of Management* 46, 1 (2020), 127–164.

[13] Jessica Colnago, Lorrie Faith Cranor, Alessandro Acquisti, and Kate Hazel Stanton. 2022. Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. 331–346.

[14] Cailing Dong, Hongxia Jin, and Bart Knijnenburg. 2015. Predicting privacy behavior on online social networks. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 9. 91–100.

[15] Denzil Ferreira, Vassilis Kostakos, Alastair R Beresford, Janne Lindqvist, and Anind K Dey. 2015. Securacy: an empirical investigation of Android applications' network usage, privacy and security. In *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. 1–11.

[16] Christian Flender and Günter Müller. 2012. Type indeterminacy in privacy decisions: the privacy paradox revisited. In *Quantum Interaction: 6th International Symposium, QI 2012, Paris, France, June 27-29, 2012, Revised Selected Papers 6*. Springer, 148–159.

[17] Alisa Frik, Juliann Kim, Joshua Rafael Sanchez, and Joanne Ma. 2022. Users' expectations about and use of smartphone privacy and security settings. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–24.

[18] Nina Gerber, Paul Gerber, and Melanie Volkamer. 2018. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & security* 77 (2018), 226–261.

[19] Hannah J Hutton and David A Ellis. 2023. Exploring User Motivations Behind iOS App Tracking Transparency Decisions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–12.

[20] Haojian Jin, Hong Shen, Mayank Jain, Swarun Kumar, and Jason I Hong. 2021. Lean privacy review: Collecting users' privacy concerns of data practices at a low cost. *ACM Transactions on Computer-Human Interaction (TOCHI)* 28, 5 (2021), 1–55.

[21] Smirity Kaushik, Yaxing Yao, Pierre Dewitte, and Yang Wang. 2021. " How I Know For Sure": People's Perspectives on Solely Automated Decision-Making. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 159–180.

[22] Aniket Kittur, Ed H Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 453–456.

[23] Bart P Knijnenburg, Alfred Kobsa, and Hongxia Jin. 2013. Dimensionality of information disclosure behavior. *International Journal of Human-Computer Studies* 71, 12 (2013), 1144–1162.

[24] Spyros Kokolakis. 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security* 64 (2017), 122–134.

[25] Terry K Koo and Mae Y Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* 15, 2 (2016), 155–163.

[26] Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24 (2014), 8788–8790.

[27] Ponnurangam Kumaraguru and Lorrie Faith Cranor. 2005. Privacy indexes: a survey of Westin's studies. (2005).

[28] Hosub Lee and Alfred Kobsa. 2017. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 276–285. https://doi.org/10.1109/PERCOM.2017.7917874

[29] Namyeon Lee and Ohbyung Kwon. 2015. A privacy-aware feature selection method for solving the personalization–privacy paradox in mobile wellness healthcare services. *Expert systems with applications* 42, 5 (2015), 2764–2771.

[30] Tianshi Li, Kayla Reiman, Yuvraj Agarwal, Lorrie Faith Cranor, and Jason I Hong. 2022. Understanding challenges for developers to create accurate privacy nutrition labels. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–24.

[31] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM conference on ubiquitous computing*. 501–510.

[32] 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 *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. 199–212.

[33] Bin Liu, Jialiu Lin, and Norman Sadeh. 2014. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help?. In *Proceedings of the 23rd international conference on World wide web*. 201–212.

[34] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[35] Kirsten Martin and Helen Nissenbaum. 2016. Measuring privacy: An empirical test using context to expose confounding variables. *Colum. Sci. & Tech. L. Rev.* 18 (2016), 176.

[36] Kirsten Martin and Katie Shilton. 2016. Why experience matters to privacy: How context-based experience moderates consumer privacy expectations for mobile applications. *Journal of the Association for Information Science and Technology* 67, 8 (2016), 1871–1882.

[37] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. 2017. Privacy expectations and preferences in an IoT world. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association Santa Clara, 399–412.

[38] Helen Nissenbaum. 2020. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

[39] Scott Plous. 1993. *The psychology of judgment and decision making*. Mcgraw-Hill Book Company.

[40] Franziska Roesner, Tadayoshi Kohno, Alexander Moshchuk, Bryan Parno, Helen J. Wang, and Crispin Cowan. 2012. User-Driven Access Control: Rethinking Permission Granting in Modern Operating Systems. In *2012 IEEE Symposium on Security and Privacy*. 224–238. https://doi.org/10.1109/SP.2012.24

[41] Rafiy Saleh, Dawn Jutla, and Peter Bodorik. 2007. Management of Users' Privacy Preferences in Context. In *2007 IEEE International Conference on Information Reuse and Integration*. IEEE, 91–97.

