Energy Reports 11 (2024) 2011–2022

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

How end users perceive their energy data within the spectrum of personal information: A two-stage clustering approach

Christian Pfeiffer [a,*], Stefanie Hatzl [b], Eva Fleiß [c], Alfred Posch [c]

[a] Forschung Burgenland GmbH, Eisenstadt, Austria
[b] Department of Information Technologies & Business Informatics, CAMPUS 02, University of Applied Sciences, Graz, Austria
[c] Department of Environmental Systems Sciences, University of Graz, Graz, Austria

ARTICLE INFO

Keywords:
Energy data
Smart meter
Willingness to share
Information sensitivity
Risk perception
Privacy concern
Cluster analysis

ABSTRACT

Flexibility in the energy system, which is needed due to the increasing proportion of decentralized and fluctuating renewable energy sources, remains limited until end users adopt digital energy services. These services require both energy data and further personal information for their operation, a setting that raises privacy concerns among end users. This paper addresses this issue by investigating the perceptions of 604 individuals of different information types regarding risk, sensitivity and willingness to share. A two-stage hierarchical clustering approach highlights the particular role of energy data within this spectrum of personal information and profiles different end user segments, aiming to provide crucial insights into existing privacy concerns of end users. The results show that smart meter data is perceived as more sensitive than other energy data, reaching a level comparable to that of social interaction data. In this regard, it can be expected that at most about 20% of the end users are willing to share their energy data. Demand side actions need to increasingly focus on fluctuating electricity production as well as demand side energy consumption for heating or individual devices. These insights are essential for designing and using digital energy services in a manner that minimizes privacy concerns and enhances the availability of household energy data and thus the provision of flexibility.

1. Introduction

As the overall electricity demand and share of – increasingly decentralized – renewable energy simultaneously increases, the balancing of the peaks and troughs of load profiles poses a major challenge for the energy transition (Aloise-Young et al., 2021). A promising approach to addressing this challenge is the opportunity to adjust the electricity infeed and purchases, i.e., providing flexibility (Plaum et al., 2022). Households can provide flexibility by shifting their loads to different times (e.g., use of washing machine at times of high photovoltaic (PV) generation) or by feeding power back into the grid (Hall and Geissler, 2020). Digital energy services (e.g., smart home energy management systems) support these measures by providing users with recommendations on how to optimize their energy consumption based on a user's consumption pattern (Pandiyan et al., 2023). However, these technologies need various personal information about end users' activities for their operation (Cagno et al., 2022; Reyes, 2022).

Thus, the basis for end user flexibility is obtaining comprehensive and accurate household energy data (Chen et al., 2023). These data can be viewed as records of the occupants' use of energy-using appliances and equipment in their home as time series, largely through everyday activities such as cleaning, food preparation, leisure and keeping warm (or cool), and are regularly used to calculate energy system models (e.g., Askeland et al., 2021). Combined with sociodemographic information (McKenna et al., 2022), consumption patterns are utilized to assess the fluctuations in energy demand over a defined time frame. Hence, energy data and other information are crucial in characterizing energy demand behavior and thus the underlying information for generating flexibility (Hussain et al., 2023).

However, the availability of household energy data remains limited until end users adopt digital energy services (Paudler et al., 2023). Consequently, smart grids remain incomplete without the adoption of smart technologies in households (Fernandes and Silva, 2022). For example, ample evidence indicates a slow progress in the deployment of

* Correspondence to: Forschung Burgenland GmbH, Campus 1, 7000 Eisenstadt, Austria.
E-mail address: christian.pfeiffer@forschung-burgenland.at (C. Pfeiffer).

https://doi.org/10.1016/j.egyr.2024.01.053
Received 12 October 2023; Received in revised form 30 November 2023; Accepted 22 January 2024
Available online 2 February 2024

advanced electricity metering infrastructure (i.e., smart meters) (Vitiello et al., 2022). Other studies underscore that privacy concerns among end users impede the collection of digital energy data through smart meters, hindering its sharing in smart grids (Ryu and Kim, 2024). These privacy concerns are prevalent not only with smart metering technologies but also extend to all digital energy services (Li et al., 2020).

The privacy debate is a recurring theme in research within the context of digital media (Ayaburi and Treku, 2020), highlighting that the primary reason for not adopting social media or other digital services is privacy concerns that arise from data sharing (Zhang and Liu, 2022). In this context, studies on data perception contribute to a deeper understanding of these concerns. For instance, researchers have explored the willingness to share personal information along with the perceived risk or the sensitivity of different information types in the marketing context (Milne et al., 2017; Belen-Saglam et al., 2022). However, empirical findings that specifically address the context of energy data are scarce. Sporadic studies only touch on single aspects, such as either sensitivity or willingness to share, rather than provide a comprehensive analysis (e.g., Deloitte, 2017; Ofgem, 2018; Pfeiffer et al., 2020).

Understanding how energy data, along with other sociodemographic information, are perceived across different user groups provides crucial insights into both the existing privacy concerns of end users and the privacy context of energy data compared to more sensitive information. This understanding is essential for designing and using digital energy services in a manner that minimizes privacy concerns and enhances privacy protection. Subsequently, appropriate energy services enhance the availability of household energy data and thus the provision of flexibility (Vigurs et al., 2021).

Thus, this work addresses this research gap by conducting a thorough analysis of end user perception of energy data. The primary contribution is the exploration of how end users perceive their energy data regarding risks, sensitivity and their willingness to share these data. Specifically, this study has two sub-goals: firstly, to compare the perception of energy data with other types of information, and secondly, to profile different end user segments based on their perception. We therefore surveyed 604 end users and employed a two-stage clustering approach. The conceptual background of our analysis is described in the following section.

2. Conceptual background

The perception of one's personal information can be measured by several privacy-related concepts. The perceived information *sensitivity* is associated with different types of risk (e.g., *psychological*, *physical*, *social*, *financial*, *privacy risk*) (Milne et al., 2017; Phelps et al., 2000). Personal information that is perceived as sensitive is seen as a risk when weighing the costs and benefits of a data disclosure and thus decrease the *willingness to share* data (Gouthier et al., 2022). Measuring these concepts (i.e., type of risk, data sensitivity and willingness to share) creates a comprehensive picture of the perception of individual information types. This allows for a better assessment of the perception of energy data within the spectrum of other personal information types. In addition, the individual assessment of each type of data by risk type, data sensitivity and willingness to share, is complemented by general attitudinal concepts such as *internet privacy concerns*, *internet activity* and *risk behavior*, generating a comprehensive picture of different user segments regarding data sharing in the context of energy systems.

The perceived information *sensitivity* can be understood as the degree of harm that is perceived as potentially caused by sharing one's data (Rumbold and Pierscionek, 2018). Milne et al. (2017) identified secure identifiers as most sensitive information types. Markos et al. (2017) examined U.S. and Brazilian consumers' perceived sensitivity for personally identifiable or linkable data with a friend, marketer, trusted marketer, or unknown marketer. They highlighted the fact that fewer sensitive personally identifiable information items (e.g., email) and more sensitive items (e.g., passwords) are not classified as personally identifiable information. Schomakers et al. (2019) and Belen-Saglam et al. (2022) compared their results with findings from Germany and the UK, both indicating that a consensus exists regarding what constitutes sensitivity in these nations. Regarding energy data, the UK government regulator Ofgem (2018) descriptively showed that energy consumption data are perceived to be the least sensitive in comparison to other personal information.

How sensitive personal information is perceived depends on the type of risk attributed to different types of data (Milne et al., 2017). Risk is a multidimensional concept that was operationalized by Jacoby and Kaplan (1972) who used the dimensions financial, physical, social, and psychological risk. As such, *financial risk* is the risk of financial loss due to fraud (Featherman and Pavlou, 2003) or exorbitant additional costs (Jacoby and Kaplan, 1972). Scholars dedicate significant effort to model dynamic price schemes, aiming to mitigate financial risks for end users due to volatile generation (Panda et al., 2023). According to Milne et al. (2017), most financial risks are assigned to the information types of credit card number, financial account number, or vehicle registration number. *Physical risk* is defined as the risk of physical danger and restriction of personal freedom of movement (Salter and Bryden, 2009). Most physical risks can be linked to secure identifiers, such as the information types of the password and social security number (Milne et al., 2017). Revealing insights into energy consumption also increases the risk of burglary (Razavi et al., 2019). *Social risk* is the risk of a threat to one's self-esteem, reputation, and/or perception by others (Featherman and Pavlou, 2003). Social risks are associated with the information types of the law enforcement file and picture of the face (Milne et al., 2017). In turn, incentive systems incorporating the reputation of prosumers foster interactive and standardized participation in peer-to-peer energy markets (Jiang et al., 2023). *Psychological risk* is defined by negative emotions such as fear, sorrow, and/or conflicts with the self-image (Featherman and Pavlou, 2003) or self-concept (Jacoby and Kaplan, 1972). In particular, Milne et al. (2017) showed that the information types of the DNA profile and medical history are frequently linked to psychological risk. With the increased use of personal information in the digital context, personal information is primarily attributed as a *privacy risk*, that is the risk of a potential loss of control, sale, or disclosure of information to third parties without consent, and the use of information for unintended purposes (Featherman and Pavlou, 2003). The same is true for smart home energy management systems. People's main barrier to adopting these energy services is their ability to reveal information about the home life (Zhang and Liu, 2022).

