Towards Usable Privacy Management for IoT TAPs: Deriving Privacy Clusters and Preference Profiles

Piero Romare[1*], Farzaneh Karegar[2] and Simone Fischer-Hübner[1,3]

[1*]Department of Computer Science and Engineering, Chalmers University of Technology, Rännvägen 6, Göteborg, 41261, Sweden.
[2]Department of Information System, Karlstad University, Universitetsgatan 2, Karlstad, 65188, Sweden.
[3]Department of Mathematics and Computer Science, Karlstad University, Universitetsgatan 2, Karlstad, 65188, Sweden.

*Corresponding author(s). E-mail(s): pieror@chalmers.se;
Contributing authors: farzaneh.karegar@kau.se; simonefi@chalmers.se;

Abstract

IoT Trigger-Action Platforms (TAPs) typically offer coarse-grained permission controls. Even when fine-grained controls are available, users are likely overwhelmed by the complexity of setting privacy preferences.

This paper contributes to usable privacy management for TAPs by deriving privacy clusters and profiles for different types of users that can be semi-automatically assigned or suggested to them.

We developed and validated a questionnaire, based on users' privacy concerns regarding confidentiality and control and their requirements towards transparency in TAPs. In an online study (N=301), where participants were informed about potential privacy risks, we clustered users by their privacy concerns and requirements into Basic, Medium and High Privacy clusters. These clusters were then characterized by the users' data sharing preferences, based on a factorial vignette approach, considering the data categories, the data recipient types, and the purpose of data sharing.

Our findings outline three distinct privacy profiles, providing a foundation for more usable privacy controls in TAPs.

Keywords: Privacy, Smart Environments, Internet of Things, Survey, Quantitative Method, End-User

1 Introduction

The fast proliferation of connected devices and services in the Internet of Things (IoT) has fundamentally transformed how we interact with and experience our environment. IoT-enabled systems offer a wide range of automation benefits, including improved energy efficiency, enhanced safety, and increased convenience for users (Khodabakhsh and Yayilgan, 2020). However, these systems' ubiquitous data collection and sharing capabilities have also raised significant privacy concerns among users (Zheng et al., 2018; Emami-Naeini et al., 2017). How users can exercise their right to informational privacy, including control over the collection and use of personal data, should match their expectations (Rao and Pfeffer, 2020).

This study focuses on Trigger-Action Platforms (TAPs), which enable users to create automated rules that control the behavior of connected devices (Ur et al., 2016), allowing for customisation of their everyday digital systems including smart homes, smart cities, and wearables. TAPs host applications that are based on the "if-this-then-that" formula, giving the end-users the role of high-level programmers by connecting different devices and/or services. These platforms, such as IFTTT, Zapier, Power Automate, Make and many others, allow users, even without extensive coding expertise, to create and submit their own applications from user-friendly web interfaces. IoT Trigger-Action applications serve as bridges ("webhooks") connecting smart devices and online services, improving their interoperability. Each app consists of an event-driven program with a code snippet that includes at least a trigger and an action. For example, a user might create a rule that turns on the smart lights when the front door is unlocked.

TAP applications, such as in IFTTT, can collect users' data from a device or service and transmit the data to third party entities involved in the application whenever certain events trigger actions to be performed by these third-party entities (Inc., 2025). Significant privacy issues may emerge when users employ this automation and customisation capability with personal or sensitive data collected from IoT devices or inferred by the TAP (Aghvamipanah et al., 2024), and then forwarded to other third party entities (Wang et al., 2025). Privacy issues also specifically arise, as "overprivilege" is a significant shortcoming of permission models in Trigger-Action IoT Platforms (Fernandes et al., 2016; Xu et al., 2019; Ahmadpanah et al., 2023). The permission systems used by many TAPs, including IFTTT, to control these personal data flows are only based on coarse-grained permissions that in most cases provide more generous access permissions and controls than needed and desired by users (Balliu et al., 2019). Nevertheless, even when fine-grained controls are in place, users, for whom privacy and security are usually only secondary goals (Fischer-Hübner and Karegar, 2024), are likely to be overwhelmed with the complexity of setting fine-grained privacy controls that will correspond with their preferences. An abundance of options actually requires more user effort to choose and can leave users feeling unsatisfied with their decisions (Schwartz, 2015). As concluded in Balliu et al. (2019), fine-grained permission and control systems are needed for IoT TAPs, which provide users with control over how their personal data are used by third-party sites, and additionally these permission and control systems should also be usable. To this end, approaches based on bundling access permissions and controls, and automatically deriving suitable profiles (or “bundles”) of permissions and controls that align with users’ preferences may be beneficial (Balliu et al., 2019). Aligning such permission and control systems with users’ requirements and data sharing preferences is key to ensuring that these systems are transparent, trustworthy and support the users’ expectations regarding control over their personal data. In other words, we need user-tailored privacy permission and control systems (Knijnenburg, 2017).

Privacy preferences refer to users’ wishes configurations that control how their personal data is shared, accessed, and utilised by others, reflecting their desired level of control over how their data is handled by various entities which is influenced by a combination of factors, including cognitive biases, contextual information, and the trade-offs between privacy and the benefits of sharing data (Acquisti et al., 2015). In the IoT context, the privacy preferences are highly variable and context-dependent, as participants may feel more or less comfortable sharing their data based on specific situations or applications (Emami-Naeini et al., 2017). To create a user-tailored privacy permission system, we need simplified privacy management based on personalized settings that align with users’ privacy preferences, supporting decision-making more intuitive and less overwhelming. This approach can involve clustering users based on either their privacy behaviors—observable actions they take to protect their data (Lin et al., 2014; Liu et al., 2014)—or their privacy attitudes, which reflect their internal thoughts and mental states (Kumaraguru and Cranor, 2005; Dupree et al., 2016).

Users often adjust their privacy preferences based on the perceived trade-offs between risks and rewards in specific situations, reflecting an ongoing negotiation between privacy concerns, individuals’ perceptions about the implications of sharing information online (Dinev and Hart, 2006)–and the benefits of data sharing, as highlighted by Acquisti et al. (2015) in their discussion of privacy decision-making dynamics and privacy calculus. However, research shows that users may have misconceptions about security and privacy risks in the context of IoT TAPs. For example, in a previous user study, real IFTTT users expressed that while secrecy and integrity of potential harms were important to them, they believed their applets were safe and did not modify their level of caution even when presented with explanations of potential violations (Cobb et al., 2020). Similarly, an online survey study found that users often struggle to understand the security and privacy implications of IoT devices, with many indicating misconceptions about the risks associated with real IFTTT applets, but with guided risk assessments considering the time of day, location, and the presence of others, users were better able to identify potential risks including leakage of sensitive data and unauthorized or unintended access (Saeidi et al., 2022). Users may also have difficulties with understanding potential privacy risks and how they can set adequate permissions that align with their preferences (Madejski et al., 2012), often due to a lack of risk awareness and knowledge about how their data might be used, which further complicates their ability to accurately assess risks (Emami-Naeini, 2020).

Therefore, solely relying on the collected privacy preferences of users who were not well informed about potential risks may not be sufficient for deriving effective controls for protecting users according to their expectations. Since privacy preferences are often shaped by users’ perceptions of risk and their privacy concerns (Lee and Kobsa, 2017), the variability in users’ risk perception and technological understanding complicates effective decision-making in IoT TAPs. Without sufficient awareness of potential risks, privacy concerns and in turn preferences may not adequately align with the controls needed to safeguard users' data (Lee and Kobsa, 2017).

Consequently, in this paper, our main goal is to derive clusters of attitudinal privacy concerns and requirements from users who have been exposed to different IoT TAPs scenarios and informed about potential privacy risks. These clusters will then be characterized by users' data sharing preferences to form privacy preference profiles. These profiles can serve as an important step towards deriving suitable bundles of permission settings that better align with users' privacy expectations and needs.

To derive clusters of privacy concerns and requirements, we conducted an online survey using a custom-designed questionnaire. Participants responded to items reflecting three a priori theorized dimensions of IoT-TAP privacy concerns and requirements in the context of four real IoT TAP application scenarios: confidentiality, transparency, and control [1]. For each scenario, participants were provided with information about the potential privacy risks associated with that specific application. Moreover, we used a factorial vignette setting to further describe and characterize these privacy clusters in terms of users' data sharing preferences, focusing on key privacy factors in the context of IoT, as identified by (Emami-Naeini et al., 2021), to form privacy profiles for IoT TAPs. These key factors include: (1) data category (personal data) to be shared: personally identifiable information (PII) (name, surname, address, IP address), location data, message and email data, image and video data; (2) purpose of data sharing: main app functionality, personalised app functionality, targeted advertisement; and (3) data recipient type: with service providers (e.g., parties that are involved in the IoT app as trigger or action providers), government and legal authorities, and (other) third parties. Further details on the factors selection are in Section 3.2.2.

Research Questions and Contributions

Our survey study was designed to answer the following research questions aimed at achieving our overall research goal for current and potential IoT TAP users with different backgrounds:

RQ1: *What* types of privacy clusters can be identified, and *how*, to serve as a foundation for a usable privacy permission and control management system that reflects concerns and requirements of users who are informed about potential privacy risks in IoT TAP applications?

RQ2: How can these privacy clusters be further characterized based on users' data sharing preferences in IoT TAP applications to form privacy preference profiles?

Our study contributes to the field with the following key findings that previously have not been investigated in the context of IoT TAPs:

- **Validated Questionnaire:** We developed and validated a novel questionnaire as a reliable data-gathering procedure to cluster participants, who were informed about potential privacy risks, for four selected IoT TAP application scenarios.

[1] Measurement validation later indicated that confidentiality and control loaded on a single factor, so subsequent analyses used two dimensions: confidentiality/control, and transparency.

- **Privacy Clusters for IoT TAP Applications and Users:** From the questionnaire responses, we derived privacy concerns and requirements clusters for participants who were informed about privacy risks in IoT TAP scenarios.
- **Privacy Preference Profiles:** We showed how the features from our data sharing factorial vignette study procedure matched and characterized the privacy clusters derived and formed privacy preference profiles.
- **Directions for Usable Privacy Management:** We provided insights into and discussed how our results can serve as the first step towards developing usable privacy management systems for IoT TAPs, based on bundles of privacy settings that can be proposed or semi-automatically assigned to users exposed to potential privacy risks, using a simple, validated 6-question questionnaire.

To the best of our knowledge, this is the first study to investigate privacy questionnaires, clusters and profiles within the specific context of IoT TAPs. Our research contributions differ considerably from earlier work that focus exclusively on IoT in general, rather than IoT TAPs, because (a) IoT TAPs introduce additional privacy risks through overprivileged permission system, automated actions, and data transfers to third parties, and (b) the privacy factors influencing users' concerns and preferences for IoT TAP use extend beyond those typically identified for IoT more broadly (Romare, 2024; Romare et al., 2023).

Organization

In Section 2, we present background about IoT TAP, privacy concerns and preferences and related work on privacy scales and privacy clustering techniques used in other contexts than IoT TAP. In Section 3, we provide the design of our validated questionnaire with the scenarios and risks proposed, address research ethics, and present the recruitment, sample and the statistical methodology that we employed. In Section 4, the details about the validity and reliability of the questionnaire are presented, and our research questions are answered. In Section 5, we discuss the relevance of our results as a step towards implementing usable privacy permission management for IoT TAPs. We also discuss our results in comparison with related work and discuss limitations as well as future research directions. Finally, we conclude this paper in Section 6 by summarizing the main results of our study and briefly discussing future research directions.

2 Related Work

In this section, we briefly summarise and discuss work related to our research contribution. This includes literature on how to capture privacy concerns and requirements through privacy scales, with a later focus on the context of IoT TAPs, as well as related work on categorising users and deriving privacy clusters and profiles for capturing users' privacy and preferences.

2.1 Capturing Users' Privacy Concerns and Requirements in IoT

Privacy concerns, as defined by Smith et al. (1996), refer to the worries and uneasiness of people regarding the loss of privacy and how well they are protected against unauthorised access and misuse of personal information. People were found to be more concerned about their privacy when the information process is not transparent (Phelps et al., 2000).

To explore privacy concerns across various contexts, privacy measurement scales and models have been developed to capture the often latent variable that expresses an individual's privacy attitudes. The Concern for Information Privacy (CFIP) model with 15 items explores organizational privacy practices as well as people's concerns towards four dimensions: collection of data, unauthorised secondary use of data, improper access to data, and errors in data (Smith et al., 1996). As the first validated instrument regarding privacy concerns, the CFIP served as a foundation for the Internet Users' Information Privacy Concerns (IUIPC) (Malhotra et al., 2004). The IUIPC refines the CFIP, and it is a scale for analysing privacy concerns for internet users that consists of 10 items, or in its shorter version of 8 items (Groß, 2021), among three dimensions such as data collection, control and awareness. The Internet General Privacy Concern (IGPC) scale was developed to understand the e-commerce customers' caution and mistrust over their personal data under the dimensions of information transfer and use of data by (Castañeda et al., 2007). Earp et al. (2005) developed another instrument to verify the alignment between the privacy policies' practices and the users' expectations regarding the employed data protections in online websites considering the personalisation, awareness, data transfer, collection, storage and data access. Regarding online social networks, the multidimensional privacy orientation scale includes four dimensions named privacy as a right, concern about own informational privacy, and concerns about others' privacy. Based on these dimensions, Baruh and Cemalcılar (2014) identified three user segments: privacy advocates, individualists, and indifferents. The Value of Other People's Privacy (VOPP) handled how much people protect others' personal information with a validated psychometric scale (Hasan et al., 2023). The Privacy Attitudes Questionnaire (PAQ) is another instrument to investigate a multidimensional construct considering exposure, willingness to be monitored, interest in privacy and privacy control or trust with 36 items (Chignell et al., 2003). The Mobile Users' Information Privacy Concerns (MUIPC) was developed by Xu et al. (2012) for measuring privacy concerns in mobile applications, including perceived surveillance, intrusion and secondary use of personal information. The MUIPC was later also extended to the IoT context (Foltz and Foltz, 2020). This scale was applied to the IoT to assess comfort levels across health, safety, user experience, and personalisation benefits. In the contexts of user experience and personalisation, comfort levels decreased when data sharing was allowed (Chawdhry et al., 2022), while providing transparent information can often ease concerns (Magrizos et al., 2025). In the context of IoT TAPs, Romare et al. (2023) used focus groups as a qualitative research method for eliciting privacy factors that impact users' concerns and preferences for using IoT TAPs, and identified transparency, control, risks, trust and confidentiality as relevant factors. Saeidi et al. (2022) were the first to use a quantitative approach for measuring users' concern scores for using IFTTT applets and explored if and how those concerns were impacted by different contextual factors such as location, time of the day and other people's access to the trigger or action service. Their results showed that concern scores were low when their participants were only exposed to the app descriptions, which indicated a limited awareness of potential risks. Considering the smart home environment, in (Apthorpe et al., 2018), the acceptability of thousands of information flows that align with indications of IFTTT user studies has been measured. Indeed, individuals' priorities are user consent and transparency in data collection practices since user awareness and control over data sharing are key factors influencing privacy concerns. While influential in the design of our questionnaire in terms of selection of items and dimensions we investigated, none of these earlier studies, in contrast to our work, focused on capturing privacy clusters and profiles in the IoT TAP context. Further details about the design, selection and modification of existing and novel items included in our questionnaire are discussed in Section 3.2.1.