[42] Stuart Schechter and Cristian Bravo-Lillo. 2014. Using ethical-response surveys to identify sources of disapproval and concern with Facebook's emotional contagion experiment and other controversial studies. (2014).

[43] Norbert Schwarz, Herbert Bless, Fritz Strack, Gisela Klumpp, Helga Rittenauer-Schatka, and Annette Simons. 1991. Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social psychology* 61, 2 (1991), 195.

[44] Marc Serramia, William Seymour, Natalia Criado, and Michael Luck. 2023. Predicting Privacy Preferences for Smart Devices as Norms. *arXiv preprint arXiv:2302.10650* (2023).

[45] Tanusree Sharma, Smirity Kaushik, Yaman Yu, Syed Ishtiaque Ahmed, and Yang Wang. 2023. User Perceptions and Experiences of Targeted Ads on Social Media Platforms: Learning from Bangladesh and India. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.

[46] Fuming Shih, Ilaria Liccardi, and Daniel Weitzner. 2015. Privacy Tipping Points in Smartphones Privacy Preferences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 807–816.

https://doi.org/10.1145/2702123.2702404
[47] Yan Shvartzshnaider, Schrasing Tong, Thomas Wies, Paula Kift, Helen Nissenbaum, Lakshminarayanan Subramanian, and Prateek Mittal. 2016. Learning privacy expectations by crowdsourcing contextual informational norms. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 4. 209–218.
[48] H Jeff Smith, Sandra J Milberg, and Sandra J Burke. 1996. Information privacy: Measuring individuals' concerns about organizational practices. *MIS quarterly* (1996), 167–196.
[49] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *science* 211, 4481 (1981), 453–458.
[50] Robin Wakefield. 2013. The influence of user affect in online information disclosure. *The Journal of Strategic Information Systems* 22, 2 (2013), 157–174.
[51] Maximiliane Windl, Niels Henze, Albrecht Schmidt, and Sebastian S Feger. 2022. Automating Contextual Privacy Policies: Design and Evaluation of a Production Tool for Digital Consumer Privacy Awareness. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
[52] Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Lauren Schmidt, Laura Brandimarte, and Alessandro Acquisti. 2014. Would a privacy fundamentalist sell their DNA for $1000... if nothing bad happened as a result? The Westin categories, behavioral intentions, and consequences. In *Symposium on Usable Privacy and Security (SOUPS)*, Vol. 5. 1.
[53] Yuxi Wu, W Keith Edwards, and Sauvik Das. 2022. "A Reasonable Thing to Ask For": Towards a Unified Voice in Privacy Collective Action. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.
[54] Shikun Zhang, Yuanyuan Feng, Lujo Bauer, Lorrie Faith Cranor, Anupam Das, and Norman Sadeh. 2021. "Did you know this camera tracks your mood?": Understanding Privacy Expectations and Preferences in the Age of Video Analytics. *Proceedings on Privacy Enhancing Technologies* 2021, 2 (2021).

A EVALUATION DETAILS

A.1 RQ1 Consistency analysis in section 4.1

Both ICC and Pearson correlations were calculated using the Statistical Package for the Social Sciences (SPSS). ICC estimates and their 95% confidence intervals were based on a mean-rating (k = 3), absolute-agreement, 2-way random-effects model. For assessing concern category consistency, we defined a triple set S_{cp} for a concern category C from participant P's response to the redundant scenario. In each of our three consistency test surveys, if P's response is labeled with C, the corresponding element in S_{cp} is set to 1, otherwise 0. Therefore, all-one (or all-zero) sets mean the specific participant' privacy concern (or non-concern) towards the same scenario stayed consistent across three surveys.

A.2 RQ2 Concern score definition in section 4.2

The concern score C_l to gauge the extent to which labels contribute to general levels of privacy concern is defined as follows

$$C_l = \sum_{i=1}^{N} \frac{s_l(i)}{S_l}$$

where $s_l(i)$ is the number of Concerns (C) user (i) expressed in stories with label (l), N is the total number of participants, and S_l is the total distinct number of stories marked with label (l).