Data and information are currently being systematically shared as technology becomes increasingly integrated into everyday actions, which suggests that there is a *willingness to share* data. The higher the information sensitivity, the lower the willingness to share (Milne et al., 2017; Pal et al., 2020). Therefore, the willingness to share data varies dramatically depending on the information type (Phelps et al., 2000). Some merely descriptive studies addressed users' willingness to share particularly energy data. The UK government regulator Ofgem (2018) illustrated that almost half of UK consumers were willing to share their energy consumption data when this sharing was connected to personal benefits (i.e., financial savings, improved energy efficiency). Similarly, Yussof et al. (2021) found a high average willingness to share contextual data on electricity consumption in Malaysia. Pfeiffer et al. (2020) distinguished between three segments of Austrian end users in this context, where about a quarter of the respondents signaled a willingness to share their energy data. However, evidence from these studies is still scarce and a comprehensive analysis is lacking.

Originally, privacy was defined as the ‘right to be alone’ (Warren and Brandeis, 2005); today, it refers to an individual’s desire to control or influence personal information (Watson et al., 2010). Therefore, an *internet privacy concern* relates to a concern about whether third parties behave opportunistically regarding the information submitted online, and particularly by the respondent (Dinev and Hart, 2006). Döbelt et al. (2015) explored end users’ internet privacy concern when adopting energy services. Their results highlight the importance of access control, information transparency and the incorporation of privacy enhancing functionalities such as user authentication. Openly accessible energy consumption data can interfere with people’s sense of autonomy and control (Vigurs et al., 2021). Chen et al. (2017) demonstrated that higher internet privacy concerns lead to lower support for and adoption of smart meters.

Furthermore, the extent of *internet activity* matters for designing energy systems. This was shown by Bjarghov et al. (2021), who identified social challenges related to users’ internet accessibility in local electricity markets. More broadly, Wang and Xu (2021) examined the impact of internet use on CO_2 emissions at a global level. Their results indicate that internet use has a positive effect on carbon emission reduction in most industrialized states.

If individuals perceive that the potential benefits outweigh the potential risks, or if they do not consider the risk significant, they are more likely to engage in risky behavior (Foster et al., 2009). The particular form of *risk behavior* differs based on the impact of privacy concerns (Jung and Park, 2018). To overcome this barrier, risk-averse end users need greater financial incentives to reduce electricity consumption (Niromandfam et al., 2020). The risk behavior affects the quality of individual decisions related to energy consumption, energy efficiency adoption and the budgeting ability in the context of energy poverty (DellaValle, 2019). Particularly the incidence of energy poverty is determined by an individual’s risk behavior (Churchill and Smyth, 2021).

3. Methods

To clarify the particular role that end users assign to energy data within the spectrum of personal information, we conducted a quantitative online survey to investigate end users’ perception of different types of data.

3.1. Study sample

We conducted the survey in July 2021 with support from the market research institute respondi in Austria. A focus was directed toward this case region for two reasons: First, like all member states of the European Union, Austria is being asked to ensure the deployment of smart metering systems that encourage the active participation of end users in the electricity market (European Union, 2019). In 2021, electric utilities in Austria had installed 3,032,222 smart meters in total, representing 47% of all metering points in Austria. However, about 3% of the Austrian end users disabled their smart meter from making granular readings of their electricity consumption data (i.e., they “opted out”). In contrast, more than 5% opted to share their quarter-hourly electricity consumption with their grid operator instead of the standard daily transmission (E-Control Austria, 2022). This discrepancy in handling energy data initially motivated us to conduct this research. Second, research in the broader context of this study focused on end users in the U.S. (Milne et al., 2017; Markos et al., 2017), Brazil (Markos et al., 2017), UK (Belen-Saglam et al., 2022) and Germany (Schomakers et al., 2019). Since the attitudes of Austrian end users towards energy policies (e.g., developing renewable energy, fighting global warming, energy affordability) (Janik et al., 2021) and their willingness to share personal information for energy purposes (Reyes, 2022) are widely comparable to those held by people in many European Union countries, the results from Austria provide evidence that can be extrapolated to other countries.

Table 1
Participant sociodemographics.

Sociodemographics	M[a]	(SD[b])	%
Gender			
Female			50.91
Male			49.09
Age (years)	43.52	(14.01)	
Education			
Apprenticeship			49.91
Intermediate vocational school			15.08
School-leaving certificate			20.27
University degree			14.74
Income			
≤€2200			37.07
€2201 - €3000			22.20
€3001 - €4000			21.38
≥€4001			19.35

$n = 604$, [a] means, [b] standard deviations.

To be included in the survey, the participants had to be (i) aged at least 18 and at most 70 years and (ii) have their principal residence in Austria. No prior exclusion criteria were defined, since the goal was to obtain a general picture of the broader public’s data perception. We set a quota by gender, age and education to obtain a representative sample. Informed consent was obtained from participants at the beginning of the survey.

Initially, we collected responses from 741 Austrian end users. Of these, 81 failed an attention check and, therefore, were excluded from the sample. In the course of data cleansing, we further identified and excluded 56 participants who provided inconsistent responses. Thus, the final sample comprises $n = 604$ participants. Detailed sociodemographic data for the participants are stated in Table 1.

3.2. Survey instrument and measures

The survey instrument comprised three parts. First, we collected participant sociodemographic data (i.e., gender, age, education and income) and information about technical devices at their home.

Second, participants were asked to evaluate different information types with respect to risk, sensitivity and willingness to share. We selected the information types based on the extensive set used by Milne et al. (2017). Additionally, we considered particular information types used by Markos et al. (2017) and Schomakers et al. (2019) as relevant. We finally selected 34 information types from these sources and added the “COVID vaccination status”, as we considered this to be an interesting information type at the time of the survey during the COVID-19 pandemic. Most importantly, four types of energy data supplemented the present study. We selected these data types (i.e., PV electricity production, heating energy consumption, washing machine electricity consumption, smart meter data) based on a literature review (Fell, 2017; Marinakis et al., 2020; Pití et al., 2017) and a workshop held with academic experts in the field of energy research. The presented set of information types, therefore, comprises 39 information types that include the supplemented energy data. Table A.3 provides an overview of the data types used for this study. Participants stated their general willingness to share their personal information on a 10-point scale ranging from 1 (“no willingness to share”) to 10 (“high willingness to share”). Similarly, they rated the information sensitivity on a 10-point scale from 1 (“not sensitive”) to 10 (“very sensitive”). Furthermore, we included a “yes/no” response, asking them to indicate whether or not they associated each of five categories of risks with the 39 information types. The risk categories were physical risk, financial risk, social risk, psychological risk and privacy risk, as adopted from Milne et al. (2017) and Featherman and Pavlou (2003). We included a description of each risk type in the survey to ensure a common understanding among the respondents. In addition, we offered the option “no risk”. For each information type, participants could assign more than one risk category. They were also provided with open response fields to comment on their statements.

Third, questions on internet privacy concern (IPC), risk behavior (RB), and internet activity (IA) were asked at the end of the survey. We measured *internet privacy concern* by using five items from Dinev and Hart (2006) and Lee et al. (2019) on a five-point scale ranging from 1 (”do not agree at all”) to 5 (“fully agree”). The *risk behavior* scale comprised four items measured on a scale ranging from 1 (”do not agree at all”) to 5 (“fully agree”). For this, we adopted one item from Khalil and Karam (2015) and adapted three items from Schäwel et al. (2021). The items to measure *internet activity* were based on those of Rosen et al. (2013) and supplemented with items from official statistics (Djahangiri et al., 2018). Eleven items comprised the scale, measured on a frequency scale from 1 (“never”) to 5 (“regularly”). The corresponding items, the descriptive statistics of the scales and their correlations are summarized in Table A.4 in the appendix.

3.3. Data analysis

We used a two-stage data analysis method to highlight the particular role of energy data within the spectrum of personal information.

First, we created information type clusters to interpret and compare the energy data types in terms of the risk assessments, sensitivity and willingness to share with the other information types. The participants' 39 information type assignments to each of the five risk categories built an aggregated data set for this first analysis stage. This procedure provides percentages of assignments for each information type assigned to each risk category. To simplify the comparisons, we transformed the 10-point scales of willingness to share and perceived information sensitivity to a scale of 0–1 that can be interpreted as a percentage with the higher values indicating a higher willingness to share and sensitivity, respectively. We calculated descriptive statistics with mean (M) assignments of each risk category, their standard deviations (SD) and bivariate correlations. A principal components analysis with the measures financial risk, physical risk, social risk, psychological risk, privacy risk, no risk, sensitivity and willingness to share reduced the dimension of the aggregated data set. The resulting orthogonal components served as input for a subsequent cluster analysis (Ringnér, 2008) conducted to create information type clusters. Because there was no explicit need to interpret the cluster analysis input, we did not apply any rotation to the extracted principal components. Using these principal components, we performed a hierarchical cluster analysis with an average linkage based on the Chebyshev distance measure. The distance between values x and y of object j is defined in Eq. (1):

$$d(x_j, y_j) = \lim_{p\to\infty}\left(\sum_{j=1}^{m}|x_j - y_j|^p\right)^{\frac{1}{p}} = max_j(|x_j - y_j|) \quad (1)$$

We chose the final number of clusters from silhouette values averaged across each object's silhouette width $s(j)$, a measure that is stated in Eq. (2):

$$s(j) = \frac{b(j) - a(j)}{max\{a(j), b(j)\}} \quad (2)$$

In Eq. (2), a(j) is the average dissimilarity of object j to all other objects in cluster A, and $b(j)$ is the minimum dissimilarity of object j to all other objects of other clusters $\neq A$ (i.e., cluster A's nearest neighbor cluster B). We validated the cluster solution using a multivariate analysis of variance (MANOVA).