A questionnaire can be used as a measurement tool that can collect data to perform the clustering in this regard when properly evaluated (Biselli et al., 2022). Thus, we validated a privacy concerns and requirements questionnaire developed by adapting existing items to the context of IoT TAPs and proposing a tool to measure privacy attitudes specific to these technologies and show how this questionnaire can be used to cluster study participants who have been informed about potential privacy risks.

2.2 Privacy Categorization of Users, Privacy Clusters and Profiles

Previously, Inverardi et al. (2023) presented a systematic literature review on privacy categorisation, including work related to our study, and showing how terminologies and methods evolved over time, concluding with discussing the potential of these approaches to support users in managing their privacy with the help of recommendation systems that can be developed. The term segmentation can be essentially expressed as *a model-driven procedure of partitioning a dataset or extracting associated features.* The term clustering is usually used to refer to the *data-driven mathematical process of grouping similar data points.* The term profiling can generally be specified as a *hybrid approach that combines multiple data sources and can involve statistical analysis to identify characteristics of individuals or groups.* Personas, on the other hand, refer to *attributing new parameters to existing segments or clusters.*

Early work on privacy categorisations of users dates back to 1990/1991 when Alan Westin proposed his first privacy segmentation from responses to four questions. His approach grouped individuals into three categories: privacy fundamentalists, pragmatists, and the unconcerned (Westin et al., 2003). His methodology provided a framework for segmenting participants based on their privacy concerns and attitudes and his segmentation has been utilised by academics in a wide range of domains for analysing and categorising users regarding their privacy attitudes and concerns, but was also discussed for its limitations (Kumaraguru and Cranor, 2005). Individuals' privacy strategies from Westin's key aspects of privacy, applied in both analogue and digital contexts, are the right to be alone, intimacy, anonymity, and limited information disclosures (Westin, 1968). These were refined to address more specific IoT privacy necessities by Ziegeldorf et al. (2014), including awareness of privacy risks posed by connected devices and services, control over personal data collection and processing, and awareness and control of third-party data sharing.

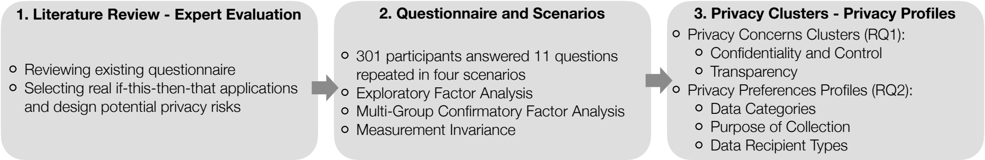
Inverardi et al. (2023) highlight several key changes as a result of their review on privacy categorization and one significant shift mentioned is the evolution from a focus on data minimization and binary privacy choices to a more nuanced understanding of privacy preferences modelled using qualitative, quantitative and hybrid approaches, where mostly three categories of users were captured. Early research often presented privacy decisions as static and simplistic, but the rise of complex digital ecosystems, such as IoT and social media, required a more dynamic approach, where users' privacy preferences vary based on the specific context or scenario. Building on this evolution, our study similarly acknowledges the importance of context dependency in privacy preferences. Accordingly, in our questionnaire, we incorporated scenario-specific considerations to capture the variability in users' privacy concerns and requirements in different IoT TAPs scenarios, as described in Section 3.1.

Clustering algorithms, such as hierarchical clustering and K-means have been used to derive privacy profiles. In Brandão et al. (2022), three privacy profiles were extracted from the hierarchical clustering algorithm, whereas six more granular and nuanced profiles emerged from the k-means approach for Android smartphones, considering detailed app category–permission combinations in a vector meticulously representing each individual user. Privacy preferences can also be modelled and used as input data to train machine learning models for clustering and predicting users' privacy settings in the form of profiles, thereby helping to address the cold start problem. For example, Bahirat et al. (2018) tested attitude-based, fit-based (achieving 82% accuracy with three profiles), and agglomerative clustering solutions. Their analysis demonstrates how participants were more comfortable when information where not continuously shared, and instead perceived more risk when the recipient was unknown using the dataset provided by Lee and Kobsa (2016). Originally, that dataset was analyzed for clustering IoT users by exploring the relationships between where data are collected, what type of data are gathered, who receives them, for what purposes, and how frequently collections occur in IoT scenarios. That analysis resulted in four distinct clusters of users based on their privacy preferences related to notifications, permissions, comfort, risk, and appropriateness (Lee and Kobsa, 2016). Another extensive study in IoT, involving over a thousand participants, used factors such as who, what, purpose, storage location, and action to create data sharing scenarios for users of household IoT devices. With the help of machine learning algorithms, ”smart” profiles were derived considering the adoption decision and the contextual factors (He et al., 2019). In the smartphone context, privacy profiles were used to implement a privacy assistant, which was, for instance, developed by Liu et al. (2016) to help Android users manage their privacy preferences with a behavioral field study. The privacy assistant could recommend privacy settings thanks to a Support Vector Machine classifier that reached 79% of acceptance rate. In another study, four privacy profiles were created considering the number of applications installed and users' permissions decisions from their phones and these four clusters were then combined with self-reported privacy attitudes and intentions of use to derive profiles (Alsoubai et al., 2022).

To the best of our knowledge, we are the first to integrate, in the context of IoT TAPs, attitudinal privacy concerns and requirements clusters collected with a questionnaire and modelled into profiles combining the data sharing preferences. By bridging self-reported privacy concerns with context-specific choices, our work can lay the foundation for future studies using behavioral data and aims to contribute to the development of usable privacy management and privacy assistants for IoT TAPs. As we will further elaborate in Section 5.1, our derived clusters not only differ domain and scope-wise but also content-wise from other clusters derived in earlier related work.

3 Method

In this section, we present our approach to address the research questions presented in Section 1. First, we conducted a literature review and an expert evaluation to set the scene and define the TAP scenarios for our questionnaire. Then, we conducted our questionnaire to capture users' attitudinal privacy concerns and requirements across three key dimensions: transparency, control, and confidentiality. To ensure the effectiveness of our questionnaire in extracting meaningful privacy clusters, we demonstrate the validity and reliability of this measurement tool. Lastly, with a factorial vignette study approach, we characterized the derived clusters based on users' data sharing preferences, resulting in distinct privacy profiles. The steps we followed are depicted in Figure 1. In addition, we discuss the ethical considerations, the participant recruitment process, and our data analysis methods in this section.



1. Literature Review - Expert Evaluation

- Reviewing existing questionnaire
- Selecting real if-this-then-that applications and design potential privacy risks

2. Questionnaire and Scenarios

- 301 participants answered 11 questions repeated in four scenarios
- Exploratory Factor Analysis
- Multi-Group Confirmatory Factor Analysis
- Measurement Invariance

3. Privacy Clusters - Privacy Profiles

- Privacy Concerns Clusters (RQ1):
  - Confidentiality and Control
  - Transparency
- Privacy Preferences Profiles (RQ2):
  - Data Categories
  - Purpose of Collection
  - Data Recipient Types

Fig. 1 The Main Steps used in this study

3.1 Expert evaluation and IoT TAP Scenarios

To capture users' attitudinal concerns and requirements related to IoT TAPs, we need a clearly defined context that frames the scenario in which users can respond to privacy-related questions. Consequently, we defined four scenarios that each presents the application's description and the privacy risk related to additional information that an adversary could infer about the user.

Information about privacy risks was included, since, as introduced in Section 1, a gap exists in participants' awareness of the risks associated with additional information that could be inferred (Alqhatani and Lipford, 2023) or, in other words, if there's a leak of data, which could also be due to a possible IFTTT application's violation as exemplified in (Surbatovich et al., 2017). The purpose of introducing the scenarios is to give examples of existing IoT TAP applications with related risks and help the participants understand how the applications can be used in their everyday lives.

Table 1 Scenario Figure Reference (#) with description and privacy risk.

#	**App Description**	**Privacy Risk**
2	If a new sleep is logged by your smartwatch, then add an event in your cloud calendar with sleep information.	If an adversary has intercepted the communication, there is a risk that sleep routine details may be leaked and used to infer a person's stress levels, fatigue, or suggest sleeping pills.
3	If your location is outside your home, then lock the door.	If an adversary collects your location information, they could predict when you won't be at home.
4	If a new photo is in your smartphone's camera roll, then upload it to your cloud storage.	If an adversary gains unauthorized access to the cloud storage, they could potentially extract sensitive information from the uploaded photos. This may include a collection of places you visited, their time, and people with you, leading to a privacy breach and the unintentional exposure of personal details.
5	If you like a video, then post it in your online social network account.	Posting content to an online social network may allow it to target you with advertisements based on your personal preferences from the media platform.

As a first step for selecting suitable applications for the scenarios, we compared the dataset available in (Kalantari et al., 2022) and the most used IFTTT applications based on their categories [2]: 1) mobile, devices and accessories, 2) news and information, 3) social media, 4) notifications 5) business tool 6) photo and video. We manually filtered and selected applications that included personal data flows that allowed us to infer sensitive information, and then categorized them into four categories: mobile devices and accessories, social media, business tools, and photos and videos. This first step resulted in 16 realistic scenarios that use common IoT devices or services.

We then conducted semi-structured interviews with six experts to form and filter out our scenarios and to gather expert opinions on potential privacy risks and their severity for each scenario. The results of our expert interviews finally led to the selection of four scenarios to be included in our study. Our expert participants were PhD students and post-doctoral researchers with backgrounds in Information Security and Privacy at Chalmers University of Technology who volunteered to take part in the interviews after our call for expertise requirements and provided their informed consent before the interview started. The interviews took 25 minutes on average and included the following steps: 1) introducing the context of information flow in IoT TAPs, 2) explaining the goal of the current study, and 3) exposing participants to 16 scenarios and asking them to identify potential risks.

The privacy risks for each scenario can have personal implications through privacy violations that may cause embarrassment or leak behavioral data (Surbatovich et al., 2017). The results of these interviews led to the selection of a suitable subset of if-this-then-that IoT TAP scenarios with certain characteristics, such as personal data

[2] https://ifttt.com/content_map (visited Nov 2023)

flow, to be used in our study (see Table 1). The scenarios are related to real IFTTT applications that can be accessed at their website [3] using the id provided under each Figure representing the application. To help our participants better grasp the scenarios and their context, each scenario was accompanied by a description and related figure (see Figure 2, Figure 3, Figure 4, Figure 5).

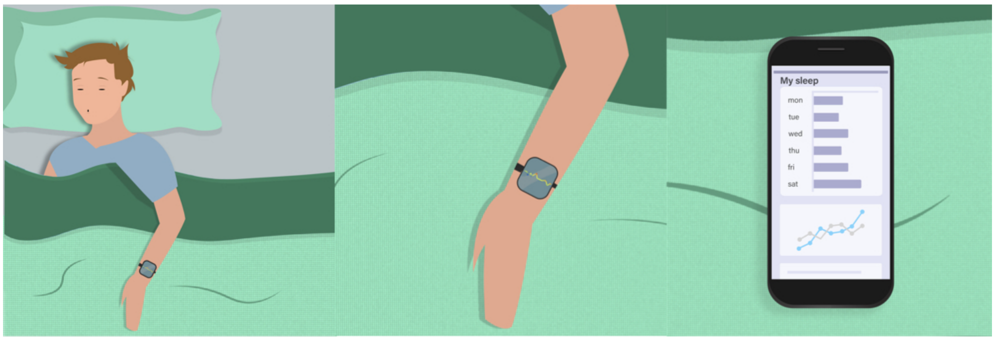


Fig. 2 Scenario #1: if new sleep log, then upload sleep information on cloud calendar - ID: wsTcJyNt

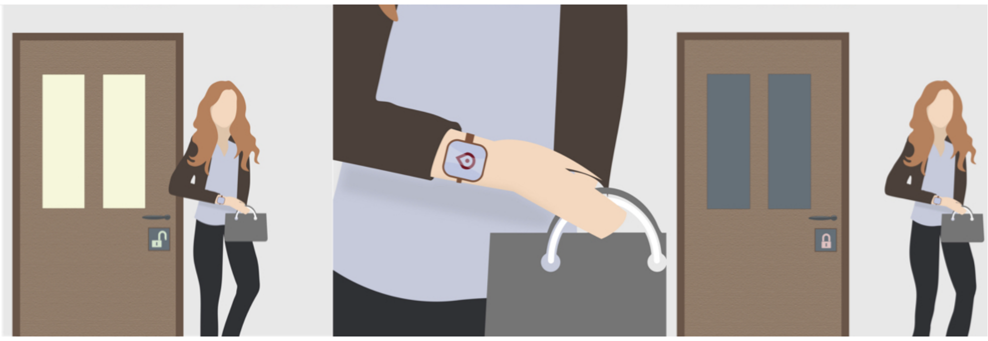


Fig. 3 Scenario #2: if outside home, then lock the door - ID: NBwbDaze

While potential privacy risks that were identified via our expert interviews may be obvious for privacy and security experts, earlier research by leading privacy scholars, including Acquisti et al. (2015) and Solove (2013), shows that lay users often misjudge and underestimate privacy risks and potential harms, and fail to connect abstract risks to concrete consequences. Also, as pointed out above, Saeidi et al. (2022) demonstrate that many people are not aware or have misconceptions about privacy risks in the context of IoT TAPs. This motivated us to inform test participants about potential privacy risks and consequences related to the IoT TAP that were identified by our experts.