B SURVEY AGGREGATED RESULTS

C PRIVACY STORIES

Figures 10 - 21 outline the privacy storyboards of the 12 real-world data applications we used in our surveys. We only offer a brief text summary for each data action in each story.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009

Table 13: Aggregated survey result. Average scores are the average ratings of all data actions in each scenario, and participants' concerns are the total number of concerns from 14 concern categories labeled from all free text responses. ContextLabel are the number of ContextLabel we annotated for all data actions in each scenario, using the codebook with label definitions in Table 1.

Scenario Index	1	2	3	4	5	6	7	8	9	10	11	12	13
Average Score	3.38	2.22	2.61	2.27	2.49	3.18	3.57	2.17	3.78	2.93	2.19	2.38	2.32
Participants' Concerns	75	164	92	129	125	49	37	117	56	87	89	71	132
ContextLabel	23	19	22	24	22	12	11	16	15	21	20	19	19

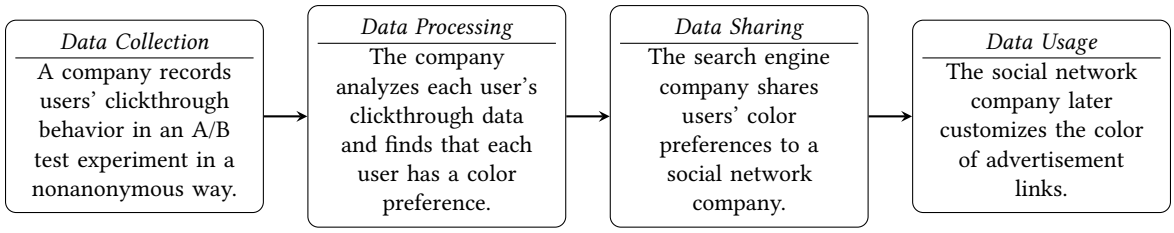

Figure 10: A privacy storyboard of "Search engine clickthrough data". A company records users' clickthrough behavior in an A/B test experiment nonanonymously and uses the data for advertising and search personalization.

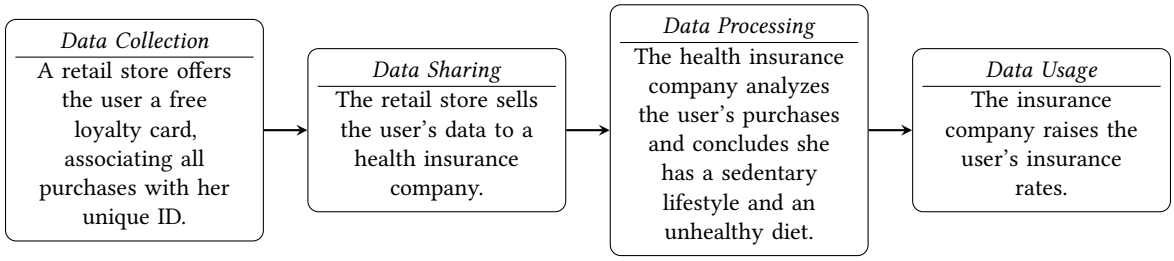

Figure 11: A privacy storyboard of "Loyalty card in a retail store". A retail store collects users' data through a loyalty card and uses the data for insurance and coupon personalization.

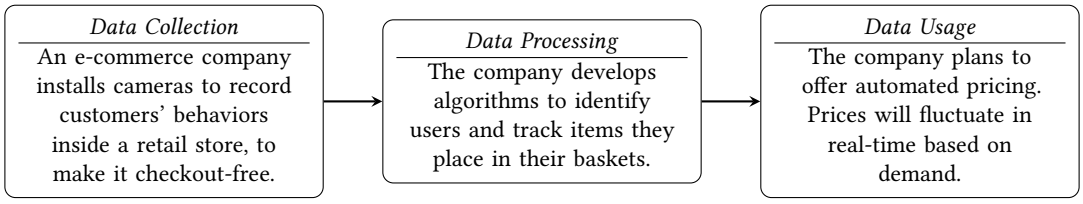

Figure 12: A privacy storyboard of "Checkout-free retail store". An e-commerce company opens a checkout-free retail store by installing various sensors inside a physical store.