Second, we deepened the results with a user segmentation. As in the first stage, we performed a principal components analysis and a subsequent hierarchical cluster analysis. Individuals' overall perceived risk, sensitivity and willingness to share for each information type cluster were used as inputs, along with their IPC, RB and IA. Since Jacoby and Kaplan (1972) defined the overall perceived risk (OPR) as a function of specific risk categories, we used the risk assignments $x_{i,j,k}$ ($\in \{0,1\}$) for subject i regarding information type j and risk category k to calculate a subject i's OPR of information type cluster J_l according to Eq. (3):

$$OPR_{i,l} = C\frac{1}{m_l}\sum_{j=1}^{m_l}\sum_{k=1}^{5} w_{j,k}x_{i,j,k}\forall j \in J_l \quad (3)$$

In Eq. (3), m_l is the number of information types that comprise cluster J_l, k indexes the five risk categories and $w_{j,k}$ is a risk weight of information type j with respect to risk category k. C represents a constant to trim the OPR to a range of [0,1] according to Eq. (4):

$$C = \frac{1}{max\left\{\frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{5} w_{j,k}x_{i,j,k}\right\}} \quad (4)$$

We further averaged the sensitivity and willingness to share across each information type cluster J_l. Furthermore, we computed mean indices for the IPC, RB and IA constructs and scaled them to a range of [0,1] to simplify the comparisons. Their reliability was considered as satisfactory for all constructs using Cronbach's alpha ($0.77 \leq \alpha \leq 0.83$). We performed all analyses with the statistical programming tool R (R Core Team, 2021).

4. Results

The descriptive statistics of mean risk assignment, sensitivity and willingness to share across all information types are summarized in Table 2.

Participants consider most information types a risk to their privacy with “passwords” (M = 0.855), “home address” (M = 0.843) and “GPS location” (M = 0.843) occurring the most. The energy data “electricity consumption of washing machine” (M = 0.219), “generation of PV plant” (M = 0.231), “heating energy consumption” (M = 0.234) were considered the least privacy risk. In contrast, “smart meter data” were assigned a comparatively higher privacy risk (M = 0.492). Nevertheless, energy data is most commonly associated with no risk, such as “electricity consumption of washing machine” (M = 0.710), “generation of PV plant” (M = 0.672), and “heating energy consumption” (M = 0.635), similar to other information types like “gender” (M = 0.620), “sports activities” (M = 0.559) and “place of birth” (M = 0.526). Again, “smart meter data” are evaluated differently (M = 0.389).

There is also a strong positive correlation of sensitivity with privacy risk (r = 0.89). On the one hand, a rather small number of participants consider “electricity consumption of washing machine” (M = 0.221), “heating energy consumption” (M = 0.286) and “generation of PV plant”

Table 2
Descriptive statistics of risk assignment, sensitivity and willingness to share.

Information type	FIN	PHY	SOC	PSY	PRI	NON	SENS	WITS
Place of birth	0.023	0.027	0.124	0.040	0.405	0.526	0.315	0.641
Religion	0.015	0.025	0.265	0.101	0.390	0.476	0.307	0.635
Body weight	0.023	0.255	0.168	0.206	0.412	0.386	0.371	0.542
Email address	0.136	0.038	0.186	0.103	0.688	0.237	0.486	0.539
Birth date	0.068	0.038	0.148	0.066	0.551	0.385	0.423	0.585
ZIP code	0.040	0.038	0.102	0.040	0.492	0.456	0.366	0.617
Number of children	0.042	0.030	0.143	0.060	0.455	0.478	0.359	0.615
Political affiliation	0.040	0.045	0.398	0.153	0.463	0.325	0.446	0.463
Sexual preference	0.025	0.150	0.280	0.221	0.521	0.335	0.400	0.546
Occupation	0.177	0.048	0.250	0.104	0.298	0.509	0.280	0.692
Shopping behavior	0.231	0.030	0.170	0.075	0.402	0.444	0.347	0.597
Sports activities	0.027	0.212	0.133	0.090	0.258	0.559	0.256	0.665
Social network profile	0.071	0.066	0.481	0.280	0.710	0.121	0.538	0.399
Medical history	0.139	0.364	0.360	0.420	0.770	0.086	0.732	0.276
Picture of face	0.078	0.194	0.294	0.279	0.708	0.149	0.669	0.294
Home address	0.129	0.128	0.189	0.126	0.843	0.103	0.685	0.312
GPS location	0.075	0.123	0.180	0.144	0.843	0.093	0.720	0.242
Phone number	0.121	0.045	0.212	0.179	0.818	0.106	0.706	0.285
Income level	0.425	0.023	0.267	0.095	0.628	0.148	0.592	0.364
Law enforcement files	0.139	0.048	0.355	0.231	0.722	0.164	0.704	0.376
Online dating activities	0.046	0.053	0.401	0.247	0.694	0.156	0.598	0.257
Credit score	0.678	0.023	0.239	0.138	0.646	0.065	0.790	0.226
Finger print	0.290	0.216	0.211	0.187	0.839	0.040	0.864	0.153
Vehicle registration number	0.186	0.038	0.100	0.050	0.656	0.229	0.613	0.276
Driver’s licence number	0.201	0.032	0.143	0.090	0.784	0.111	0.793	0.156
Credit card number	0.812	0.027	0.101	0.101	0.624	0.017	0.917	0.118
Financial account number	0.812	0.035	0.121	0.124	0.619	0.018	0.906	0.101
IP address	0.334	0.051	0.204	0.139	0.823	0.063	0.788	0.182
Social security number	0.292	0.126	0.287	0.143	0.794	0.055	0.818	0.187
Passwords	0.643	0.095	0.309	0.275	0.855	0.023	0.961	0.035
DNA profile	0.104	0.340	0.260	0.270	0.829	0.065	0.857	0.143
Digital signature	0.470	0.043	0.221	0.111	0.750	0.071	0.817	0.240
Passport number	0.242	0.063	0.191	0.100	0.819	0.071	0.857	0.140
Gender	0.032	0.106	0.138	0.081	0.249	0.620	0.181	0.829
COVID vaccination status	0.056	0.132	0.292	0.178	0.439	0.428	0.378	0.642
Smart meter data	0.126	0.055	0.110	0.066	0.492	0.389	0.527	0.286
Generation of PV plant	0.107	0.013	0.045	0.020	0.231	0.672	0.296	0.426
Heating energy consumption	0.172	0.005	0.053	0.015	0.234	0.635	0.286	0.555
Electricity consumption of washing machine	0.103	0.010	0.037	0.018	0.219	0.710	0.221	0.591

Mean assignments, FIN = financial risk, PHY = physical risk, SOC = social risk, PSY = psychological risk, PRI = privacy risk, NON = no risk, SENS = sensitivity, WITS = willingness to share.

(M = 0.296) as sensitive. On the other hand, a rather large number of participants consider “smart meter data” (M = 0.526) as sensitive. By contrast, there is a strong negative correlation between the willingness to share and the risk category privacy risk (r = −0.81). Regarding energy data, participants would least commonly share their “smart meter data” (M = 0.286). They were, however, much more willing to share data about “generation of PV plant” (M = 0.426), “heating energy consumption” (M = 0.555) and “electricity consumption of washing machine” (M = 0.591). Psychological risk positively correlates with both physical risk (r = 0.69) and social risk (r = 0.79). Participants are more willing to share data associated with no risk (r = 0.87) and consider such data less sensitive (r = −0.95).

4.1. Information type clusters

A principal components analysis revealed that more than 75% of the total object variance could be explained using the first two principal components. The first one explains 54.81% of the total variance (with a sum of squared loadings of 4.39) comprising financial risk, privacy risk, no risk, sensitivity and willingness to share. The second principal component explains 21.97% of the total variance (with a sum of squared loadings of 1.76) comprising physical risk, social risk and psychological risk. Although the measures can be assigned to the principal components, some fuzziness still remains since the risk categories, the sensitivity and the willingness to share are intercorrelated. The hierarchical cluster analysis suggested a six to nine cluster solution. Comparing the cluster silhouette scores with the model simplicity and interpretability yields to a seven-cluster solution with an average silhouette score of 0.42. The resulting clusters differ significantly in terms of their risk, sensitivity and willingness to share values according to a MANOVA (Pillai’s trace V = 0.60, $F(8, 30) = 5.74$, $p < 0.001$). Fig. 1 shows a graphical representation of these seven clusters in two principal components. The data contained therein indicate the following cluster descriptions: *Demographics*, *Attitude and Status*, *Social Interaction*, *Secure Identifiers*, *Financial Information*, *Residuals* and *Energy Data*. For detailed information on cluster profiles, please see Table A.5 in the appendix.