[3] https://ifttt.com/applets/(id)

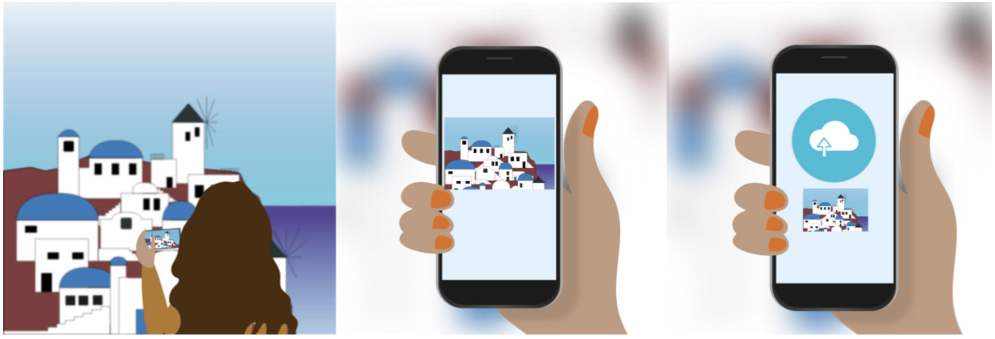


Fig. 4 Scenario #3: if new photo, then upload on cloud storage - ID: th9yp6nk

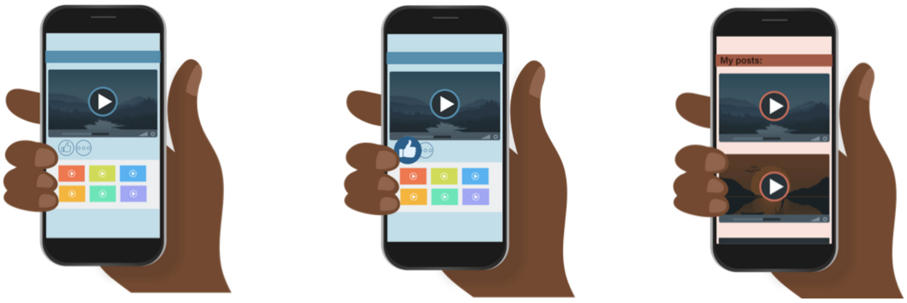


Fig. 5 Scenario #4: if you like a new video, then post it on online social media - ID: k6LpJhGg

3.2 Questionnaire Design

In the recruitment invitation message, we briefly introduced the context of the questionnaire to the participants. We provided the Wikipedia definition of IoT and introduced IoT Trigger Action platforms with an example: ”If enter a room, then turn on the lights”. In addition to a self-explained image which described the data flow between the two entities, we added text details about the trigger, the action and the interconnection between them. If the participant decided to continue and take part in the study, the consent form and the information were provided to the participant. Upon giving their consent, they answered a few demographic questions and questions related to their usage of IoT devices. We asked about the region of residence (EEA / US), the age group (18-25 / 26-35 / 36-45 / 46-55 / 55+), the gender (Woman, Man, Non-binary, Prefer not to say), how many IoT devices they owned (0 / 1-2 / 3-4 / 5 or more), if they had data breach experiences (Yes / No), whether they used or heard of any IoT TAPs (Yes / No) and lastly their highest level of education (Less than High School / High School / University Degree). After these questions, participants were exposed again to the introduction to IoT, IoT TAP, and IFTTT with an example.

The core part of our questionnaire then followed, which included two main parts for deriving privacy concerns and requirements, and data sharing preferences.

In the first part, we presented four scenarios in the context of IoT TAP to each participant and after each scenario, they answered 11 question using a 5-point Likert scale (from strongly disagree to strongly agree) for capturing attitudinal privacy concerns and requirements related to three dimensions of confidentiality, control, and transparency (presented in Table 2 and in Appendix B). In the second part, we captured users' data sharing preferences in the context of IoT TAPs. Participants had to indicate their willingness (yes or no) to adopt/accept 36 specific IoT TAP scenarios which varied in three aspects: 1) the data category to be shared, the purpose of data sharing, and the data recipient type (see Table 3 and Appendix C).

3.2.1 Privacy Concerns and Requirements

We designed a questionnaire that captures attitudinal privacy concerns and requirements in three dimensions related to confidentiality of personal data, user control, and transparency for our defined IoT TAP scenarios. A systematic literature review analyzed the developments of 16 privacy concern scales and concluded that there is a lack in deriving the constructs from existing theoretical frameworks to get common definitions (Bartol et al., 2021). Our three dimensions correspond to three dimensions of the six-axis privacy protection goals framework's taxonomy presented by Hansen et al. (2015). These three data protection goals play an important role in the design of a transparent permission control system, which allows users to protect and control the disclosure of and access to their personal data. Anonymity, secrecy and confidentiality were shown as significant contributors to the levels of control over users' personal data (Tamara Dinev and Hart, 2013). When users can control and manage their data, they have fewer privacy concerns (Xu et al., 2011; Tamara Dinev and Hart, 2013). Transparency is a key aspect for building trust between the users and those who are processing their data (Xu et al., 2024). Other important data protection goals of unlinkability/anonymity or data minimisation of the taxonomy by Hansen et al. (2015) can be more effectively protected by a privacy-preserving TAP re-design as suggested in (Ahmadpanah et al., 2023) rather than by permission controls, and were therefore not directly captured by our questionnaire.

Basing our questionnaire on these three dimensions was also motivated by earlier work conducted by Romare et al. (2023) that identified transparency, control and confidentiality as important privacy factors playing a role in users' concerns and requirements for using IoT TAPs and its automation function. In their qualitative user study, participants particularly demanded control over the final action, expected to be able to restrict or disallow full automation, and required transparency of automation settings, rules, and data recipients, as well as the confidentiality of communicated and stored data.

When possible, a facilitating practice to get empirical evidence from a scale would be to reuse, combine or calibrate the items of an existing one (Preibusch, 2013). Our questions, related to these three privacy dimensions, were designed by examining the literature on similar constructs applied in the general privacy context. Such selections were partially sustained by an existing literature review (Romare, 2024) within the context of if-this-then-than applications. The confidentiality dimension was developed by carefully selecting questions from (Cobb et al., 2020; Tamara Dinev and Hart, 2013; Maus et al., 2021) that belong to definitions such as integrity, secrecy, confidentiality, and secondary use of personal information. The transparency dimension was entirely developed by Awad and Krishnan to measure the customers' perceived information transparency in online personalized services and advertising (Awad and Krishnan, 2006). However, their instrument may not be applicable directly to the IoT context. Therefore, we performed substantial updates on the items (see Table D1) as practice done in (Dinev and Hart, 2005) considering the IoT Trigger-Action platforms and the presentation format to our participants with the scenario descriptions and images (Figure 2, Figure 3, Figure 4, Figure 5).

Table 2 Study part 1: Questionnaire related to privacy concerns and requirements repeated per scenario with Dimensions and Items.

Dimension	id	Item
Confidentiality	u1	I feel that as a result of using this app, my personal information may not remain confidential (Tamara Dinev and Hart, 2013)
	u2	I am concerned that this app may sell my information to other companies or institutions without my permission (Maus et al., 2021)
	u3	I believe that if I use this app, unauthorized people will have access to my data (Tamara Dinev and Hart, 2013)
	u4	I know that by using this app, an incidental data leakage may happen (Cobb et al., 2020)
Control	c1	I would be upset if this app unintentionally triggered and processed my data (Cobb et al., 2020)
	c2	I am concerned that I may lose control over my data by using this app (Cobb et al., 2020)
	c3	I feel I have control over my data in this app if the data collection happens in compliance with the Privacy Policies, Rules, and Standards (Awad and Krishnan, 2006)
	c4	For this app, I don't want it automatically running; I prefer to press a button before the action runs (Romare et al., 2023)
Transparency	t1	It's important for me to know what information the service companies involved in this app store about me in their database (Awad and Krishnan, 2006)
	t2	For this app, I want to receive a summary about the data processing that occurs (e.g., how my data are manipulated to produce meaningful information) (Awad and Krishnan, 2006)
	t3	It's important for me to know if the data from this app will be sold to third-parties for marketing purposes (Awad and Krishnan, 2006)

- **Confidentiality** addresses potential unintended disclosure through automated data transfers between trigger and action services.
- **Control** addresses the preference to initiate actions manually rather than having them run automatically.

- **Transparency** addresses desire about information collected and shared in cross-service data flows and sharing with third parties.

In Table 2, we provide the details of each item and the source from which they were initially derived. The items (statements) of the questionnaire were repeated four times, once after each scenario. Both the scenarios and questionnaire items were shown to participants in random order and answered by a 5-point Likert Scale from ”strongly disagree” to ”strongly agree” (e.g. a participant answers the questions in random order for each of the four scenarios, which appeared in a random order for each participant).

3.2.2 Data Sharing Preferences

The second part of our questionnaire is composed of features which describe a context that can have an impact on users' decisions for sharing their data, and thus can be important parameters for a privacy permission management system or access control system. These features serve as independent variables in our factorial vignette study: data category, the purpose of data sharing, and data recipient type. We chose these factors as they represent core elements for understanding the context of personal data processing and thus for judging the sensitivity of personal data, which is highly context-dependent (see also Census Decision Federal Constitutional Court of Germany (1983)). In addition, Art. 13 of the EU General Data Protection Regulation (GDPR) and other privacy laws require privacy notices to inform data subjects at least about these factors. These three factors are also directly aligned with the core structure of the IoT Privacy Label framework (Emami-Naeini et al., 2021), which defines them as essential for informing users about how their data is handled. Our factorial vignette study (4x3x3) was performed using the related sub-levels of each feature as shown in Table 3.

Table 3 Study part 2: factorial vignette study related to data sharing preferences.

Features	Sub-levels
Data Category	Location Data Image and Video Data Personal Information Message and Email Data
Purpose of Data Sharing	Main app functionality Personalised app functionality Targeted Advertising
Data Recipient Type	Third Parties Government and Legal Authorities Service Providers

Each participant answered a total of 36 yes-no questions, in a random order, in the form of: ”Would you accept the following for running an IoT application? Your {data category} is shared for {purpose of data sharing} and with {data recipient type}”.

3.3 Participants and Ethics

The study was conducted in compliance with the EU General Data Protection Regulation (GDPR) and the ethical review act of Sweden. The study design was reviewed and accepted by the data protection officer of Chalmers University of Technology and Karlstad University's ethics advisor. We asked for and obtained informed consent from the six experts we interviewed to design the study and from the participants of our online questionnaire (before they took part in the study). All collected data has been securely pseudonymized and is protected following the rules of the GDPR. We selected LimeSurvey as a survey tool because it has an EU-based (Germany) data controller with data processing policies that comply with the GDPR.

Table 4 Demographics from our participants sample.

Gender	**Age**	**Education**	
Female 48%	18-25 32%	Less than High School 1%	
Male 50%	26-35 33%	High School 34%	
Non-binary 1%	36-45 17%	University Degree 65%	
Prefer not to say ¡1%	46-55 11%		
	55+ 7%		
Region of Residence	**# Devices**	**TAP**	**Breach**
EEA 50%	0 18%	Yes 20%	Yes 37%
United States 50%	1-2 43%	No 80%	No 63%
	3-4 25%		
	5+ 14%		

We recruited a total number of 301 participants from the Prolific platform for our study (see subsection 3.2). We selected a balanced distribution between males and females (≈ 50% - 50%) in the European Economic Area (EEA) and the United States (US) as western regions, where IoT TAPs are more diffused [4]. Our gender-balanced sample was weighted towards younger participants, which we, however, considered as well acceptable, as current and future IoT TAP users should typically be younger. Moreover, we chose to include both participants who are using TAPs already, as well as participants who are not using IoT TAPs yet. Participants who use IoT TAPs already could likely be mostly classified as early adopters, who often tend to be ”privacy unconcerned” (Gollust et al., 2011). Therefore, we opted for not involving just the current users of IFTTT since they could introduce a bias in our study towards a less concerned user population, their extroversion that impacts information sharing (Lynn et al., 2017) and different perceptions of convenience (Lafontaine et al., 2021) - and future users might not have been well represented.

Participants were compensated 11£ / hour for completing the study, with a median time required of 10 minutes. Because of the relatively quick time of completing the questionnaire, we did not include any attention check questions, but instead, we combined the completion times with outlier analysis and Cronbach's α (Meade and Craig, 2012) in answering the 11 questions in each scenario in the first part (see Section 3.2.1)

[4] see also: https://6sense.com/tech/integration/ifttt-market-share (visited Nov 2023)

and the total time spent in answering the 36 yes/no questions in the second part related to the factorial vignette study (see Section 3.2.2). No outliers were detected with the Interquartile Range (IQR) rule.

3.4 Statistical Testing

For our questionnaire's reliability and validation, and to derive clusters, we adapted the procedure outlined in (Faklaris et al., 2019; Moritz Büchi and Latzer, 2017; Deng et al., 2005). First, we started with Exploratory Factor Analysis (EFA). After that, we investigated the validity (whether the questionnaire's items actually measure what we propose), reliability (whether the questionnaire's items are consistent), and global fit (how collectively the items are performing) of our questionnaire, by employing the Multi-Group Confirmatory Factor Analysis (MGCFA). Lastly, we implemented hierarchical clustering to obtain groups of participants to characterize the privacy concerns and requirements in clusters and later with data sharing preferences in profiles.

3.4.1 Questionnaire Validation and Reliability

The EFA is a data-driven methodology which provides relations between the items in the questionnaire and it is relevant for developing and refining the questionnaire instrument (Conway and Huffcutt, 2003). It gives the direction of potentially aggregating more than one item under a dimension. To verify the EFA results which discern the factors accounting for the correlation between observed variables without necessitating underlying theoretical frameworks (Thomas G. Reio and Shuck, 2015; Byrne, 2005), it is relevant to perform the Confirmatory Factor Analysis (CFA).

The CFA is essential to establish reliability and validity to ensure that the constructs assessed by the questionnaire accurately represent the theoretical concepts. When, as in our study, more groups or conditions are employed, the Multi-Group Confirmatory Factor Analysis (MGCFA) is suitable to ensure that the questionnaire is measured equivalently across groups or conditions. It has been used in technology acceptance model studies to measure the invariance across populations (Saritepeci et al., 2024; Li and Zhang, 2021) and contexts (Bansal et al., 2016). By testing the measurement invariance, we can verify that the questions are valid and with the same meanings in the four scenarios we proposed. Such a test implies five steps for demonstrating the full invariance among scenarios: 1) fit CFA models independently one from the others (baseline models); 2) fit MGCFA without constraining the factor loadings estimated (configural model); 3) fit the MGCFA by constraining the factor loadings to be the same among the four scenarios (metric model); 4) fit the MGCFA by constraining the factor loadings and the intercepts to be equal across the four scenarios (scalar model); 5) fit the MGCFA by constraining the factor loadings, the intercepts and the residuals to be equal across the four scenarios (full model). The demonstration of the measurement invariance process was finalized by comparing the Comparative Fit Index (CFI) between the ones obtained in step 2, step 3, step 4 and step 5 which is supported when the difference is ≤ 0.01 (Chen et al., 2010). When demonstrated, the measurement invariance ensures that the dimensions hold across different conditions, so measured equally across them (Putnick and Bornstein, 2016), allowing merging the questionnaire data from each scenario to perform the clustering. The overall performances of the MGCFA models are evaluated with global fit and by verifying the validity and reliability (Groß, 2023) considering Cronbach's α ($\alpha > 0.7$), Root Mean Square Error of Approximation of (RMSEA < 0.08), Standardized Root Mean Square Residual (SRMR < 0.08), Comparative Fit Index (CFI > 0.95), Tucker-Lewis Index (TLI > 0.95), Average Variance Extracted (AVE > 0.5), the HeteroTrain-MonoTrait (HTMT < 0.85) and the Congeneric Reliability ($\omega > 0.7$).