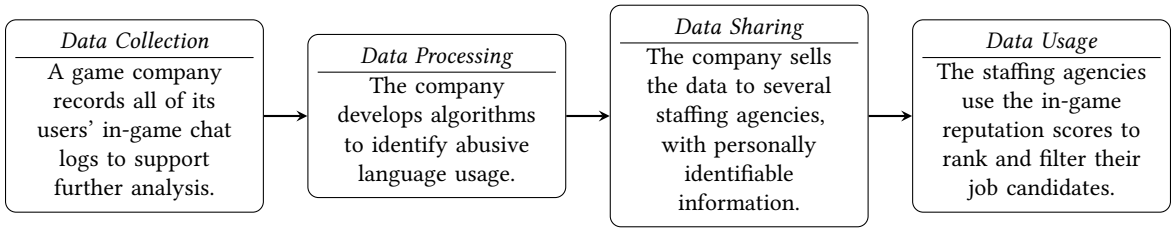

Figure 13: A privacy storyboard of "Game chat log". An online game company uses its chat logs to identify potential problems in the workplace.

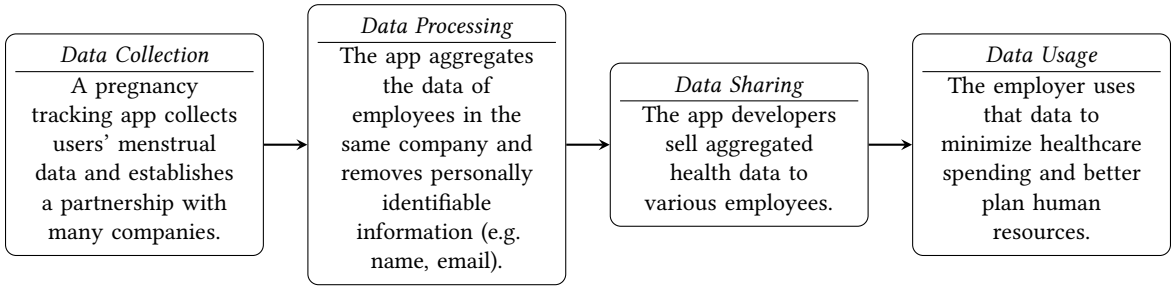

Figure 14: A privacy storyboard of "Pregnancy intimate data". A pregnancy app shares users' intimate body data with their employers.

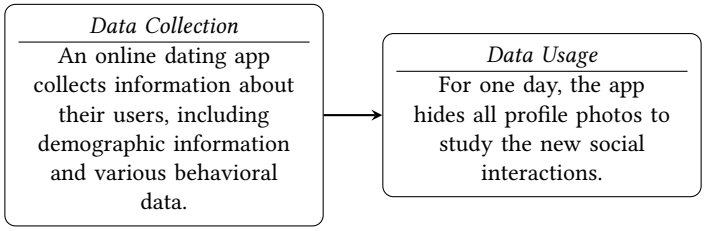

Figure 15: A privacy storyboard of "Data science experiments in a dating app". An online dating app conducts several experiments to understand the nature of romance.

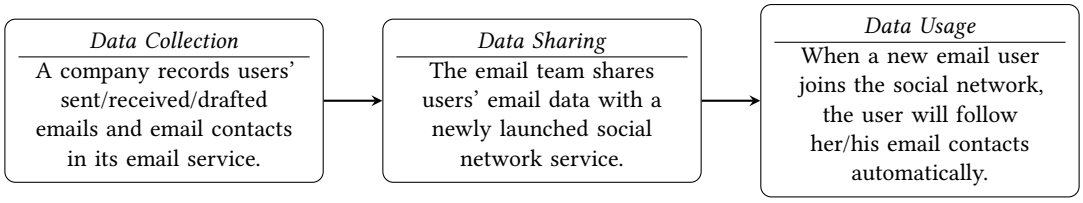

Figure 16: A privacy storyboard of "Email contacts for social network bootstrapping". A technology company appropriates users' email data to bootstrap a new social network service.

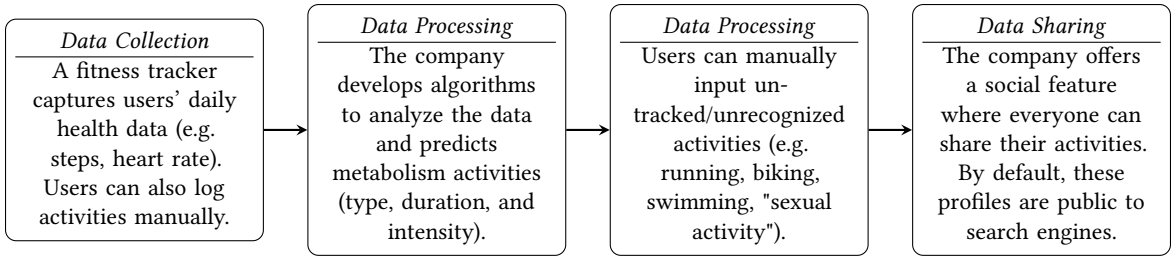

Figure 17: A privacy storyboard of "Fitness tracking". A wearable technology company collects users' intimate behavior data and makes them public by default.