The *Demographics* cluster contains nine information types (e.g., gender, birth date, occupation, religion), information that is associated with comparatively low amounts of financial (M = 0.073, SD = 0.77) and privacy risk (M = 0.389, SD = 0.104). Hence, participants frequently attested them as having no risk (M = 0.495, SD = 0.069) and not only perceive this segment as insensitive (M = 0.315, SD = 0.071)

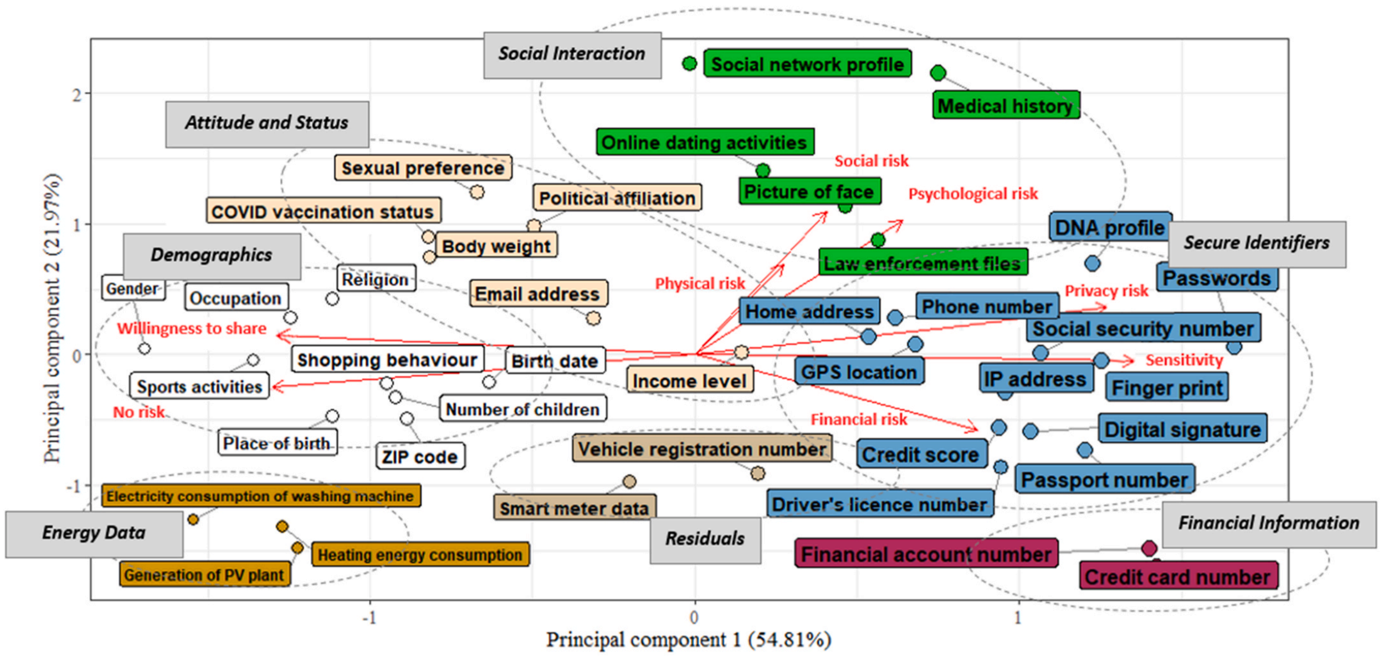

Fig. 1. Cluster solution of information types.

but exhibited the highest willingness to share them (M = 0.653, SD = 0.074).

A cluster with balanced sensitivity (M = 0.445, SD = 0.084) and willingness to share (M = 0.516, SD = 0.094) comprises the participants' *Attitude and Status* regarding six information types (e.g., income level, COVID vaccination status, sexual preference). This information is assigned a comparatively high social risk (M = 0.265, SD = 0.083).

Participants associate the cluster called *Social Interaction* with the comparatively highest physical (M = 0.145, SD = 0.136), social (M = 0.378, SD = 0.069) and psychological (M = 0.291, SD = 0.075) risks. Furthermore, they assign the five corresponding information types (social network profile, medical history, online dating activities, picture of face, law enforcement files) with a high privacy risk (M = 0.721, SD = 0.029), consider them as rather sensitive (M = 0.648, SD = 0.079) and have a rather low willingness to share them (M = 0.320, SD = 0.063).

The *Secure Identifiers* cluster comprises twelve information types (e.g., phone number, DNA profile, passwords, passport number) that yield a high privacy risk (M = 0.804, SD = 0.058) and sensitivity (M = 0.805, SD = 0.077) perception. Secure identifiers are associated with a low willingness to share (M = 0.192, SD = 0.075).

The cluster associated with the highest financial risk (M = 0.812, SD = 0.000) comprises two information types: financial account number and credit card number. This *Financial Information* is perceived as highly sensitive (M = 0.912, SD = 0.008) and a very low willingness to share was observed (M = 0.109, SD = 0.012).

Participants ascribe the *Energy Data* cluster not only a very low financial risk (M = 0.127, SD = 0.039) but also the least physical (M = 0.009, SD = 0.004), social (M = 0.045, SD = 0.008), psychological (M = 0.018, SD = 0.003) and privacy (M = 0.228, SD = 0.008) risks. Furthermore, they consider this cluster, comprising the data types electricity consumption of washing machine, heating energy consumption and generation of PV plant, as the least sensitive (M = 0.268, SD = 0.041) even when compared to *Demographics* (M = 0.315, SD = 0.071). Although *Energy Data* is a no-risk cluster (M = 0.672, SD = 0.037), the willingness to share is moderate (M = 0.514, SD = 0.087) and can be classified as similar to that seen in *Attitude and Status*.

Interestingly, the smart meter data type is not part of the *Energy Data* cluster but builds a unique cluster together with vehicle registration number. Since there is no prior association regarding these data types, the remaining cluster is called *Residuals*. A high privacy risk (M = 0.574, SD = 0.116) is seen in this segment, a result that falls between *Attitude and Status* and *Financial Information*. Regarding sensitivity (M = 0.570, SD = 0.061) and willingness to share (M = 0.281, SD = 0.007), the *Residuals* are somewhat comparable to *Social Interaction* and exhibit rather low willingness to share.

4.2. User segmentation

Individual users perceive the highest OPR with sharing *Financial Information* (M = 0.51, SD = 0.18). They perceive further moderate overall risks regarding *Secure Identifiers* (M = 0.39, SD = 0.12) and *Social Interaction* (M = 0.37, SD = 0.16), while *Energy Data* is associated with lowest overall risk (M = 0.03, SD = 0.05). *Residuals* are perceived as rather not risky (M = 0.18, SD = 0.11). Furthermore, participants are moderately concerned about their internet privacy (M = 0.59, SD = 0.23) and similarly report a moderate risk behavior (M = 0.58, SD = 0.22). These constructs also somewhat correlate positively (r = 0.43). On average, participants use the internet occasionally (M = 0.56, SD = 0.15).

To build user segments, a principal components analysis revealed that about 53% of the total subject variance could be explained using the first four principal components. The first one explains 19.28% of the total variance (with a sum of squared loadings of 4.62) comprising risk, sensitivity and willingness to share for the information type cluster *Demographics*, *Attitude and Status*, *Social Interaction* and *Energy Data*. The second principal component explains 18.04% of the total variance (with a sum of squared loadings of 4.33) comprising either the sensitivity of *Social Interaction*, *Secure Identifiers*, *Financial Information* and *Residuals*, as well as the willingness to share of *Secure Identifiers* and *Residuals*. The third and fourth principal component together explain 15.31% of the total variance (with a sum of squared loadings of 3.68). The third one comprises risk perceptions of further data type clusters, and the fourth one comprises further clusters regarding their willingness to share.

The subsequent hierarchical cluster analysis suggested a six cluster solution with an average silhouette score of 0.22. The resulting user segments differ significantly in their perceptions according to a MANOVA (Pillai's trace V = 0.52, $F(24, 575) = 25.72$, $p < 0.001$). Clusters 5 (n = 6) and 6 (n = 3) were identified as outliers; therefore, they were