3.4.2 Extraction of Privacy Clusters and Profiles

Unsupervised learning is a branch of machine learning that includes clustering and dimensionality reduction algorithms. Hierarchical clustering is an agglomerative clustering algorithm and it is built in a tree form as a dendrogram. In agglomerative clustering, each data point starts as its own cluster, and clusters are merged iteratively considering the distance between those data points using distance metrics. Previous works employed hierarchical clustering to derive privacy clusters and then profiles to manage privacy preferences (Lin et al., 2014). Unlike algorithms such as K-means, hierarchical clustering does not require a pre-defined number of clusters. Instead, the number of clusters can be determined by cutting the dendrogram at a specific threshold or by evaluating the silhouette score (Shahapure and Nicholas, 2020), ranging from $[-1, +1]$. The dendrogram, created in the hierarchical clustering procedure, shows the levels of similarity between the clusters and allows a visual representation of how the participants, in our context, relate to their concerns and requirements.

K-means, another popular algorithm (Liu et al., 2014), requires the number of clusters to be specified in advance, which can influence the clustering outcome. Other methods used, and often combined, for such a quantitative analysis are principal component analysis, latent semantic analysis and non-negative matrix factorization (Salminen et al., 2020). Despite the hierarchical clustering being more computationally expensive (Abdalla, 2022), we chose to employ it due to the relatively small size of our dataset and the interpretability benefits.

To derive profiles, as defined in Section section 2, the clusters required multiple data sources to be combined based, for example, on correlated data (Inverardi et al., 2023). To this end, we describe the clusters based on privacy concerns and requirements with the data-preference features collected from the factorial vignette study procedure. The goal of the factorial approach is to gain a better understanding of an individual's judgement and decision-making by using scenarios that consist of features that vary according to their associated (sub)levels and transforming the clusters in profiles. The sub-levels are often presented in the context of a scenario, which serves to situate the survey participant in a specific context. In our work, we are focused on the isolation of those parameters, in terms of allowing or denying potential adoption, regarding their combinations.

4 Results

In this section, we present the results of our analysis. The first step is to ensure that our questionnaire is a valid and reliable measurement tool with the EFA and

MGCFA. We ran the MGCFA, since our study setup implied the repetition of the same questionnaire in four scenarios that belong to real IFTTT applications. The second step is to derive the privacy concerns and requirements clusters per scenario. Lastly, we used the privacy clusters as ground-truth labels to improve their characterization with the factorial vignette study procedure related to data sharing preferences.

4.1 EFA and MGCFA of our Questionnaire

We start our analysis by verifying the feasibility of running the EFA using the Kaiser-Meyer-Olkin factor adequacy (KMO). We ran independently such analysis for all four scenarios which resulted to be with a score of 0.90 for Scenario #1, 0.89 for Scenario #2, 0.90 for Scenario #3, and 0.90 for Scenario #4. We then run the EFA using the polychoric correlation matrix and the *oblimin* rotation method. We evaluate the EFA considering the factor loadings that are recommended to load with a threshold of > 0.32 and they should be dropped if the related cross load with more than one factor is above this threshold (Tabachnick et al., 2013) and when the commonalities are $h^2 \geq 0.40$ (Costello and Osborne, 2005). This exploration shows the possible structures of the model. The evidences suggested a transparency latent construct and an overlap between confidentiality and control latent constructs in all four scenarios. The items (see Table 2) for the transparency construct *t1, t2, t3* show communalities and factor loadings above the recommended thresholds and small cross-loadings. The candidate items for the confidentiality and control construct are *u1, u2, u3, c2*. We excluded *u4, c1, c3 and c4* due to cross loadings and since they do not achieve the commonalities threshold. The details of the EFA results can be viewed in Table 5 and in Appendix E. Thus, we tested and verified possible combinations of items under these two factors in the MGCFA procedure.

Following the steps exemplified in Section 3, we evaluated our questionnaire with the data-driven insights from the EFA by using the MGCFA. Our model was valid and satisfied the recommended metric thresholds by including the items *u1, u2, c2* for Confidentiality and Control's dimension, and items *t1, t2, t3* for the dimension related to Transparency. The four baseline models, each scenario fitted independently one from the other in the CFA model, were obtained with an RMSEA of 0.00 for Scenario #1 and #2 and of 0.03 for Scenario #3 and #4. RMSEA had excellent fits in all the scenarios, with the upper bounds for scenarios #3 and #4 that are with mediocre fits and excellent for Scenario #1 and #2. The SRMR was in all the scenarios between $0.02 - 0.03$. Similarly, in all the scenarios (baseline models), the values for CFI and TLI that we obtained were in the range of $0.99 - 1.00$ for both metrics. All the metrics explored are satisfying and all of them are providing the goodness fit of the models. The internal reliability was measured with Cronbach's alphas (α) and found in a range between $0.78 - 0.84$, considering independently all the scenarios. We integrate the average variance extracted (AVE) that, in all the scenarios, showed the convergent validity to be above the accepted threshold of 0.5. We measured the HeteroTrait-MonoTrait ratio of correlations (HTMT) to demonstrate discriminant validity which was satisfactory with a value of 0.75. We further calculated the Congeneric Reliability (ω) that in both confidentiality and control as well as transparency dimensions yielded a score of 0.84. The configural (scenarios together without constraints), the metric (scenarios

Table 5 Exploratory Factor Analysis among the four Scenarios (S) with loadings and communalities (h^2). Items meeting the thresholds are shown in **bold**.

Item	S1 MR1 / MR2	S1 h2	S2 MR1-MR2	S2 h2
u1	**0.78 / 0.10**	**0.72**	**0.82 / 0.05**	**0.72**
u2	**0.62 / 0.21**	**0.63**	**0.63 / 0.20**	**0.61**
u3	**0.90 / -0.16**	**0.63**	**0.89 / -0.12**	**0.67**
u4	0.45 / 0.16	0.33	0.43 / 0.29	0.43
c1	0.33 / 0.48	0.57	0.37 / 0.40	0.49
c2	**0.71 / 0.19**	**0.73**	**0.76 / 0.11**	**0.70**
c3	0.39 / -0.25	0.07	0.47 / -0.29	0.13
c4	0.42 / 0.18	0.31	0.48 / 0.18	0.38
t1	**0.02 / 0.86**	**0.77**	**0.13 / 0.74**	**0.69**
t2	**0.08 / 0.69**	**0.57**	**-0.04 / 0.76**	**0.55**
t3	**-0.01 / 0.84**	**0.69**	**0.01 / 0.88**	**0.78**
Item	**S3 MR1-MR2**	**S3 h2**	**S4 MR1-MR2**	**S4 h2**
u1	**0.80 / 0.07**	**0.71**	**0.76 / 0.08**	**0.66**
u2	**0.66 / 0.21**	**0.64**	**0.51 / 0.30**	**0.57**
u3	**0.83 / -0.04**	**0.65**	**0.83 / -0.10**	**0.58**
u4	0.24 / 0.53	0.50	0.48 / 0.14	0.34
c1	0.35 / 0.49	0.56	0.31 / 0.58	0.60
c2	**0.90 / 0.00**	**0.81**	**0.80 / 0.10**	**0.76**
c3	-0.52 / 0.37	0.18	0.49 / -0.24	0.12
c4	0.53 / 0.12	0.37	0.11 / 0.64	0.53
t1	**-0.01 / 0.93**	**0.85**	**-0.06 / 0.96**	**0.84**
t2	**0.15 / 0.67**	**0.59**	**0.10 / 0.71**	**0.62**
t3	**-0.03 / 0.90**	**0.78**	**0.01 / 0.83**	**0.69**

together with the fixed factor loadings constraint), the scalar model (scenarios together with fixed factor loadings and item intercepts) and the full model (scenarios together with fixed factor loadings, intercepts and residuals) provided satisfactory results that respected all the metrics' thresholds. Thus, we accepted the full invariance among the four scenarios since the $\Delta CFI \leq 0.01$ between the configural, the metric, the scalar and the full models. A summary of all the performances related to the metrics for the MGCFA can be seen in Table 6 and are in line with the literature recommendations to assure reliability, validity and global fit of the CFA model (Groß, 2023).

These results confirm that our questionnaire consistently measures the intended constructs across all scenarios. Thus, we can continue with our validated measurement tool (e.g., questionnaire) to perform the clustering considering the scenarios that we proposed including the information about their potential privacy risks.

4.2 Privacy Clusters and their Characterization with Profiles

To answer the research question **RQ1**, we proceed with the data analysis by employing the hierarchical clustering using the Euclidean metric and the Ward method. The Ward method aims to minimize the increase in variance when merging clusters (Jr., 1963). We obtained the input data for the clustering algorithm from the *lavPredict* function (*lavaan* package in R) which estimates latent response values for each item by

Table 6 Multi-Group CFA global fit performances with Configural, Metric, Scalar and Full models regarding the measurement invariance.

Model	RMSEA (CI 90%)	SRMR	CFI	TLI
Scenario #1	0.00 (0.00 - 0.07)	0.02	1.00	1.00
Scenario #2	0.00 (0.00 - 0.08)	0.02	1.00	1.00
Scenario #3	0.05 (0.00 - 0.10)	0.03	0.99	0.99
Scenario #4	0.05 (0.00 - 0.09)	0.02	0.99	0.99
Configural	0.03 (0.00 - 0.06)	0.02	0.99	0.99
Metric	0.03 (0.00 - 0.06)	0.03	0.99	0.99
Scalar	0.05 (0.03 - 0.07)	0.04	0.99	0.98
Full	0.05 (0.03 - 0.07)	0.05	0.98	0.99

utilizing the thresholds corresponding to the ordinal answer options as in (Biselli et al., 2022). The predicted score values are then derived by calculating the weighted sum of these latent values for each factor, where the factor loadings serve as the weights.

Since the proven measurement invariance was established across the configural, metric, scalar and full models, we considered the two dimensions (e.g., Control and Confidentiality, and Transparency) across all the scenarios. Thus, we tested 8 parameters (2 dimensions * 4 scenarios) with the number of clusters (k) from 2 to 7. The optimal k was decided by visualizing the dendrogram (see Figure E1) and maximising the silhouette score. All together the scenarios reached their maximum values in the silhouette score when the $k = 2$ with 0.31, but considering the dendrograms, we opted for $k = 3$, which achieved a silhouette score of 0.21 To select the number of clusters, we also considered a trade-off between the data-driven approach (induction) as well as the consideration of our goal for a potential real-world implementation (deduction) (Salminen et al., 2022). Specifically, the silhouette scores varied by no more than 5% between $k = 2$ and $k = 3$ considering the $[-1, +1]$ silhouette score range, and the identification of only two clusters would not reflect and capture a broader range of privacy concerns and requirements. More precisely deriving three, instead of two clusters as a basis for expressing data sharing preferences, will allow us to subsequently derive more fine-grained bundles of privacy permission settings for at least three different types of users.

Table 7 shows the results of how participants populated the three clusters in each scenario with related descriptive statistics divided by questionnaire items. We named the three clusters as follows:

- Basic Privacy
- Medium Privacy
- High Privacy

Users are in general demanding transparency and show concerns regarding confidentiality and control - these two dimensions of the questionnaire positively correlate - considering the high score answers from the participants that answered the questionnaire (see Table 7 for items' mean, standard deviation and median aggregated across scenarios), thus privacy concerns and requirements for transparency are basically visible for all privacy clusters and need to be addressed in some form for users in all clusters as important privacy feature (and for this reason, we chose the name ”Basic Privacy” instead of ”Low Privacy” for the cluster with the lowest privacy concerns and requirements). We did not find any statistically significant evidence in relation to the demographics (including gender, age or region) and the privacy clusters.

Table 7 Mean, Standard Deviation, and Median per item— after calculating the mean among the four scenarios— within each privacy cluster with %=Number of Participants in percentage.

Privacy Level (% Part.)	u1	u2	c2	t1	t2	t3
High Privacy (19%)						
Mean	4.28	4.42	4.19	4.61	4.44	4.67
Standard Deviation	0.66	0.80	0.75	0.55	0.61	0.51
Median	4.50	4.75	4.25	4.75	4.50	5.00
Medium Privacy (65%)						
Mean	3.63	3.79	3.69	4.18	4.01	4.29
Standard Deviation	0.60	0.67	0.64	0.53	0.65	0.56
Median	3.75	4.00	3.75	4.25	4.00	4.25
Basic Privacy (16%)						
Mean	3.14	3.22	3.04	3.37	3.38	3.28
Standard Deviation	0.84	0.87	0.83	0.79	0.80	0.86
Median	3.12	3.12	3.00	3.25	3.50	3.25

In the **Basic Privacy cluster** (16%), the participants are overall less concerned than the participants of other clusters and have fewer requirements for transparency, with mean values between 3.04 and 3.38 across all dimensions Table 7). These scores indicate a generally neutral placement, reflecting limited sensitivity to privacy risks. Nevertheless, the cluster still shows a consistent expectation for transparency (*t1–t3*), where the means (3.28–3.38) are slightly higher than those for confidentiality and control (*u1–u2-c2*, 3.14–3.22 – 3.04). This suggests that even participants with basic privacy concerns value being informed about how their data is handled and have protection and control of their personal data flows in the TAP applications. Among scenarios, the highest scores were observed for Scenario #2 (”lock the door”), while the lowest was in Scenario #1 (”sleep log on cloud calendar”), at 3.07 slightly suggesting that tangible or verifiable actions generate, in comparison, more concerns even in the basic privacy cluster than a nowadays widely adopted solution of connecting health data from smartwatch (e.g., often one of its main functionality) with other smartphone services such as the calendar.

Participants belonging to the **Medium Privacy cluster** (65%) have higher concerns and requirements than the ones in the Basic cluster and overall, moderate requirements, with means ranging from 3.63 to 4.29. Transparency dimensions consistently scored higher than control and confidentiality. Scenario #2 (”lock the door”) reached the highest scores within this group, while Scenario #1 (”sleep log on cloud calendar”) showed the lowest, suggesting, similar to the Basic Privacy cluster, that physical security actions elicit stronger concerns and requirements than cloud-based automation also in this cluster.

The **High Privacy cluster** (19%) is the one with participants that showed overall highest concerns and requirements. The tendency is similar to the other clusters, participants demand more transparency than they are concerned about confidentiality and control. However, their mean scores are less different compared to the other two clusters with all of them exceeding 4.19 and medians approaching the maximum of the scale. Transparency was the dominant requirement across all scenarios. Regarding the scenarios, all means are above 4.32. Scenario #2 again achieved the maximum value within this cluster (4.63), while Scenario #1 recorded the lowest (4.32), although still close to the top of the scale. So, participants are concerned about losing control, and they demand transparency about the processing of their data, including how their data will be used and who will use them. The unpacked descriptive statistics per cluster considering the different scenarios can be found in Table 8.