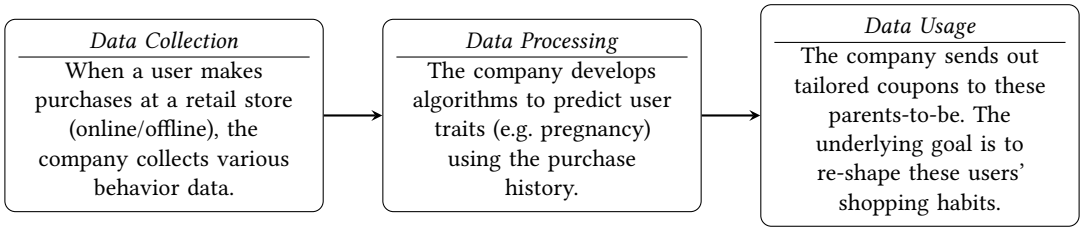

Figure 18: A privacy storyboard of "Retail store pregnancy". A retail store predicts users' pregnancy status by analyzing their purchase history.

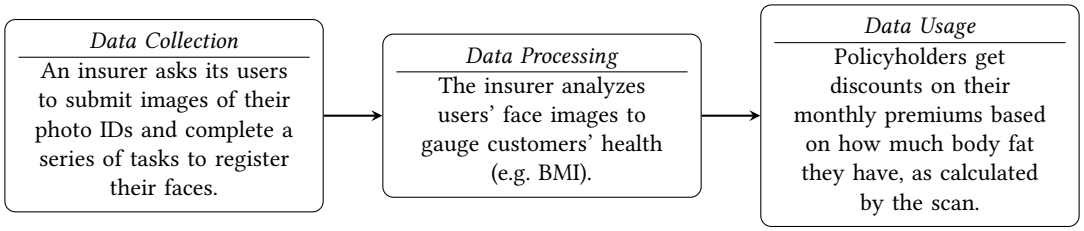

Figure 19: A privacy storyboard of "Insurer employs AI". An insurance company uses facial-recognition technology to identify untrustworthy and unprofitable customers.

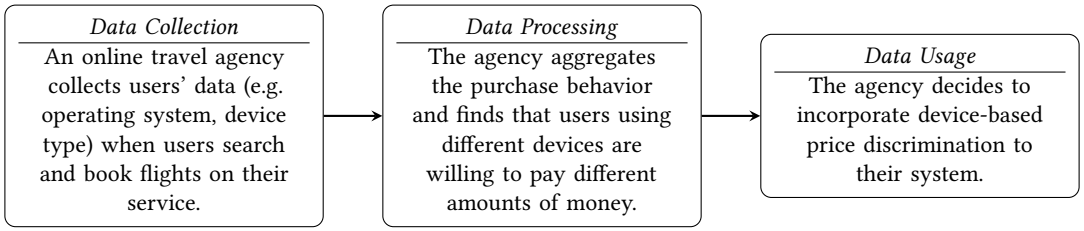

Figure 20: A privacy storyboard of "Dynamic pricing". Technology companies collect users' behavior data to adjust the service price dynamically.

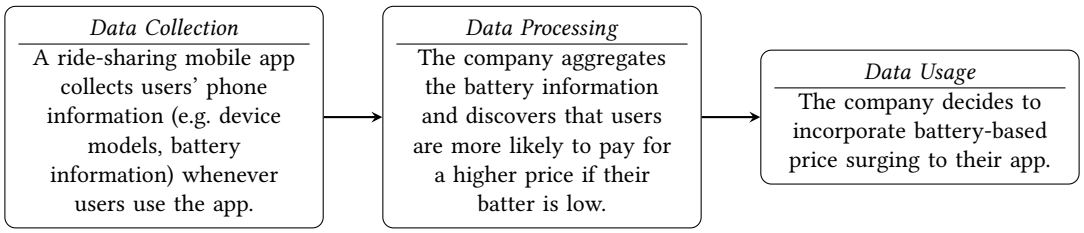

Figure 21: A privacy storyboard of "Dynamic pricing". Technology companies collect users' behavior data to adjust the service price dynamically.