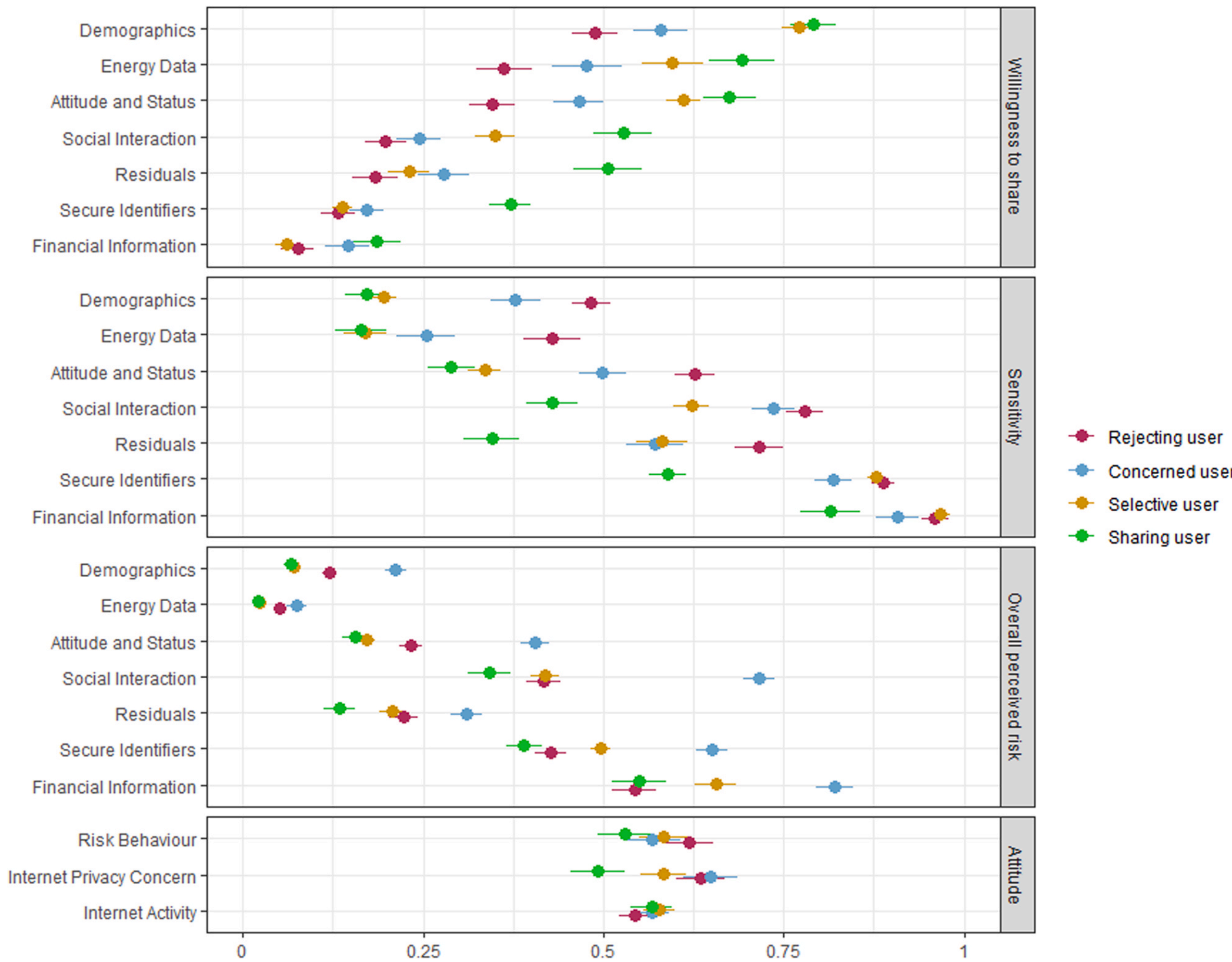

Fig. 2. Cluster solution of user segments.

excluded from further analyses. Fig. 2 shows a graphical representation of the four main user segments. For detailed information about the user profiles, please see Table A.6 in the appendix.

Segment A: "Rejecting user" (29.5%). The prototype of the *rejecting user* is 44.79 years old (SD = 13.83) and predominantly female (56%). Although they do not perceive any information type cluster as particularly risky, their sensitivity perception is the highest among all user segments and thus their willingness to share data is the lowest. Their sensitivity perception of energy data is significantly higher (M = 0.429, SD = 0.273) and willingness to share significantly lower (M = 0.362, SD = 0.266) than those of other segments. Similarly, they perceive the residuals group (which contains smart meter data) as comparatively sensitive (M = 0.716, SD = 0.222) and are thus not willing to share such data or information (M = 0.183, SD = 0.216). As such, the rejecting user has high internet privacy concerns (M = 0.634, SD = 0.231) and is comparatively most careful when using the internet (M = 0.619, SD = 0.221). The level of education corresponds to that of the entire sample. However, lower income levels are slightly overrepresented in this cluster.

Segment B: "Concerned user" (18.7%): The *concerned user* is on average 41.70 years old (SD = 13.27) and characterized by a larger number of males (56%). Compared to other segments, this segment has a relatively high education level, with 25% completing secondary school with a school-leaving certificate and another 27% completing tertiary education. Concerned users have incomes that are either low or high, with few having average incomes. Persons in this segment perceive data sharing as relatively risky, particularly information about social interaction (M = 0.715, SD = 0.120). In general, they report comparatively high information sensitivity and are less willing to share data. Although concerned users practically assign no OPR (M = 0.075, SD = 0.076) and low sensitivity (M = 0.254, SD = 0.218) to energy data, they are only moderately willing to share them (M = 0.477, SD = 0.260). Furthermore, they show a low willingness to share data that corresponds to the residuals group (M = 0.278, SD = 0.195). Although the risk behavior and internet activity of concerned users do not differ from those of other segments, their internet privacy concerns are comparatively the highest (M = 0.649, SD = 0.203).

Segment C: "Selective user" (30.7%): The typical *selective user* is female (58%) with an average age of 42.26 (SD = 13.65), who has completed an apprenticeship (59%) and has a moderate income (27%). Selective users assign moderate OPR to data sharing but consider Secure Identifiers (M = 0.877, SD = 0.073) and Financial Information (M = 0.967, SD = 0.096) as highly sensitive. Therefore, they are unwilling to share these information types. However, persons in this segment assign low sensitivity not only to Demographics (0.195, SD = 0.126) and Attitudes and Status (M = 0.335, SD = 0.150), but also Energy Data (M =

0.169, SD = 0.208). Their willingness to share Energy Data is also comparatively high (M = 0.596, SD = 0.290). The Residuals have an average sensitivity (M = 0.581, SD = 0.242), but low willingness to share (M = 0.230, SD = 0.201). Their IA is highest among all segments (M = 0.577, SD = 0.148), with a considerable number of users possessing smart household devices (14%) and energy management systems (16%). In this largest user segment, 7% own a PV system.

Segment D: "Sharing user" (19.7%): The *sharing user* is typically male (62%), with an average age of 44.88 (SD = 14.82) and tends to have completed a secondary school with a school-leaving certificate (24%). Persons in this segment assign both the lowest OPR and sensitivities to each information type cluster. Furthermore, their willingness to share is highest overall. Although they perceive the sensitivity of Energy Data in ways similar to the selective user (M = 0.164, SD = 0.197), it is striking that the sharing user perceives significantly less sensitivity regarding the Residuals, a cluster that contains smart meter data (M = 0.345, SD = 0.212). Hence, their willingness to share Residuals is highest (M = 0.506, SD = 0.259). A comparatively large proportion of users have smart household devices (16%), while only 7% own a PV system. The IPC is also lowest in this segment (M = 0.492, SD = 0.209).

5. Discussion

In this study, we provided insight into how end users perceive their energy data as compared to other types of information regarding risks, sensitivity and the willingness to share these data for designing and using digital energy services in a manner that minimizes privacy concerns and enhances the availability of household energy data and thus the provision of flexibility. There is compelling evidence that end users perceive energy data as the least sensitive when compared to other information. However, their willingness to share these data remains only moderate. Smart meter data occupy a unique role, as the willingness to share these data is comparatively low. Our general results agree with the findings of Milne et al. (2017) regarding secure identifiers as most sensitive information types. Furthermore, we found that end users are most likely willing to share their own demographic information, a result that supports the previous findings of Milne et al. (2017), Markos et al. (2017) and Schomakers et al. (2019), who obtained similar results in studies performed in the countries of Brazil, the U.S. and Germany.

Previous studies show that a risk of burglary exists, along with information leakage about working, dining and vacationing habits of households, when the meter readings are compromised (Razavi et al., 2019). Thus, third-party insights into a household's energy consumption can even lead to financial and physical risks. Nevertheless, most energy data are perceived as the least sensitive and assigned as having practically no risk compared to other personal information. Even end users that are generally concerned about risks associated with data sharing do not assign any relevant risk to energy data. Still, the willingness to share remains questionable, as this willingness is less pronounced than that for information types in the demographics cluster, but higher than for the other data clusters. Thus, energy service providers should capitalize on the willingness of end users to share their demographic information, leveraging potential synergies for energy data collection when developing demand response solutions. In respect to energy data, the highest actionable potential in utilizing data is related to PV electricity production, heating energy consumption and washing machine electricity consumption. Hence, demand side actions should prioritize strategic load growth based on data that provide information about high PV production, potential for thermal load shifting and individual device usage (Hussain et al., 2023), without relying on electricity consumption data.

A key finding of our study regarding smart meter data differs from recent findings about the perception of energy data. While Ofgem (2018) reported that electricity consumption data are perceived as being the least sensitive when compared with other personal information, our results show a shift of smart meter data within the sensitivity spectrum towards a level that is comparable to social interaction information, such as the social network profile and online dating activities. In contrast to other types of energy data, end users classify smart meter data as belonging to a separate cluster, together with vehicle registration number. Not only are they perceived as more sensitive than PV electricity production, heating energy consumption and washing machine electricity consumption, but they are also less likely to be shared. This is a surprising finding because most privacy concerns regarding energy services arise from their potential to reveal information about daily household activities (Vigurs et al., 2021). However, both heating energy and washing machine electricity consumption have the potential to provide more detailed information about the home life than aggregated smart meter data. While grid operators can – with a time delay – already utilize the latter (Chen et al., 2023), the availability of real-time electricity consumption data will remain limited.

Our results also show that a user segment comprising about 20% of end users are rather willing to share most of their energy data but are only moderately willing to share their smart meter data. Although not completely comparable, this proportion is lower than that reported by previous studies (e.g., Ofgem, 2018; Pfeiffer et al., 2020), emphasizing the current limited availability of household consumption data (Thorve et al., 2023) and thus constraining household flexibility. Despite scholars advocate for energy services that promote the sharing of proprietary data (Paudler et al., 2023), the value of sharing energy data has not yet become widespread among end users (Fernandes and Silva, 2022). Coupled with a lack of awareness of energy transition topics, such as smart grids and demand response, brought about by inadequate communication from decision makers (Geels et al., 2021), end users may have adopted a position of rational distrust towards systematic data collection and use (de Godoy et al., 2021). This distrust may have been heightened by the ongoing smart meter roll-out, a top-down implementation that somewhat restricts the voluntary nature of data sharing (European Union, 2019) as compared to, e.g., the use of social media. Open participant responses reinforce the assumption of distrust, expressing concerns about potential restrictions in the energy supply. Furthermore, participants emphasized the role of privacy concerns and question the added value in sharing energy data. This may also explain why energy data are frequently considered as the least risky and sensitive but are less likely to be shared than other information types. The smart meter controversy is notably present in another user segment, comprising 31%, which exhibits a similar willingness to share energy data but perceives smart meter data as significantly more sensitive than other energy data. Despite this, users in this segment are relatively receptive to smart technologies, a characteristic that holds flexibility potentials through appropriate energy services.