Once we obtained the cluster labels, we trained logistic regression classifiers. In our procedure, the supervised learning algorithms have the goal of predicting the cluster labels from the ordinal scale data directly collected from the participants regarding the attitudinal privacy concerns and requirements asked in the first part of the questionnaire. The Scenario #1 data can predict the cluster label with 58% of balanced accuracy. Scenario #2, it is 63%, followed by 69% for Scenario #3 and Scenario #4. The performances increased when combining two scenarios data to predict the cluster label. The combination of Scenario #1 and #2 achieve 80% of accuracy, Scenario #3 with #4 the 75%, Scenario #1 and #3 the 81%, Scenario #2 and #4 the 76%, Scenario #1 and #4 the 73% and finally Scenario #2 with Scenario #3 the 82%. All scenarios together as input data for the logistic regression classifier reach 85% of accuracy. Thanks to these logistic regression models, it will be possible to classify a new user into a specific privacy cluster by skipping the transformation employed to get the input data for the clustering algorithms (e.g., MGCFA and *lavPredict* function) and just answer the six questions of our questionnaire in a scenario where information about the privacy risks is provided, while still recommending the use of Principal Component Analysis (PCA) to verify that the developed scales' items load correctly (Islami et al., 2024).

Table 8 Mean, standard deviation, and median for scenarios S1–S4 within each privacy cluster, grouped in scenario pairs.

Privacy Level	Mean	Std	Median	Mean	Std	Median
	Scenario 1			**Scenario 2**		
High Privacy	4.32	0.74	4.50	4.63	0.46	4.83
Medium Privacy	3.61	0.71	3.67	4.17	0.61	4.17
Basic Privacy	3.07	0.82	3.00	3.53	0.87	3.50
	Scenario 3			**Scenario 4**		
High Privacy	4.40	0.73	4.58	4.38	0.61	4.50
Medium Privacy	3.92	0.60	4.00	4.02	0.65	4.00
Basic Privacy	3.14	0.83	3.17	3.21	0.85	3.25

To answer the **RQ2**, we further characterized the clusters by including the data obtained in the second part of the questionnaire related to the data sharing preferences.

In this regard, we describe the clusters with the frequencies of the isolated features (of the types of data categories, purpose of collection, and data recipient types) from the factorial vignette study procedure according to the related cluster labels. The participants resulted in a distribution where 16% of them are in the Basic Privacy Cluster, 65% in the Medium Privacy Cluster and 19% in the High Privacy Cluster.

We plotted the percentage of data sharing to get an intuitive overview of how data category, purpose of data sharing and data recipient type varied among the three clusters (see Figure 6). In this second part of the study, we collected higher percentages of no than yes answers related to their potential acceptance to run an IoT app that shares data for all features with yes answers not exceeding 35% for all features. We focused our attention on the differences in the yes answers (i.e. the difference in the acceptance to share data) by characterizing the 3 derived privacy clusters with the privacy data sharing preferences per feature.

As Figure 6 illustrates, the clustering of users into the three hierarchical privacy clusters (Basic, Medium, High), which we derived from the privacy attitudinal concerns and requirements questionnaire (part 1), is confirmed by the data sharing preferences factorial vignette study procedure (part 2). The factorial vignette study showed that, across the three groups, the preference for potentially accepting data sharing decreased from the High Privacy cluster to the Medium Privacy cluster, and then to the Basic Privacy cluster (even though, as mentioned, users in all clusters rather prefer to not accept data sharing). Correlating the data with each privacy cluster yields corresponding privacy profiles (Inverardi et al., 2023).

Participants belonging to the **Basic Privacy profile** showed the highest acceptance to share data by running an IoT app, even though the acceptance was still on average $\leq 35\%$ for all features. For this cluster, under the data category feature, location data is the type of data that they prefer to share more than others, followed by message and email data, while image and video data are the data types with the lowest willingness to share. The purpose of data sharing, do not differ among the main app functionality and the personalized one, while the targeting advertising is the less accepted among those. No particular percentage differences were detected regarding the data recipient, with a slightly more willingness to share to government and legal authorities.

The **Medium Privacy profile** got a lower percentage for a potential acceptance of sharing location data than the Basic privacy cluster and they are less willing to share message and email data with a difference of 13% on average. Putting in a rank the purpose of data sharing sub-levels, higher acceptance of sharing is for the main app functionality followed by the personalized one and ending with targeting advertising. As for the Basic Privacy cluster, no particular differences were found in considering the data recipient type.

The **High Privacy profile** showed the overall lowest scores compared to the other two clusters, with less than the $\leq 20\%$ acceptance to use an app that shares data for all features on average. The purpose of data sharing had a similar pattern as the Basic Privacy cluster, but with 15% less of acceptance as expected for the cluster who demand more privacy than the others. The data recipient type had the service providers as higher acceptance to share with followed by the government and legal authorities and finally the third parties.

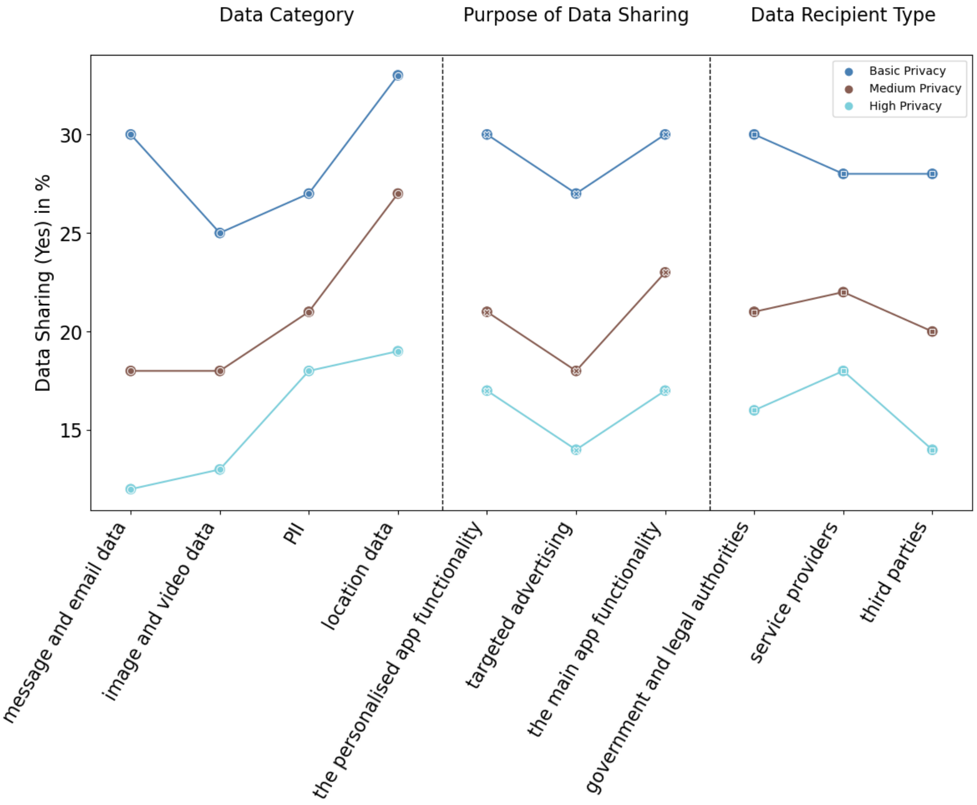


Fig. 6 Percentage of Data Sharing in relation to Data Category, Purpose of Data Sharing and Data Recipient Type

5 Discussion

The key results in relation to our research questions can be summarized as follows: we show that when participants are informed about the privacy risks of leaked data of our selected IoT TAP scenarios, three privacy clusters (High Privacy, Medium Privacy and Basic Privacy) can be derived for IoT TAPs (**RQ1**). To this end, we developed, used and validated a novel questionnaire as a reliable procedure for clustering participants for conducting our study. We further described the clusters with data sharing preferences to extend them in privacy profiles (**RQ2**) as a step towards finding bundles of privacy settings that could be recommended to users. No evidence was found regarding the impact of the demographics that we collected and the profiles. These key results and their implications will be discussed further in this section.

Clustering or segmentation processes aim to identify groups of users to better understand their characteristics and needs. A critical step in these methods involves designing a capable measurement tool to collect the necessary data. The data collected by these tools typically includes data collected using questionnaires, behavioral data, or a combination of both (Salminen et al., 2020). In this work, we validate a questionnaire with two dimensions: Control and Confidentiality, and Transparency to gather privacy concerns and requirements which were analysed to form clusters. These clusters serve as an approximation of the complex decision-making processes involved in setting data sharing preferences. In contrast, current approaches for configuring privacy settings — whether in online services via browsers or on devices through smartphone interfaces — often present users with an abundance of information that they must first comprehend before making informed choices. This can lead to a significant risk of users either making decisions without full awareness due to cognitive overload or lack of knowledge, or even worse by neglecting the process entirely. Thus, the privacy profiles we derived offer an initial framework to guide semi-automated decisions based on users' privacy concerns. Further improving the profiles could later enable individual users to receive recommendations informed by the preferences of similar users. Over time, as users become more familiar with the permission management system, they can additionally personalize their privacy settings or change to a different cluster to adopt more or less stringent privacy controls.

5.1 Novelty of the Privacy Clusters

Our three privacy clusters were the first privacy clusters that were derived for the context of IoT TAPs for users who were informed about potential risks. Privacy clusters and profiles derived by earlier related work, which we reviewed in Section 2, were developed for other domains and scopes (e.g. IT or IoT in general and not specifically for IoT TAPs, with users that were not necessarily informed about potential privacy risks and consequences). Also, for this reason, privacy clusters and profiles that were developed earlier for other domains/scopes also differ content-wise from our derived profiles. In particular, the well-known three privacy clusters/categories (”privacy unconcerned”, ”privacy pragmatists”, ”privacy fundamentalists”) that Alan Westin derived (Westin, 1968), differ not only domain- and scope-wise but also content-wise significantly from our three privacy clusters (Basic Privacy, Medium Privacy, High Privacy). All participants in our three clusters (including our “Basic Privacy cluster”) share privacy concerns regarding confidentiality and control over IoT TAPs, still to different degrees, possibly also for the reason that our participants were informed about potential risks and consequences. In contrast, users in Westin's cluster “privacy unconcerned” generally do not share privacy concerns and trust organizations processing their personal information and are comfortable with existing organizational procedures. Also, other related works reviewed by Inverardi et al. (2023) that derive three privacy clusters for other domains include, in contrast to our work, a cluster for ”privacy unconcerned”. Moreover, participants in all three of our clusters are, in general, demanding transparency, a dimension that is not directly captured by Westin's categories or by most other privacy clusters.

More details on how our results can facilitate the users' decision-making process regarding IoT TAP permission and controls are discussed in the subsection 5.3 and subsection 5.4.

5.2 Informing on risk

Our study was based on a questionnaire including IoT TAP application scenarios together with a description of the risks that potential adversaries could infer sensitive information from these applications. As discussed above and in line with related work Section 2, we assume that informing users about privacy risks is relevant for their decision-making (Morgan et al., 2022) and suggest that IoT TAP users may also need help to understand and reason about privacy risks for developing preferences and concerns (Saeidi et al., 2022). Moreover, our survey results, which show an overall high requirement for transparency about data processing and storage by participants, are also motivating our approach of informing users about possible personal data inferences that could be made when they decide about permitting the use of their data for an application.

The GDPR also well recognizes the importance of informing users about potential risks in the context of automated decision-making by requiring in its Art. 13 (2) (f) and 15 (1) (h) that data subjects need to be informed about the significance and the envisaged consequences of such processing. While these rules are rather meant for regulating automated decisions made by data controllers without involving the data subjects, our results rely on the literature assumptions (see Section 1) that, even for automation by IoT apps initiated by the users, transparency about privacy risks and their potential consequences is important for users' decision-making when installing apps or setting permissions.

However, since app providers may not be interested in informing about potential risks and consequences in an unbiased manner, the question of who should provide this risk communication and how it should be provided still needs to be addressed. Another possible direction could be the development of transparency-enhancing tools (TETs) that provide information on potential privacy risks for different types of TAP applications (Breve et al., 2022). TETS can help individuals to exercise their rights for transparency (e.g. pursuant to Art. 12- 15 GDPR), and reduce information asymmetry by providing users with insights about the providers' data collection and usage and possible consequences that might arise (Murmann and Fischer-Hübner, 2017; Zimmermann, 2015). A TET providing transparency about risks and potential consequences for IoT TAP applications could become part of or could be invoked by future permission control systems implemented for TAPs. When an IoT TAP application of a certain type is installed or first run by users, or at least at the moment when an action by such an application is triggered that requires to transfer of personal data to a third-party entity, a permission control system at the TAP would request the users for providing their permissions for executing the app or transferring personal data to the third party entity (unless such permissions have already previously set by the users). In this context, the TET could provide or link users to information about potential risks and consequences related to applications of that type.

5.3 Towards usable privacy permission and control profiles

This survey was conducted within the scope of a larger research project on IoT TAP privacy, which has as a future goal to develop usable and transparent privacy control systems. Usability should be achieved via semi-automation with the help of profiles (”bundles”) of privacy permission and control settings that can be easily chosen or (semi-)automatically suggested and assigned ”on the fly”.

Our results show that three hierarchical privacy clusters capturing Basic, Medium and High attitudinal privacy concerns and requirements can be derived for TAP applications for individuals who are informed about potential risks. Additionally, the factorial vignette design confirmed that each of the three privacy clusters could be described by one of three distinct attitudinal data sharing preference profiles reflecting users' preferences for Basic, Medium and High privacy protection.

However, the derived profiles of attitudinal data sharing preferences cannot directly be transferred to different profiles of data sharing permission settings since for all three data sharing preference profiles, users rather preferred not to accept the sharing of their data for different data categories, purposes and data recipient types, even though the extent to which they were accepting data sharing was increasing from profiles that can be associated with the Basic to the Medium and to the High Privacy clusters. In compliance with the Data Protection by Default principle of Art. 25 GDPR, users should, by default, be assigned to a Data Protection by Default profile with the most privacy-friendly permission settings. Based on our results, no further profiles can be suggested and offered as alternative profiles which different types of users could then pick. Nonetheless, our research should be complemented in future with an analysis of behavioral privacy data sharing preferences (Salminen et al., 2020), especially in the context where users have an interest in using certain types of IoT TAP apps and may therefore be more willing to accept data sharing for being able to use the applications. This could be done by using the clusters as classification labels to obtain more specific profiles of implementable permission settings that can offer users tailored privacy controls that align with their typical preferences as in (Lipford et al., 2022). As we expect, this may lead to profiles with behavioral data sharing preferences that have on average higher preferences for accepting data sharing than the average attitudinal preferences, and may still capture increasing sharing preferences from Basic to Medium and to High and may thus still be useful for characterizing our three privacy clusters (to be confirmed by our future work). Any profiles with high behavioral preferences/acceptance for data sharing (e.g., above 85%) for certain items could lead to predefined profiles with permission settings authorizing the data sharing, and these profiles could be chosen and offered to users belonging to the cluster, which the respective preference profile characterizes. Also, machine learning-based privacy assistants, with a design similar to (Das et al., 2018), running on the users' (mobile) devices (i.e., under their control) could be implemented and used to evaluate the users' decisions for developing or adapting privacy profiles for IoT TAPs based on the behavioral user data. Those decisions might have different sources, and it should be up to the user to decide which one to use, for example, from expert opinions or other users' opinions in a collaborative manner (Colnago et al., 2020; Lipford et al., 2022)

Even though our derived attitudinal data sharing preference profiles cannot lead to predefined profiles of data sharing permission settings, our three clusters capturing both increasing privacy concerns, from Basic to Medium and to High, regarding the loss of control as well as increasing transparency requirements, can be used as a basis for deriving and suggesting three profiles of settings related to increasing user control and transparency. We describe these three different profiles based on our findings and adapt them to the actual essence of IoT TAPs which is automation.

We assume that users in the High Privacy cluster prefer a profile with high privacy controls and low automation. The users in this cluster should therefore always (even if they have already provided informed consent, i.e., set permissions, at the time when the app was installed) be asked to confirm the execution of triggered actions involving certain types of data transfers and processing at third-party servers (”control before final action”), especially if this triggering of actions may impose higher risks (if e.g., it involves personal data that is perceived as sensitive) that they could then be again informed about.

Users in the Medium Privacy cluster find themselves in the middle score profile. Their evaluations have been found with higher variability under the data category and purpose of data sharing, while more stable in terms of data recipient type. They might be asked again to confirm any previously (at installation or run-time) given informed consent in regular intervals when executing apps or sharing their data via IoT TAPs based on the purpose and data category, e.g. with a once-per-week option (or other time intervals suitable for a specific app). Requesting a re-confirmation also allows reminding the users of any potential risks on a regular basis, which they may then be able to contextualise better after having been using the app.

Users in the Basic Privacy cluster, on the other hand, may rather trade high automation and thus more convenience for less control and less privacy. They may thus only be asked for consent for data processing and transfers, and receive information about any potential risks at app installation time.

Privacy profiles deliver solutions that represent groups. Despite that, we interpret these profiles as a starting point for users to eventually adapt, and thus personalise their privacy settings as single individuals (Marky et al., 2024). Users can receive profile suggestions (with a summary of the profile that suits them best (Marky et al., 2024)), based on the profiles to which they belong, by answering our validated questionnaire in the same setting as our study, which employed the application descriptions and privacy risks.

Privacy profiles can also be semi-automatically developed or updated and personalised ”on the fly” depending on how the users behave and interact with consent requests when installing an app, using it for the first time, or when consent needs to be dynamically re-obtained. This will allow us to consider also behavioral privacy preferences. For instance, if a user installs an app ”Share my location on Slack”, which at certain time intervals posts the user’s location on Slack, at app installation, the user consents to share their location. However, if the users’ location becomes sensitive in a certain context (e.g. at specific times when a religious service or political event takes place at a certain location), explicit consent needs to be dynamically requested and obtained according to Art. 9 GDPR. If users, especially those who do not belong to a profile representing the High Privacy cluster, refuse to provide explicit consent, they can be asked if they would like to adapt their permission settings to stricter ones.

5.4 Too varied to generalize—context-specific scale and profiles as a way forward

IoT TAPs have very different characteristics when it comes to different data processing practices, such as the type of data in use, purpose of data usage, data transfer, and data flows between different parties, to name a few and all impact users' data sharing preferences and concerns in this context. Considering the substantial number of IoT TAP applications—numbering in the thousands—it was not feasible to include all possible variations without overwhelming participants or introducing cognitive fatigue. Capturing such vast variability is challenging, especially considering that privacy concerns and preferences are highly contextual. Consequently, in our study, we employed four carefully selected IoT TAP scenarios, which were formed and filtered based on interview results with experts, see Section 3, to capture a broader spectrum of user concerns and preferences. Despite this, our questionnaire is not limited to the scenarios we selected, since during the election period, we considered a widely categorized group of applications in IoT TAPs, and each of the selected scenarios is a representative of a group of scenarios based on certain characteristics, such as personal data sensitivity and data recipient types. For the same reason of avoiding participants' cognitive fatigue, we captured their concerns and preferences using a relatively small number of questions in the questionnaire and features in the factorial vignette study procedure concerning three dimensions of data protection goals that proved to be role-players in users' privacy concerns and preferences.

However, while our questionnaire was designed to be broadly applicable, capturing concerns across key dimensions of data protection goals (Hansen et al., 2015), further evaluations are necessary to determine if our developed questionnaire can reliably capture privacy concerns and preferences across various IoT TAP applications, especially those with specific privacy risks not covered in our study (Butori and Miltgen, 2023). For instance, applications involving technologies like biometric data or those enabling complex data-sharing networks may present unique challenges and privacy risks not addressed in our study. Additionally, these evaluations will help in deriving a comprehensive scale to accurately capture users' preferences and concerns within this context. This will ensure that the resulting profiles remain valid and reliable for a broader range of IoT TAP scenarios. Nonetheless, we should point out that it may not be feasible to generalize privacy concerns and preferences and consequently, the clusters across various ranges of IoT TAPs applications. Further scenario-specific validation is essential for future standardization.

Indeed, data practices within the IoT TAPs context are nuanced and varied, often differing significantly between applications due to factors like technical complexity, regulatory environment, and type of data processed for different purposes. This underscores the importance of tailoring privacy management strategies to specific (groups of) applications rather than assuming a one-size-fits-all approach. It also highlights the need for ongoing monitoring and adaptation of these solutions to accommodate changes in data practices over time (Smith et al., 1996). Ultimately, recognizing the application-specific nature of privacy concerns in IoT TAPs can lead to more effective and user-centric privacy management approaches. Future research should focus on expanding the scope of scenarios and exploring adaptive privacy settings that respond to the evolving landscape of IoT technologies.

5.5 Limitations

First, we recognize that data practices across IoT TAP applications are complex and diverse; therefore, we highlight the need of customizing privacy management strategies to each context rather than applying a one-size-fits-all solution. As discussed in more detail in Section 5.4, our results are based on data collected in four IoT TAPs selected scenarios. Although the scenarios represent a range of common characteristics in IoT TAP applications, the complexity and abundance of significantly different data practices and the continual emergence of new technologies and applications mean that additional research could help expand our findings by exploring additional scenarios and adaptive privacy settings that respond to the evolving landscape of IoT technologies.

The results of this study focus on privacy attitudes rather than observable behaviour. While investigating attitudes is relevant—as they reflect how much users care about the processing and sharing of their data—it may not fully represent actual user behaviour. We acknowledge the existence of a privacy paradox, where individuals' behaviours do not align with their expressed attitudes. In the complex IoT context, users often prioritize functionality and convenience over privacy concerns, potentially leading to discrepancies between stated preferences and actions, i.e. privacy paradox. However, users' concerns and preferences still represent important demands that need to be considered when developing tools, ensuring that privacy settings and protections align with these dispositions. It remains complicated to investigate actual privacy behaviour due to various factors, such as the lack or absence of privacy management tools where to gather the privacy settings or logs like in Android, as well as the vast diversity of applications and contexts in which they are used (Barth and de Jong, 2017). These complexities further highlight the importance of understanding privacy attitudes as a foundation for designing systems that can accommodate users' expectations and concerns.

In the online survey, 301 participants answered the questions. Looking at Table 4, most of our participants were young. Among them, 65% were below 36 years old. Furthermore, we focused on people living in the EEA and the US, without considering other regions where IFTTT is present. The US and EEA are Western regions with similar basic privacy principles/standards that are enforced by laws (GDPR, US state privacy acts and federal consumer protection laws) and self-regulation and where IFTTT is broadly in use. While this focus ensures some consistency in the legal context, it limits the applicability of our findings to other regions with different cultural attitudes toward privacy and varying regulatory environments. Nonetheless, our survey evaluations showed no evidence that the demographic background, including the region (US, EEA), age or gender, could have any significant impact on our results.

Moreover, 80% of participants had never used or heard of IoT TAPs like IFTTT. Including non-users was intentional to capture potential future users and avoid bias toward early adopters who might be less privacy-concerned. However, this choice may have influenced the results, as non-users might have different perceptions or heightened concerns due to unfamiliarity with the technology. The hesitation to share personal data observed in our results (see Section 4) might be partially attributed to this lack of prior experience. Still, as we discussed above, if we had only conducted the study with IoT TAP users as early adopters, our data set would have likely been biased to represent rather the privacy unconcerned individuals, and hence we made a conscious choice to include a broader range of participants.

Furthermore, the participants were exposed, together with the description of the IoT TAP application, to related potential privacy risks in the first part of the study; thus, this probably impacted their answers. However, as we motivate in Section 1 and in Section 5.2, informing users was an essential choice in our study setup, for enabling users to make informed and conscious decisions.

6 Conclusion

In an era where IoT TAPs are becoming integral to the personalisation and automation of everyday environments, the challenge of managing privacy effectively has never been more pressing. Our study addresses this issue by exploring how current and potential IoT TAP users' data sharing preferences and concerns can be systematically understood and managed. We showed that when users are informed about privacy risks, their privacy concerns can be clustered using a validated two-dimensional questionnaire focused on confidentiality and control, as well as transparency. This clustering approach was further described using factorial vignette study features into privacy profiles considering data sharing preferences such as data category, purpose of data sharing and data recipient type that linked, as the first step, the privacy disposition to a usable permission and control management system.

Our findings are based on data collected from four selected IoT TAP scenarios. However, these scenarios include a range of common characteristics within IoT TAP applications, the complexity and number of significantly different data practices, coupled with the continuous emergence of new technologies and applications in the context of IoT TAPs, indicate that additional research could enhance our findings. Therefore, our research suggests that the variability in data sharing preferences and concerns across different IoT TAP applications requires a context-specific approach to privacy management. Privacy management permission systems should be tailored to specific (groups of) applications, with ongoing monitoring and adaptation to accommodate changes in data practices and user expectations over time.

Future work on implementing a comprehensive permission management system is necessary. The next step following this study might involve the development and evaluation of user interfaces or dashboards that allow users, assigned with a profile, to monitor data sharing activities within IoT TAP applications in a transparent way capable to iterate and update dynamically the privacy profiles based on the actual user choices. Such a procedure would implement the privacy recommendations to future IoT TAP users also by showing anonymized comparisons with users of similar profiles.

Through the interaction with a transparency dashboard, it may be feasible to incorporate behavioral measures as well to refine, adapt and further validate the privacy profiles based on IoT TAP automation applications review, acceptance or rejection. Additionally, a subsequent phase would focus on enhancing the privacy profiles created in this research by incorporating, through an additional or extended questionnaire, the privacy principles choice and notice/transparency (Iravantchi et al., 2025) with finer granularity (e.g., how frequently a user wants to be notified, via what channels they want to be notified and where they share their data). These enhancements would assist developers in understanding how frequently each user related to a specific profile prefers to receive notifications and how privacy information should be communicated effectively (see also Schaub et al. (2015); Feng et al. (2021) for the discussion of different choice and notice design spaces). Utilizing these clusters as classification labels could enable the creation of more refined privacy profiles including behavioral measures as well as more features or more fine-grained existing features regarding for example sensor types, retention time and data storage location in addition to purpose of collection, the data categories, the recipient type that we used in the current study.

Acknowledgements

We thank all our participants and those who contributed to the scenario selection. We thank our anonymized reviewers for their useful comments and feedbacks. This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Appendix A Introduction and Demographics

Useful Information

”The Internet of Things (IoT) describes devices with sensors, processing ability, software, and other technologies that connect and exchange data with other devices and systems over the Internet or other communications networks” (Wikipedia definition retrieved Jan 2024).

This questionnaire is related to IoT platforms that are online services allowing end-users to connect services and devices through a simple recipe. This recipe follows the formula: IF THIS THEN THAT. For example: if you enter a room, then turn on the lights. In this example, the trigger is when the end user enters a room (where a sensor detects presence) and the action is turning on the lights.

Between the trigger and action there is an interconnection via the IoT platform. Due to this interconnection, the trigger company provider and the action company provider automatically communicate and know the data related to the app.

- Where are you from? *(Options: European Economic Area, United States of America)*
- What’s your age group? *(Options: 18-25, 26-35, 36-45, 46-55, 55+)*
- What’s your gender? *(Options: Woman, Man, Non-binary, Prefer not to say)*
- How many IoT devices do you own? *(Options: 0, 1-2, 3-4, 5+)*
- How much time per day do you interact with your IoT devices on average? *(Options: 0’-30’, 30’-60’, 60’-120’, 120’+)*
- Have you ever had experiences with a data breach (e.g., someone stole your data by hacking an online service)? *(Options: Yes, No)*
- Have you ever heard of or used an IoT Trigger-Action platform such as IFTTT, Microsoft Power Automate, or Zapier? *(Options: Yes, No)*
- What is your highest level of education? *(Options: Less than High School, High School, University degree)*

You will be presented with 4 scenarios where such interconnection is used, followed by 11 statements. You will have five response options: strongly disagree, disagree, neutral, agree, strongly agree. After the 4 scenarios, additional yes/no questions will be asked.

Appendix B Study part I: Questionnaire

Questionnaire Items

Here are the 11 questions of the questionnaire. Note that both the order of questions and the order of scenarios were randomized.

1. I feel that as a result of using this app, my personal information may not remain confidential *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
2. I am concerned that this app may sell my information to other companies or institutions without my permission *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
3. I believe that if I use this app, unauthorized people will have access to my data *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
4. I know that by using this app, an incidental data leakage may happen *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
5. I would be upset if this app unintentionally triggered and processed my data *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
6. I am concerned that I may lose control over my data by using this app *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
7. I feel I have control over my data in this app if the data collection happens in compliance with Privacy Policies, Rules, and Standards *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
8. For this app, I don't want it to run automatically; I prefer to press a button before the action runs *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
9. It's important for me to know what information the service companies involved in this app store about me in their database *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
10. For this app, I want to receive a summary about the data processing (e.g., how my data are manipulated to produce meaningful information) that occurs *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.
11. It's important for me to know if the data from this app will be sold to third parties for marketing purposes *(Options: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)*.

Study Scenarios

Scenario 1

App: When a new sleep pattern is logged by your smartwatch, this application will add an event in your cloud calendar containing your sleep information.

Risk: If an adversary intercepts the communication, there is a risk that sleep routine details may be leaked and used to infer stress levels, fatigue, or even suggest sleeping pills.

Note: Along with the app description, we presented the image 2 to help participants understand the IoT TAP context.

The 11 questions of the questionnaire are presented here.