5.1. Implications for research and practice

Our research provides comprehensive analysis results about the perception of energy data, offering a novel perspective on energy system modeling. These models incorporating household data attain validity if end users provide their energy data. However, according to our results, at most about 20% of the end users can be expected to share their data, and at most half could be willing to voluntarily share other energy data, though not necessarily electricity consumption data. Hence, our results provide scholars with benchmarks for achieving greater accuracy in energy system modeling.

Unlike prior studies that have generally focused on information sensitivity and the willingness of end users to share this information (Markos et al., 2017; Rumbold and Pierscionek, 2018; Schomakers et al., 2019; Belen-Saglam et al., 2022), we followed the approach of Milne et al. (2017) who clustered a set of information types according to their risk, sensitivity and willingness to share. However, we contribute to the methodology by providing an even more comprehensive comparison of several information types by considering not only physical risk, financial risk, social risk and psychological risk, but also the privacy risk.

Furthermore, we consider an individual's overall perceived risk as a function of specific risk categories and differentiate among several information type clusters, an approach that improved our ability to perform in-depth end user segmentation.

Our work has several practical implications. Policymakers should make more effort to communicate the benefits of sharing energy data, and particularly smart meter data, for the energy system. Although the energy transition is a complex issue, awareness in the general public is still lacking. Therefore, we need a broader and deeper debate among stakeholders about the necessity of energy data for the energy transition to more effectively highlight the added value and reduce privacy concerns (Radtke, 2022). In this regard, there are initial indications that the latter have recently started to deteriorate, possibly as a result of governmental efforts (Schallehn and Valogianni, 2022). Targeted communication strategies also enable end users to fully consider the potential risks of sharing energy data. Although they do not consider this data to be sensitive, only the privacy risk is considered as relevant with reference to smart meter data. Potential financial and physical risks are also associated with sharing energy data. More generally, this work contributes to a comprehensive understanding of information sensitivity and its privacy aspects. By comparing energy data with other information types, we provide a more finely nuanced view of the end users' perception of data. This ensures that targeted recommendations for action can be made not only in the area of energy transition but also in the whole discussion about data protection.

5.2. Limitations and future research

Two features of this work limit the conclusions we can draw about the perception of energy data. First, we aimed to gain a general perspective of the perception of energy data rather than to focus on specific recipients with whom the data should be shared. Previous studies indicate that the willingness of individuals to share data depends on the particular recipient (Markos et al., 2017). Second, the willingness to share data is linked to personal benefits (e.g., financial savings) (Tschersich et al., 2016), an area that was beyond the explicit scope of this study.

Whether sharing energy data with specific recipients or to obtain incentives would yield comparable results is an open question. Moreover, further research should address more antecedents regarding the willingness to share energy data and smart meter data in particular.

6. Conclusion

Sharing energy data provides added value through digital services for both the personal household and the whole energy system. However, certain drawbacks associated with sharing data such as privacy risks limit the potential of using the available energy data to ensure a more sustainable energy system. Our work clarifies the particular role played by energy data within the spectrum of personal information. While energy data are generally perceived as the least sensitive, smart meter data stands out as an exception. Moreover, a relatively small number of end users express willingness to share their energy data. Given that it is possible to raise awareness about the benefits of sharing smart meter data, given that we find ways to harness further energy data, and given that we effectively communicate these benefits to overcome existing privacy concerns, we conclude that sharing energy data not only mitigates the current controversy regarding digital services but significantly contributes to a sustainable energy system.

CRediT authorship contribution statement

Christian Pfeiffer: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing – original draft. **Stefanie Hatzl:** Conceptualization, Data curation, Investigation, Methodology, Resources, Validation, Writing – review & editing. **Eva Fleiß:** Conceptualization, Data curation, Investigation, Methodology, Resources, Software, Validation, Writing – review & editing. **Alfred Posch:** Funding acquisition, Project administration, Supervision, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data can be retrieved from our Github repository.

Acknowledgements

The financial support by the Climate and Energy Fund [grant number 868693]; and the Austrian Federal Ministry of Labour and Economy, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association [JR-Centre *LiSA*] is gratefully acknowledged. Data were collected in cooperation with respondi AG, Cologne (GER). We thank Sara Crockett for proofreading.

Appendix A. Additional tables

Table A.3
Information types used.

Information type	Information type	Information type
Place of birth[a,b,c,d]	Medical history[a,b,c,d]	Financial account number[a,b,c,d]
Religion[a,b,c,d]	Picture of face[a,b,c,d]	IP address[a,b,c,d]
Body weight[a,b,c,d]	Home address[a,b,c,d]	Social security number[a,b,c]
Email address[a,b,c,d]	GPS location[a,b,c,d]	Passwords[a,b,c,d]
Birth date[a,b,c,d]	Phone number[a,b,c,d]	DNA profile[a,b,c]
ZIP code[a,b,c,d]	Income level[a,b,c,d]	Digital signature[a,b,c]
Number of children[a,b,c,d]	Law enforcement files[a,b,c]	Passport number[a,b,c,d]
Political affiliation[a,b,c,d]	Online dating activities[c,d]	Gender[a,d]
Sexual preference[a,b,c,d]	Credit score[a,b,c,d]	COVID vaccination status[e]
Occupation[a,b,c,d]	Finger print[a,b,c]	Smart meter data[e]
Shopping behavior[a,b,d]	Vehicle registration number[a,b,c,d]	Generation of PV plant[e]
Sports activities[c]	Driver's licence number[a,b]	Heating energy consumption[e]
Social network profile[a,b,c,d]	Credit card number[a,b,c,d]	Electricity consumption of washing machine[e]

Sources: [a] Milne et al. (2017), [b] Markos et al. (2017), [c] Schomakers et al. (2019), [d] Belen-Saglam et al. (2022), [e] new.

Table A.4
Scale items, descriptive statistics and correlations.

Scale	α[a]	M[b]	(SD[c])	RB	IA
Internet privacy concern (IPC)	0.83	0.60	(0.23)	0.43	0.09
I am concerned about submitting information on the Internet because I am unsure what others might do with it.[d,e]					
I am concerned that websites are requesting too much personal information for my registration.[e]					
I am concerned that someone might use my personal information illegally.[d,e]					
I am generally concerned about my privacy when I use the internet.[e]					
I am concerned that someone will use my password illegally.[e]					
Scale	α[a]	M[b]	(SD[c])		IA
Risk behavior (RB)	0.79	0.58	(0.22)		-0.14
I read the privacy policy for mobile devices before downloading an app.[f]					
I actively protect my data when I use the internet.[g]					
Before I use an online service, I find out whether my data is safe there.[g]					
I choose privacy-friendly apps or internet services.[g]					
Scale	α[a]	M[b]	(SD[c])		
Internet activity (IA)	0.77	0.56	(0.15)		
Send instant messages via apps (e.g. WhatsApp, Signal, Telegram)[h,i]					
Use internet banking[i]					
Search for information on the internet[h]					
Plan leisure activities and vacations[i]					
Watch TV and/or listen to music (e.g., web radio, online streaming services such as Spotify, Netflix, Amazon Prime)[h,i]					
Make phone calls or video calls (e.g., via apps such as WhatsApp, Skype, Facebook)[h,i]					
Upload self-created content to websites (e.g., text, images, photos, videos, music, software)[h,i]					
Sell goods or services (e.g., via willhaben, shpock, or at auctions via eBay)[i]					
Use a citizen card/mobile phone signature for official purposes[i]					
Shop online[i]					
Use social networks (e.g. Facebook, Instagram, Snapchat, LinkedIn or Twitter)[i]					
Track or record fitness activities[i]					
Online dating[i]					

$n = 604$, [a] Cronbach alpha, [b] means $\in [0,1]$, [c] standard deviations.
Sources: [d] Dinev and Hart (2006), [e] Lee et al. (2019), [f] Khalil and Karam (2015), [g] Schäwel et al. (2021), [h] Rosen et al. (2013), [i] Djahangiri et al. (2018).

Table A.5
Cluster profiles of information types from hierarchical cluster analysis.

	Demographics (23.08%)		Attitude and Status (15.38%)		Social Interaction (12.82%)		Secure Identifiers (30.77%)		Residuals (5.13%)		Financial Information (5.13%)		Energy Data (7.69%)	
Dimension	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
FIN	0.073	0.077	0.118	0.156	0.095	0.042	0.298	0.203	0.156	0.042	0.812	0	0.127	0.039
PHY	0.062	0.062	0.107	0.090	0.145	0.136	0.107	0.092	0.046	0.012	0.031	0.006	0.009	0.004
SOC	0.163	0.056	0.265	0.083	0.378	0.069	0.220	0.047	0.105	0.007	0.111	0.014	0.045	0.008
PSY	0.073	0.024	0.159	0.052	0.291	0.075	0.159	0.060	0.058	0.012	0.113	0.016	0.018	0.003
PRI	0.389	0.104	0.525	0.110	0.721	0.029	0.804	0.058	0.574	0.116	0.621	0.004	0.228	0.008
NON	0.495	0.069	0.310	0.102	0.135	0.032	0.072	0.027	0.309	0.113	0.017	0.001	0.672	0.037
SENS	0.315	0.071	0.445	0.084	0.648	0.079	0.805	0.077	0.570	0.061	0.912	0.008	0.268	0.041
WITS	0.653	0.074	0.516	0.094	0.320	0.063	0.192	0.075	0.281	0.007	0.109	0.012	0.524	0.087

Table A.6
Profiles of user segments.