Scenario 2

App: If your location is outside your home, this application will lock the door.

Risk: If an adversary collects your location information, they could predict when you are not at home.

Note: Along with the app description, we presented the image 3 to help participants understand the IoT TAP context.

The 11 questions of the questionnaire are presented here.

Scenario 3

App: When a new photo is added to your smartphone's photo gallery, this application will upload the photo to your designated cloud storage.

Risk: If an adversary gains unauthorized access to the cloud storage, they could extract sensitive information from the uploaded photos. This could include places you visited, timestamps, and people with you, leading to a privacy breach and unintentional exposure of personal details.

Note: Along with the app description, we presented the image 4 to help participants understand the IoT TAP context.

The 11 questions of the questionnaire are presented here.

Scenario 4

App: When you like a video on a media platform (e.g., video, photo, news), this application will post the content to your online social network account.

Risk: Posting content to an online social network may allow it to target you with advertisements based on your personal preferences from the media platform.

Note: Along with the app description, we presented the image 5 to help participants understand the IoT TAP context.

The 11 questions of the questionnaire are presented here.

Appendix C Study part II: factorial vignette study

Factorial Vignette Study

Features and Levels: This part of the study used the features and related sub-levels listed in Table 3.

Main Question:
”Would you accept the following for running an IoT application?”

Formula for each of the 36 combinations (4×3×4):
Your {*data category*} *is shared for* {*purpose of data sharing*} *and with* {*data recipient type*}. (Options: Yes/No)

Example:
Your location data is shared for the main app functionality and with third parties.

All 36 factorial vignette study applications were presented in random order.

Appendix D Template to Assign a User to a Privacy Clusters

App Evaluation

App Description: If This Then That

Potential Privacy Risk: to be manually derived

Questions (5-point Likert Scale):

- I feel that as a result of using this app, my personal information may not remain confidential.
- I am concerned that this app may sell my information to other companies or institutions without my permission.
- I am concerned that I may lose control over my data by using this app.
- It’s important for me to know what information the service companies involved in this app store about me in their database.
- For this app, I want to receive a summary about the data processing that occurs (e.g., how my data are manipulated to produce meaningful information).
- It’s important for me to know if the data from this app will be sold to third parties for marketing purposes.

Appendix E Methods

Table D1 Study part 1: Original statements from literature that inspired the questionnaire items.

Dimension	id	Original Item or Focus Group quotation
Confidentiality	u1	I believe my personal information provided to these Web sites remains confidential (Tamara Dinev and Hart, 2013).
	u2	I am afraid that this data will be used for commercial reasons or that the wrong person will get access to my data. Moreover, I am afraid that insurance companies will use the data to personalize insurances (Maus et al., 2021).
	u3	I believe my personal information is accessible only to those authorized to have access (Tamara Dinev and Hart, 2013).
	u4	Thinking about the possible data leakage, has your desire to keeping using any of these applets changed? (Cobb et al., 2020).
Control	c1	Would you be upset if the applet contributed to the following situations occurring: private information gets posted online unintentionally, possibly embarrassing you (Cobb et al., 2020).
	c2	Thinking about the possible loss of control, has your desire to keeping using any of these applets changed? (Cobb et al., 2020).
	c3	Importance of whether or not the site posts a privacy policy (Awad and Krishnan, 2006).
	c4	I would definitely not like it to upload immediately. It would be very scary if it just uploaded without you like pressing upload (Romare et al., 2023).
Transparency	t1	Importance of whether a company will allow me to find out what information about me they keep in their databases (Awad and Krishnan, 2006).
	t2	Importance of whether a site is going to use the information they collect from me in a way that will identify me (Awad and Krishnan, 2006).
	t3	Some Web sites use special identification numbers not only to personalize site content, but also to personalize advertising that appears on the site and make sure that visitors are not repeatedly shown the same advertisements. If a site that you frequented asked you whether it could assign you an identification number so that it could provide you with personalized advertising, would you agree? (Awad and Krishnan, 2006).

References

Aghvamipanah, M., Amini, M., Artho, C., Balliu, M.: Activity recognition protection for iot trigger-action platforms. In: 2024 IEEE 9th European Symposium on Security and Privacy (EuroS&P), pp. 600–616 (2024). IEEE

Abdalla, H.I.: A brief comparison of k-means and agglomerative hierarchical clustering algorithms on small datasets. In: Qian, Z., Jabbar, M.A., Li, X. (eds.) Proceeding of 2021 International Conference on Wireless Communications, Networking and Applications, pp. 623–632. Springer, Singapore (2022)

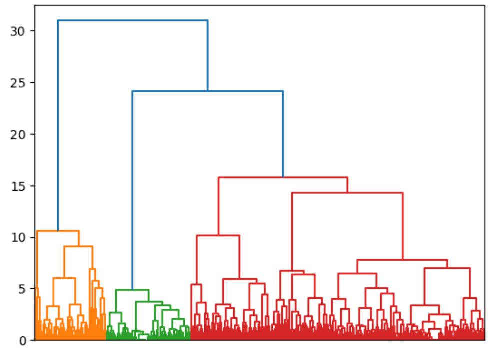


Fig. E1 Dendrogram

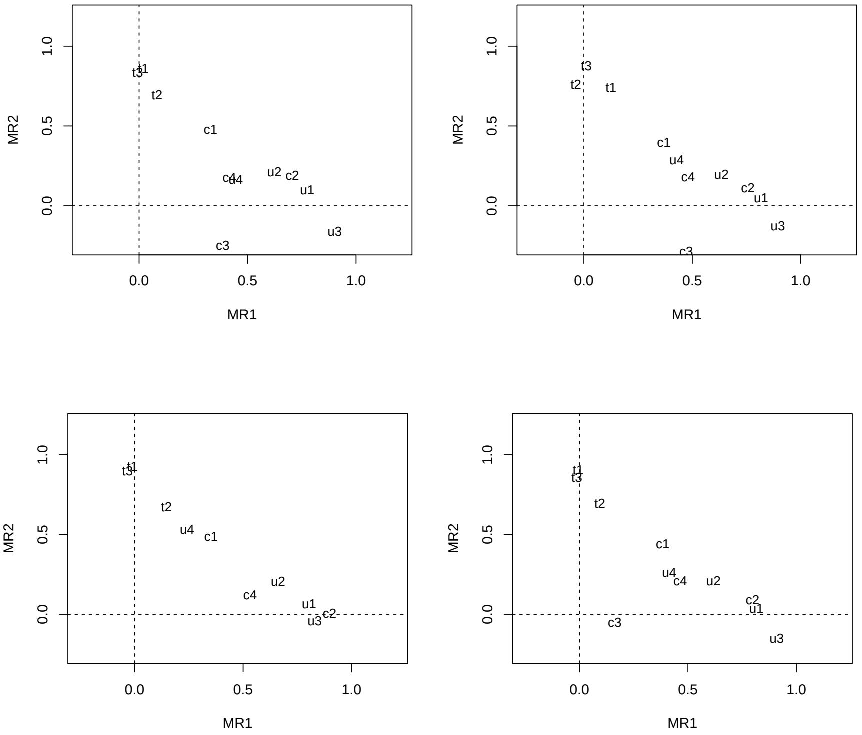


Fig. E2 Exploratory Factor Analysis per Scenario, Scenario #1 (top left), Scenario #2 (top right), Scenario #3 (bottom left) and Scenario #4 (bottom right). The items that can be considered and tested in the MGCFA should load with a value > 0.32 and with communalities $h^2 \geq 0.40$, see Table 5.

Acquisti, A., Brandimarte, L., Loewenstein, G.: Privacy and human behavior in the age of information. Science **347**(6221), 509–514 (2015) https://doi.org/10.1126/science.aaa1465 https://www.science.org/doi/pdf/10.1126/science.aaa1465

Alsoubai, A., Ghaiumy Anaraky, R., Li, Y., Page, X., Knijnenburg, B., Wisniewski, P.J.: 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. CHI '22. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3491102.3517652

Ahmadpanah, M.M., Hedin, D., Sabelfeld, A.: Lazytap: On-demand data minimization for trigger-action applications. In: 2023 IEEE Symposium on Security and Privacy (SP), pp. 3079–3097. IEEE, New York, NY, USA (2023)

Awad, N.F., Krishnan, M.S.: The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Quarterly **30**(1), 13–28 (2006). Accessed 2024-03-20

Alqhatani, A., Lipford, H.R.: Look before you leap! perceptions and attitudes towards inferences in wearable fitness trackers. In: Moallem, A. (ed.) HCI for Cybersecurity, Privacy and Trust, pp. 399–418. Springer, Cham (2023)

Apthorpe, N., Shvartzshnaider, Y., Mathur, A., Reisman, D., Feamster, N.: Discovering smart home internet of things privacy norms using contextual integrity **2**(2) (2018) https://doi.org/10.1145/3214262

Brandão, A., Mendes, R., Vilela, J.a.P.: Prediction of mobile app privacy preferences with user profiles via federated learning. In: CODASPY '22, pp. 89–100. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3508398.3511526

Balliu, M., Bastys, I., Sabelfeld, A.: Securing iot apps. IEEE Security & Privacy **17**(5), 22–29 (2019)

Baruh, L., Cemalcılar, Z.: It is more than personal: Development and validation of a multidimensional privacy orientation scale. Personality and Individual Differences **70**, 165–170 (2014)

Breve, B., Cimino, G., Deufemia, V.: Identifying security and privacy violation rules in trigger-action iot platforms with nlp models. IEEE Internet of Things Journal **10**(6), 5607–5622 (2022)

Barth, S., de Jong, M.D.T.: The privacy paradox – investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review. Telematics and Informatics **34**(7), 1038–1058 (2017) https://doi.org/10.1016/j.tele.2017.04.013

Bahirat, P., He, Y., Menon, A., Knijnenburg, B.: A data-driven approach to developing iot privacy-setting interfaces. In: Proceedings of the 23rd International Conference on Intelligent User Interfaces. IUI '18, pp. 165–176. Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3172944.3172982

Butori, R., Miltgen, C.L.: A construal level theory approach to privacy protection: The conjoint impact of benefits and risks of information disclosure. Journal of Business Research **168**, 114205 (2023)

Biselli, T., Steinbrink, E., Herbert, F., Schmidbauer-Wolf, G.M., Reuter, C.: On the challenges of developing a concise questionnaire to identify privacy personas. Proceedings on Privacy Enhancing Technologies **2022**, 645–669 (2022) https://doi.org/10.56553/popets-2022-0126

Bartol, J., Vehovar, V., Petrovčič, A.: Should we be concerned about how information privacy concerns are measured in online contexts? a systematic review of survey scale development studies. Informatics **8**(2) (2021) https://doi.org/10.3390/informatics8020031

Byrne, B.M.: Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. Journal of Personality Assessment **85**(1), 17–32 (2005) https://doi.org/10.1207/s15327752jpa8501_02 . PMID: 16083381

Bansal, G., Zahedi, F.M., Gefen, D.: Do context and personality matter? trust and privacy concerns in disclosing private information online. Information & Management **53**(1), 1–21 (2016) https://doi.org/10.1016/j.im.2015.08.001

Colnago, J., Feng, Y., Palanivel, T., Pearman, S., Ung, M., Acquisti, A., Cranor, L.F., Sadeh, N.: Informing the design of a personalized privacy assistant for the internet of things. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20, pp. 1–13. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376389

Conway, J.M., Huffcutt, A.I.: A review and evaluation of exploratory factor analysis practices in organizational research. Organizational Research Methods **6**(2), 147–168 (2003) https://doi.org/10.1177/1094428103251541

Chen, H., Keith, T.Z., Weiss, L., Zhu, J., Li, Y.: Testing for multigroup invariance of second-order wisc-iv structure across china, hong kong, macau, and taiwan. Personality and Individual Differences **49**(7), 677–682 (2010) https://doi.org/10.1016/j.paid.2010.06.004

Costello, A., Osborne, J.: Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment, Research & Evaluation **10**, 1–9 (2005)

Chawdhry, A., Paullet, K., Pinchot, J.: Internet of things: Measuring data privacy concerns of users. Issues In Information Systems **23**, 30–45 (2022) https://doi.org/10.48009/4_iis_2022_109

Chignell, M.H., Quan-Haase, A., Gwizdka, J.: The privacy attitudes questionnaire (paq): initial development and validation. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 47, pp. 1326–1330 (2003). SAGE Publications Sage CA: Los Angeles, CA

Castañeda, A., Ríos, F., Luque Martínez, T.: The dimensionality of customer privacy concern on the internet. Online Information Review **31**, 420–439 (2007) https://doi.org/10.1108/14684520710780395

Cobb, C., Surbatovich, M., Kawakami, A., Sharif, M., Bauer, L., Das, A., Jia, L.: How risky are real users' ifttt applets? In: Proceedings of the Sixteenth USENIX Conference on Usable Privacy and Security. SOUPS'20. USENIX Association, USA (2020)

Dupree, J.L., Devries, R., Berry, D.M., Lank, E.: Privacy personas: Clustering users via attitudes and behaviors toward security practices. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. CHI '16, pp. 5228–5239. Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858214

Deng, X., Doll, W.J., Hendrickson, A.R., Scazzero, J.A.: A multi-group analysis of structural invariance: an illustration using the technology acceptance model. Information & Management **42**(5), 745–759 (2005) https://doi.org/10.1016/j.im.2004.08.001

Das, A., Degeling, M., Smullen, D., Sadeh, N.: Personalized privacy assistants for the internet of things: Providing users with notice and choice. IEEE Pervasive Computing **17**(3), 35–46 (2018) https://doi.org/10.1109/MPRV.2018.03367733

Dinev, T., Hart, P.: Internet privacy concerns and social awareness as determinants of intention to transact. International Journal of Electronic Commerce **10**(2), 7–29 (2005). Accessed 2024-04-14

Dinev, T., Hart, P.: An extended privacy calculus model for e-commerce transactions. Information Systems Research **17**(1), 61–80 (2006) https://doi.org/10.1287/isre.1060.0080

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

Emami-Naeini, P.: Informing privacy and security decision making in an iot world. PhD thesis, Carnegie Mellon Pittsburgh, PA, USA (2020)

Emami-Naeini, P., Agarwal, Y., Cranor, L.: iotsecurityprivacy.org. https://www.iotsecurityprivacy.org/downloads/Privacy_and_Security_Specifications.pdf. [Accessed 01-03-2024] (2021)

Emami-Naeini, P., Bhagavatula, S., Habib, H., Degeling, M., Bauer, L., Cranor, L.F., Sadeh, N.: Privacy expectations and preferences in an iot world. In: Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security. SOUPS '17, pp. 399–412. USENIX Association, USA (2017)

Faklaris, C., Dabbish, L.A., Hong, J.I.: A Self-Report measure of End-User security attitudes (SA-6). In: Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019), pp. 61–77. USENIX Association, Santa Clara, CA (2019). https://www.usenix.org/conference/soups2019/presentation/faklaris

Federal Constitutional Court of Germany: Decision of 15 December 1983 – Census Verdict. BVerfGE 65, 1 et seq. Docket Nos. 1 BvR 209/83 et al. (1983). https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs19831215_1bvr020983en.html

Foltz, C., Foltz, L.: Mobile users' information privacy concerns instrument and iot. Information & Computer Security **ahead-of-print** (2020) https://doi.org/10.1108/ICS-07-2019-0090

Fischer-Hübner, S., Karegar, F.: Challenges of Usable Privacy, pp. 103–131. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-54158-2_4

Fernandes, E., Jung, J., Prakash, A.: Security analysis of emerging smart home applications. In: 2016 IEEE Symposium on Security and Privacy (SP), pp. 636–654 (2016). IEEE

Feng, Y., Yao, Y., Sadeh, N.: A design space for privacy choices: Towards meaningful privacy control in the internet of things. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21. Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3411764.3445148

Gollust, S.E., Gordon, E.S., Zayac, C., Griffin, G., Christman, M., Pyeritz, R., Wawak, L., Bernhardt, B.A.: Motivations and perceptions of early adopters of personalized genomics: perspectives from research participants. Public health genomics **15**(1), 22–30 (2011)

Groß, T.: Validity and reliability of the scale internet users' information privacy concerns (iuipc). Proceedings on Privacy Enhancing Technologies **2021**, 235–258 (2021) https://doi.org/10.2478/popets-2021-0026

Groß, T.: Toward Valid and Reliable Privacy Concern Scales: The Example of IUIPC-8, pp. 55–81. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-28643-8_4

He, Y., Bahirat, P., Knijnenburg, B.P., Menon, A.: A data-driven approach to designing for privacy in household iot. ACM Trans. Interact. Intell. Syst. **10**(1) (2019) https://doi.org/10.1145/3241378

Hansen, M., Jensen, M., Rost, M.: Protection goals for privacy engineering. In: 2015 IEEE Security and Privacy Workshops, pp. 159–166. IEEE, New York, NY, USA (2015). https://doi.org/10.1109/SPW.2015.13

Hasan, R., Weil, R., Siegel, R., Krombholz, K.: A psychometric scale to measure individuals' value of other people's privacy (vopp). In: Proceedings of the 2023 Chi Conference on Human Factors in Computing Systems, pp. 1–14 (2023)

Iravantchi, Y., Emami-Naeini, P., Sample, A.: Sok:(un) usable privacy: the lack of overlap between privacy-aware sensing and usable privacy research. Proceedings on Privacy Enhancing Technologies (2025)

Islami, L., Kitkowska, A., Fischer-Hübner, S.: Inter-regional lens on the privacy preferences of drivers for its and future vanets. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, pp. 1–20 (2024)

Inverardi, P., Migliarini, P., Palmiero, M.: Systematic review on privacy categorisation. Computer Science Review **49**, 100574 (2023) https://doi.org/10.1016/j.cosrev.2023.100574

Inc., I.: IFTTT Terms of Service. https://ifttt.com/terms. Accessed: 2025-08-15 (2025)

Jr., J.H.W.: Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association **58**(301), 236–244 (1963) https://doi.org/10.1080/01621459.1963.10500845

Kumaraguru, P., Cranor, L.F.: Privacy indexes: a survey of westin's studies. Institute for Software Research International (2005)

Kalantari, S., Hughes, D., De Decker, B.: Listing the ingredients for ifttt recipes. In: 2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), pp. 1376–1383. IEEE, New York, NY, USA (2022). https://doi.org/10.1109/TrustCom56396.2022.00194

Knijnenburg, B.P.: Privacy? i can't even! making a case for user-tailored privacy. IEEE Security & Privacy **15**(4), 62–67 (2017) https://doi.org/10.1109/MSP.2017.3151331

Khodabakhsh, A., Yayilgan, S.Y.: Data Privacy in IoT Equipped Future Smart Homes (2020). https://arxiv.org/abs/2008.04979

Liu, B., Andersen, M.S., Schaub, F., Almuhimedi, H., Zhang, S.A., Sadeh, N., Agarwal, Y., Acquisti, A.: Follow my recommendations: A personalized privacy assistant for mobile app permissions. In: Twelfth Symposium on Usable Privacy and Security (SOUPS 2016), pp. 27–41. USENIX Association, Denver, CO (2016). https://www.usenix.org/conference/soups2016/technical-sessions/presentation/liu

Lee, H., Kobsa, A.: Understanding user privacy in internet of things environments. In: 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), pp. 407–412. IEEE, New York, NY, USA (2016). https://doi.org/10.1109/WF-IoT.2016.7845392

Lee, H., Kobsa, A.: Privacy preference modeling and prediction in a simulated campuswide iot environment. In: 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 276–285. IEEE, New York, NY, USA (2017). https://doi.org/10.1109/PERCOM.2017.7917874

Liu, B., Lin, J., Sadeh, N.: Reconciling mobile app privacy and usability on smartphones: could user privacy profiles help? In: Proceedings of the 23rd International Conference on World Wide Web. WWW '14, pp. 201–212. Association for Computing Machinery, New York, NY, USA (2014). https://doi.org/10.1145/2566486.2568035

Lin, J., Liu, B., Sadeh, N., Hong, J.I.: 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 '14, pp. 199–212. USENIX Association, USA (2014)

Lynn, T., Muzellec, L., Caemmerer, B., Turley, D.: Social network sites: early adopters' personality and influence. Journal of Product & Brand Management **26**(1), 42–51 (2017)

Lafontaine, E., Sabir, A., Das, A.: Understanding people's attitude and concerns towards adopting iot devices. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. CHI EA '21. Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3411763.3451633

Lipford, H.R., Tabassum, M., Bahirat, P., Yao, Y., Knijnenburg, B.P.: Privacy and the Internet of Things, pp. 233–264. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-82786-1_11

Li, L., Zhang, Y.: An extended theory of planned behavior to explain the intention to use carsharing: a multi-group analysis of different sociodemographic characteristics. Transportation **50** (2021) https://doi.org/10.1007/s11116-021-10240-1

Moritz Büchi, N.J., Latzer, M.: Caring is not enough: the importance of internet skills for online privacy protection. Information, Communication & Society **20**(8), 1261–1278 (2017) https://doi.org/10.1080/1369118X.2016.1229001

Meade, A., Craig, B.: Identifying careless responses in survey data. Psychological methods **17**, 437–55 (2012) https://doi.org/10.1037/a0028085

Magrizos, S., Campora, M., Lamprinakos, G., Giovanis, A., Christofi, M.: Transparency by design: the effect of privacy policies visualisation on brand trust and perceived intrusion. Behaviour & Information Technology, 1–13 (2025)

Morgan, P.L., Collins, E.I.M., Spiliotopoulos, T., Greeno, D.J., Jones, D.M.: Reducing risk to security and privacy in the selection of trigger-action rules: Implicit vs. explicit priming for domestic smart devices. International Journal of Human-Computer Studies **168**, 102902 (2022) https://doi.org/10.1016/j.ijhcs.2022.102902

Murmann, P., Fischer-Hübner, S.: Tools for achieving usable ex post transparency: A survey. IEEE Access **5**, 22965–22991 (2017) https://doi.org/10.1109/ACCESS.2017.2765539

Madejski, M., Johnson, M., Bellovin, S.M.: A study of privacy settings errors in an online social network. In: 2012 IEEE International Conference on Pervasive Computing and Communications Workshops, pp. 340–345. IEEE, New York, NY, USA (2012). https://doi.org/10.1109/PerComW.2012.6197507

Malhotra, N.K., Kim, S.S., Agarwal, J.: Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model. Information systems research **15**(4), 336–355 (2004)

Maus, B., Olsson, C.M., Salvi, D.: Privacy personas for IoT-based health research: A privacy calculus approach. Front. Digit. Health **3**, 675754 (2021)

Marky, K., Stöver, A., Prange, S., Bleck, K., Gerber, P., Zimmermann, V., Müller, F., Alt, F., Mühlhäuser, M.: Decide yourself or delegate-user preferences regarding the autonomy of personal privacy assistants in private iot-equipped environments. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, pp. 1–20. Association for Computing Machinery, New York, NY, USA (2024)

Putnick, D.L., Bornstein, M.H.: Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review **41**, 71–90 (2016) https://doi.org/10.1016/j.dr.2016.06.004

Phelps, J., Nowak, G., Ferrell, E.: Privacy concerns and consumer willingness to provide personal information. Journal of Public Policy and Marketing **19**(1), 27–41 (2000) https://doi.org/10.1509/jppm.19.1.27.16941 . Cited by: 757

Preibusch, S.: Guide to measuring privacy concern: Review of survey and observational instruments. International Journal of Human-Computer Studies **71**(12), 1133–1143 (2013) https://doi.org/10.1016/j.ijhcs.2013.09.002

Romare, P., Morel, V., Karegar, F., Fischer-Hübner, S.: Tapping into privacy: A study of user preferences and concerns on trigger-action platforms. In: 2023 20th Annual International Conference on Privacy, Security and Trust (PST), pp. 1–12. IEEE, New York, NY, USA (2023). https://doi.org/10.1109/PST58708.2023.10320180

Romare, P.: User-driven privacy factors in trigger-action apps: A comparative analysis with general iot. In: Bieker, F., Conca, S., Gruschka, N., Jensen, M., Schiering, I. (eds.) Privacy and Identity Management. Sharing in a Digital World, pp. 244–264. Springer, Cham (2024)

Rao, A., Pfeffer, J.: Types of privacy expectations. Frontiers Big Data **3**, 7 (2020) https://doi.org/10.3389/FDATA.2020.00007

Surbatovich, M., Aljuraidan, J., Bauer, L., Das, A., Jia, L.: Some recipes can do more than spoil your appetite: Analyzing the security and privacy risks of ifttt recipes. In: WWW '17, pp. 1501–1510. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE (2017). https://doi.org/10.1145/3038912.3052709

Schaub, F., Balebako, R., Durity, A.L., Cranor, L.F.: A design space for effective privacy notices. In: Eleventh Symposium On Usable Privacy and Security (SOUPS 2015), pp. 1–17. USENIX Association, Ottawa (2015). https://www.usenix.org/conference/soups2015/proceedings/presentation/schaub

Saeidi, M., Calvert, M., Au, A., Sarma, A., Bobba, R.: If this context then that concern : Exploring users' concerns with ifttt applets. Proceedings on Privacy Enhancing Technologies **2022**, 166–186 (2022) https://doi.org/10.2478/popets-2022-0009

Schwartz, B.: The paradox of choice. Positive psychology in practice: Promoting human flourishing in work, health, education, and everyday life, 121–138 (2015)

Salminen, J., Guan, K., Jung, S.-G., Chowdhury, S.A., Jansen, B.J.: A literature review of quantitative persona creation. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20, pp. 1–14. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376502

Salminen, J., Jung, S.-G., Nielsen, L., Jansen, B.: Creating more personas improves representation of demographically diverse populations: Implications towards interactive persona systems. In: Nordic Human-Computer Interaction Conference. NordiCHI '22. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3546155.3546654

Smith, H.J., Milberg, S.J., Burke, S.J.: Information privacy: Measuring individuals' concerns about organizational practices. MIS Quarterly **20**(2), 167–196 (1996). Accessed 2024-08-28

Shahapure, K.R., Nicholas, C.: Cluster quality analysis using silhouette score. In: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), pp. 747–748. IEEE, USA (2020). https://doi.org/10.1109/DSAA49011.2020.00096

Solove, D.J.: Privacy self-management and the consent dilemma. Harvard Law Review **126**, 1880–1903 (2013). GWU Legal Studies Research Paper No. 2012-141, GWU Law School Public Law Research Paper No. 2012-141

Saritepeci, M., Yildiz- Durak, H., Özüdoğru, G., Atman Uslu, N.: The role of digital literacy and digital data security awareness in online privacy concerns: a multi-group analysis with gender. Online Information Review **48** (2024) https://doi.org/10.1108/OIR-03-2023-0122

Tamara Dinev, J.H.S. Heng Xu, Hart, P.: Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts. European Journal of Information Systems **22**(3), 295–316 (2013) https://doi.org/10.1057/ejis.2012.23

Tabachnick, B.G., Fidell, L.S., Ullman, J.B.: Using Multivariate Statistics vol. 6. pearson Boston, MA, USA (2013)

Thomas G. Reio, J., Shuck, B.: Exploratory factor analysis: Implications for theory, research, and practice. Advances in Developing Human Resources **17**(1), 12–25 (2015) https://doi.org/10.1177/1523422314559804

Ur, B., Pak Yong Ho, M., Brawner, S., Lee, J., Mennicken, S., Picard, N., Schulze, D., Littman, M.L.: Trigger-action programming in the wild: An analysis of 200,000 ifttt recipes. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. CHI '16, pp. 3227–3231. Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858556

Westin, A., *et al.*: Bibliography of surveys of the us public, 1970-2003. Accessed on August **11**, 2016 (2003)

Westin, A.F.: Privacy and freedom. Washington and Lee Law Review **25**, 166 (1968)

Wang, Z., Luo, B., Li, F.: Privacyguard: Exploring hidden cross-app privacy leakage threats in iot apps. Proceedings on Privacy Enhancing Technologies (2025)

Xu, H., Dinev, T., Smith, H., Hart, P.: Information privacy concerns: Linking individual perceptions with institutional privacy assurances. J. AIS **12** (2011) https://doi.org/10.17705/1jais.00281

Xu, H., Gupta, S., Rosson, M., Carroll, J.: Measuring mobile users' concerns for information privacy. In: International Conference on Information Systems, ICIS 2012. International Conference on Information Systems, ICIS 2012, pp. 2278–2293. AIS, Atlanta USA (2012). International Conference on Information Systems, ICIS 2012 ; Conference date: 16-12-2012 Through 19-12-2012

Xu, M., Jug, Z., Tamò-Larrieux, A.: A cross-cultural analysis of transparency: the interplay of law, privacy policies, and user perceptions. International Data Privacy Law, 011 (2024) https://doi.org/10.1093/idpl/ipae011

Xu, R., Zeng, Q., Zhu, L., Chi, H., Du, X., Guizani, M.: Privacy leakage in smart homes and its mitigation: Ifttt as a case study. IEEE Access **7**, 63457–63471 (2019)

Zheng, S., Apthorpe, N., Chetty, M., Feamster, N.: User perceptions of smart home iot privacy. Proc. ACM Hum.-Comput. Interact. **2**(CSCW) (2018) https://doi.org/10.1145/3274469

Zimmermann, C.: A categorization of transparency-enhancing technologies. CoRR **abs/1507.04914** (2015) 1507.04914

Ziegeldorf, J., Morchon, O., Wehrle, K.: Privacy in the internet of things: Threats and challenges. Security and Communication Networks **7** (2014) https://doi.org/10.1002/sec.795