	Rejecting user (29.5%)		Concerned user (18.7%)		Selective user (30.7%)		Sharing user (19.7%)	
	M	SD	M	SD	M	SD	M	SD
Age	44.785	13.829	41.696	13.266	42.261	13.650	44.881	14.822
Overall perceived risk								
Attitudes and Status	0.233	0.107	0.405	0.101	0.171	0.088	0.156	0.104
Social Interaction	0.417	0.165	0.715	0.120	0.419	0.139	0.342	0.168
Secure Identifiers	0.427	0.141	0.650	0.121	0.496	0.099	0.390	0.139
Residuals	0.223	0.132	0.310	0.118	0.208	0.128	0.134	0.123
Financial Information	0.543	0.209	0.820	0.139	0.656	0.201	0.550	0.211
Energy Data	0.051	0.056	0.075	0.076	0.023	0.037	0.021	0.040
Demographics	0.120	0.071	0.212	0.082	0.071	0.057	0.067	0.059
Sensitivity								
Attitudes and Status	0.626	0.193	0.499	0.176	0.335	0.150	0.289	0.180
Social Interaction	0.779	0.177	0.736	0.163	0.622	0.176	0.429	0.201
Secure Identifiers	0.888	0.105	0.818	0.138	0.877	0.073	0.589	0.139
Residuals	0.716	0.222	0.571	0.214	0.581	0.242	0.345	0.212
Financial Information	0.960	0.121	0.907	0.162	0.967	0.096	0.815	0.229
Energy Data	0.429	0.273	0.254	0.218	0.169	0.208	0.164	0.197
Demographics	0.483	0.186	0.378	0.181	0.195	0.126	0.171	0.169
Willingness to share								
Attitudes and Status	0.346	0.211	0.466	0.192	0.611	0.166	0.675	0.199
Social Interaction	0.198	0.191	0.244	0.165	0.350	0.197	0.527	0.224
Secure Identifiers	0.132	0.158	0.172	0.131	0.138	0.091	0.371	0.159
Residuals	0.183	0.216	0.278	0.195	0.230	0.201	0.506	0.259
Financial Information	0.076	0.153	0.145	0.164	0.061	0.110	0.185	0.186
Energy Data	0.362	0.266	0.477	0.260	0.596	0.290	0.692	0.247
Demographics	0.488	0.213	0.579	0.205	0.771	0.168	0.791	0.178
Attitudes								
Internet privacy concern	0.634	0.231	0.649	0.203	0.583	0.220	0.492	0.209
Risk behavior	0.619	0.221	0.568	0.209	0.583	0.230	0.530	0.208
Internet activity	0.543	0.152	0.568	0.127	0.577	0.148	0.567	0.160

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.egyr.2024.01.053.

References

Aloise-Young, P.A., Lurbe, S., Isley, S., Kadavil, R., Suryanarayanan, S., Christensen, D., 2021. Dirty dishes or dirty laundry? Comparing two methods for quantifying American consumers' preferences for load management in a smart home. Energy Res. Soc. Sci. 71, 101781 https://doi.org/10.1016/j.erss.2020.101781.

Askeland, M., Backe, S., Bjarghov, S., Lindberg, K.B., Korpas, M., 2021. Activating the potential of decentralized flexibility and energy resources to increase the EV hosting capacity: a case study of a multi-stakeholder local electricity system in Norway. Smart Energy 3, 100034. https://doi.org/10.1016/j.segy.2021.100034.

Ayaburi, E.W., Treku, D.N., 2020. Effect of penitence on social media trust and privacy concerns: the case of Facebook. Int. J. Inf. Manag. 50, 171–181. https://doi.org/10.1016/j.ijinfomgt.2019.05.014.

Belen-Saglam, R., Nurse, J.R., Hodges, D., 2022. An investigation into the sensitivity of personal information and implications for disclosure: a UK perspective. Front. Comput. Sci. 4, 1–22. https://doi.org/10.3389/fcomp.2022.908245.

Bjarghov, S., Löschenbrand, M., Ibn Saif, A.U.N., Alonso Pedrero, R., Pfeiffer, C., Khadem, S.K., Rabelhofer, M., Revheim, F., Farahmand, H., 2021. Developments and challenges in local electricity markets: a comprehensive review. IEEE Access 9, 58910–58943. https://doi.org/10.1109/ACCESS.2021.3071830.

Cagno, E., Franzó, S., Storoni, E., Trianni, A., 2022. A characterisation framework of energy services offered by energy service companies. Appl. Energy 324. https://doi.org/10.1016/j.apenergy.2022.119674.

Chen, C.F., Xu, X., Arpan, L., States, 2017. Between the technology acceptance model and sustainable energy technology acceptance model: Investigating smart meter acceptance in the United. Energy Res. Soc. Sci. 25, 93–104. https://doi.org/10.1016/j.erss.2016.12.011.

Chen, Y.-H., Li, Y.-Z., Jiang, H., Huang, Z., 2023. Research on household energy demand patterns, data acquisition and influencing factors: a review. Sustain. Cities Soc. 99, 104916 https://doi.org/10.1016/j.scs.2023.104916.

Churchill, S.A., Smyth, R., 2021. Locus of control and energy poverty. Energy Econ. 104, 105648 https://doi.org/10.1016/j.eneco.2021.105648.

DellaValle, N., 2019. People's decisions matter: understanding and addressing energy poverty with behavioral economics. Energy Build. 204, 109515 https://doi.org/10.1016/j.enbuild.2019.109515.

Deloitte, 2017. New Technologies Case Study: Data Sharing in Infrastructure, 1–50URL: https://nic.org.uk/app/uploads//Data-sharing-in-infrastructure.pdf.

Dinev, T., Hart, P., 2006. An extended privacy calculus model for ecommerce transactions. Inf. Syst. Res. 17, 61–80. https://doi.org/10.1287/isre.1060.0080.

Djahangiri, N., Weinauer, M., Haßl, G., 2018. ICT usage in households. Technical Report. URL: ⟨http://www.statistik.at⟩.

Döbelt, S., Jung, M., Busch, M., Tscheligi, M., 2015. Consumers' privacy concerns and implications for a privacy preserving Smart Grid architecture - Results of an Austrian study. Energy Res. Soc. Sci. 9, 137–145. https://doi.org/10.1016/j.erss.2015.08.022.

E-Control Austria, 2022. Bericht zur Einführung von intelligenten Messgeräten in Österreich. Technical Report. E-Control. URL: ⟨https://www.e-control.at/documents/1785851/1811582/E-Control-Smart-Meter-Monitoringbericht_2022.pdf/5dc7d56a-5f12–1a02–5662-232e4b41089c?t=1665589737746⟩ [Access on 28/11/2023].

European Union, 2019. Regulation (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU (recast). Online; accessed 05 April 2023.

Featherman, M.S., Pavlou, P.A., 2003. Predicting e-services adoption: a perceived risk facets perspective. Int. J. Hum. Comput. Stud. 59, 451–474. https://doi.org/10.1016/S1071-5819(03)00111-3.

Fell, M.J., 2017. Energy services: a conceptual review. Energy Res. Soc. Sci. 27, 129–140. https://doi.org/10.1016/j.erss.2017.02.010.

Fernandes, D.V., Silva, C.S., 2022. Open energy data—a regulatory framework proposal under the Portuguese electric system context. Energy Policy 170, 113240. https://doi.org/10.1016/j.enpol.2022.113240.

Foster, J.D., Shenesey, J.W., Goff, J.S., 2009. Why do narcissists take more risks? Testing the roles of perceived risks and benefits of risky behaviors. Personal. Individ. Differ. 47, 885–889. https://doi.org/10.1016/j.paid.2009.07.008.

Geels, F.W., Sareen, S., Hook, A., Sovacool, B.K., 2021. Navigating implementation dilemmas in technology-forcing policies: a comparative analysis of accelerated smart meter diffusion in the Netherlands, UK, Norway, and Portugal (2000-2019). Res. Policy 50, 104272. https://doi.org/10.1016/j.respol.2021.104272.

de Godoy, J., Otrel-Cass, K., Toft, K.H., 2021. Transformations of trust in society: a systematic review of how access to big data in energy systems challenges

Scandinavian culture. Energy AI 5, 100079. https://doi.org/10.1016/j.egyai.2021.100079.
Gouthier, M.H., Nennstiel, C., Kern, N., Wendel, L., 2022. The more the better? Data disclosure between the conflicting priorities of privacy concerns, information sensitivity and personalization in e-commerce. J. Bus. Res. 148, 174–189. https://doi.org/10.1016/j.jbusres.2022.04.034.
Hall, M., Geissler, A., 2020. Load control by demand side management to support grid stability in building clusters. Energies 13. https://doi.org/10.3390/en13195112.
Hussain, S., Lai, C., Eicker, U., 2023. Flexibility: literature review on concepts, modeling, and provision method in smart grid. Sustain. Energy Grids Netw. 35, 101113 https://doi.org/10.1016/j.segan.2023.101113.
Jacoby, J., Kaplan, L.B., 1972. The components of perceived risk. Adv. Consum. Res. 382–393.
Janik, A., Ryszko, A., Szafraniec, M., 2021. Determinants of the EU citizens' attitudes towards the European Energy Union priorities. Energies 14. https://doi.org/10.3390/en14175237.
Jiang, T., Hua, T., Xiao, H., Fu, L., Pei, W., Ma, T., 2023. P2P transaction method for distributed energy prosumers based on reputation value. Glob. Energy Interconnect. 6 (3), 308–323. https://doi.org/10.1016/j.gloei.2023.06.005.
Jung, Y., Park, J., 2018. An investigation of relationships among privacy concerns, affective responses, and coping behaviors in locationbased services. Int. J. Inf. Manag. 43, 15–24. https://doi.org/10.1016/j.ijinfomgt.2018.05.007.
Khalil, L., Karam, N.A., 2015. Security management: real versus perceived risk of commercial exploitation of social media personal data. Procedia Comput. Sci. 65, 304–313. https://doi.org/10.1016/j.procs.2015.09.087.
Lee, H., Wong, S.F., Oh, J., Chang, Y., 2019. Information privacy concerns and demographic characteristics: data from a Korean media panel survey. Gov. Inf. Q. 36, 294–303. https://doi.org/10.1016/j.giq.2019.01.002.
Li, W., Yigitcanlar, T., Erol, I., Liu, A., 2020. Motivations, barriers and risks of smart home adoption: from systematic literature review to conceptual framework. Energy Res. Soc. Sci. 80, 102211 https://doi.org/10.1016/j.erss.2021.102211.
Marinakis, V., Doukas, H., Tsapelas, J., Mouzakitis, S., Sicilia, A., Madrazo, L., Sgouridis, S., 2020. From big data to smart energy services: an application for intelligent energy management. Future Gener. Comput. Syst. 110, 572–586. https://doi.org/10.1016/j.future.2018.04.062.
Markos, E., Milne, G.R., Peltier, J.W., 2017. Information sensitivity and willingness to provide continua: a comparative privacy study of the United States and Brazil. J. Public Policy Mark. 36, 79–96. https://doi.org/10.1509/jppm.15.159.
McKenna, E., Few, J., Webbord, E., Anderson, B., Elam, S., Shipworth, D., Cooper, A., Pullinger, M., Oreszczyn, T., 2022. Explaining daily energy demand in British housing using linked smart meter and socio-technical data in a bottom-up statistical model. Energy Build. 258, 111845 https://doi.org/10.1016/j.enbuild.2022.111845.
Milne, G.R., Pettinico, G., Hajjat, F.M., Markos, E., 2017. Information sensitivity typology: mapping the degree and type of risk consumers perceive in personal data sharing. J. Consum. Aff. 51, 133–161. https://doi.org/10.1111/joca.12111.
Niromandfam, A., Yazdankhah, A.S., Kazemzadeh, R., 2020. Modeling demand response based on utility function considering wind profit maximization in the day-ahead market. J. Clean. Prod. 251, 119317 https://doi.org/10.1016/j.jclepro.2019.119317.
Ofgem, 2018. Consumer views on sharing half-hourly settlement data, 1–13. URL: ⟨https://www.ofgem.gov.uk/publications-and-updates/consumer-research-datasets⟩.
Pal, D., Funilkul, S., Zhang, X., 2020. Should i disclose my personal data? Perspectives from internet of things services. IEEE Access 9. https://doi.org/10.1109/ACCESS.2020.3048163.
Panda, S., Mohanty, S., Rout, B.K., Sahu, B.K., Parida, S.M., Samanta, I.S., Bajaj, M., Piecha, M., Blazek, V., Prokop, L., 2023. A comprehensive review on demand side management and market design for renewable energy support and integration. Energy Rep 10, 2228–2250. https://doi.org/10.1016/j.egyr.2023.09.049.
Pandiyan, P., Saravanan, S., Usha, K., Kannadasan, R., Alsharif, M.H., Kim, M.K., 2023. Technological advancements toward smart energy management in smart cities. Energy Rep 10, 648–677. https://doi.org/10.1016/j.egyr.2023.07.021.
Paudler, H.J.A., Schwanitz, V.J., Wierling, A., 2023. Realizing a human-centered digitalization of the energy sector. Open Res. Eur. 2022 2, 140. https://doi.org/10.12688/openreseurope.15129.1.
Pfeiffer, C., Hatzl, S., Fleiß, E., Maier, C., Kremsner, T.P., 2020. A typology of end users' willingness to share energy data. in: Gremmel-Simon, H. (ed.), Science.Research. Pannonia. Leykam. volume 22, 57–66. URL: ⟨https://www.4wardenergy.at/fileadmin/user_upload/Tagungsband_enova2020.pdf⟩.
Phelps, J., Nowak, G., Ferrell, E., 2000. Privacy concerns and consumer willingness to provide personal information. J. Public Policy Mark. 27 (19), 27–41. https://doi.org/10.5005/jp/books/10966-2.
Pití, A., Verticale, G., Rottondi, C., Capone, A., Lo Schiavo, L., 2017. The role of smart meters in enabling real-time energy services for households: the Italian case. Energies 10. https://doi.org/10.3390/en10020199.
Plaum, F., Ahmadiahangar, R., Rosin, A., Kilter, J., 2022. Aggregated demand-side energy flexibility: a comprehensive review on characterization, forcasting and market prospects. Energy Rep. 2022. https://doi.org/10.1016/j.egyr.2022.07.038.
R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: ⟨https://www.R-project.org/⟩.
Radtke, J., 2022. Smart energy systems beyond the age of COVID19: Towards a new order of monitoring, disciplining and sanctioning energy behavior? Energy Res. Soc. Sci. 84 https://doi.org/10.1016/j.erss.2021.102355.
Razavi, R., Gharipour, A., Fleury, M., Akpan, I.J., 2019. Occupancy detection of residential buildings using smart meter data: a large-scale study. Energy Build. 183, 195–208. https://doi.org/10.1016/j.enbuild.2018.11.025.
Reyes, J.A.L., 2022. Willingness to share information for energy efficiency: exploring differences and drivers across the Nordic countries. Energy Sustain. Soc. 12, 1–18. https://doi.org/10.1186/s13705022-00363-3.
Ringnér, M., 2008. What is principal components analysis? Nat. Biotechnol. 26, 303–304.
Rosen, L.D., Whaling, K., Carrier, L.M., Cheever, N.A., Rokkum, J., 2013. The media and technology usage and attitudes scale: an empirical investigation. Comput. Hum. Behav. 29, 2501–2511. https://doi.org/10.1016/j.chb.2013.06.006.
Rumbold, J.M., Pierscionek, B.K., 2018. What are data? A categorization of the data sensitivity spectrum. Big Data Res. 12, 49–59. https://doi.org/10.1016/j.bdr.2017.11.001.
Ryu, D.-H., Kim, K.-J., 2024. The influence of information privacy concerns and perceived electricity usage habits on the usage intention of advanced metering infrastructure. Renew. Sustain. Energy Rev. 189, 113851 https://doi.org/10.1016/j.rser.2023.113851.
Salter, M., Bryden, C., 2009. I can see you: harassment and stalking on the Internet. Inf. Commun. Technol. Law 18, 99–122. https://doi.org/10.1080/13600830902812830.
Schallehn, F., Valogianni, K., 2022. Sustainability awareness and smart meter privacy concerns: the cases of US and Germany. Energy Policy 161, 112756. https://doi.org/10.1016/j.enpol.2021.112756.
Schäwel, J., Frener, R., Masur, P.K., Trepte, S., 2021. Learning by doing or doing by learning? The interdependency between online privacy literacy and data protection behavior. Medien und Kommunikationswissenschaft, 69, 221–246. https://doi.org/10.5771/1615-634X-2021-2-221.
Schomakers, E.M., Lidynia, C., Müllmann, D., Ziefle, M., 2019. Internet users' perceptions of information sensitivity – insights from Germany. Int. J. Inf. Manag. 46, 142–150. https://doi.org/10.1016/j.ijinfomgt.2018.11.018.
Thorve, S., Baek, Y.Y., Swarup, S., Mortveit, H., Marathe, A., Vullikanti, A., Marathe, M., 2023. High resolution synthetic residential energy use profiles for the United States. Sci. Data 10, 76. https://doi.org/10.1038/s41597-022-01914-1.
Tschersich, M., Kiyomoto, S., Pape, S., Nakamura, T., Bal, G., Takasaki, H., Rannenberg, K., 2016. On gender specific perception of data sharing in Japan, in: Hoepman, J., Katzenbeisser, S. (Eds.), IFIP Advances in Information and Communication Technology. Springer. volume 471, pp. 150–160. URL: ⟨http://link.springer.com/10.1007/978–3-319–33630-5–11⟩, doi:10.1007/978-3-319-33630-5-11.
Vigurs, C., Maidment, C., Fell, M., Shipworth, D., 2021. Customer privacy concerns as a barrier to sharing data about energy use in smart local energy systems: a rapid realist review. Energies 14. https://doi.org/10.3390/en14051285.
Vitiello, S., Andreadou, N., Ardelean, M., Fulli, G., 2022. Smart metering roll-out in europe: where do we stand? Cost benefit analyses in the clean energy package and tesearch trends in the green Deal. Energies 15, 2340. https://doi.org/10.3390/en15072340.
Wang, J., Xu, Y., 2021. Internet usage, human capital and CO_2 emissions: a global perspective. Sustainability 13, 8268. https://doi.org/10.3390/su13158268.
Warren, S.D., Brandeis, L.D., 2005. The right to privacy. Inf. Ethics.: Priv., Prop., Power 4, 209–225. https://doi.org/10.7312/gold91730-002.
Watson, R.T., Boudreau, M.C., Chen, J., A., 2010. Management Information Systems Research Center, University of Minnesota. MIS Quarter 34, 567–594.
Yussof, S., Mohd Zulkefle, N.N., Yusoff, Y., Abu Bakar, A., 2021. Willingness of Electricity Consumer in Malaysia, in: Advances in Visual Informatics. IVIC 2021. Lecture Notes in Computer Science. Springer International Publishing, pp. 689–700. URL: ⟨http://dx.doi.org/10.1007/978–3-030–90235-3–59⟩, doi:10.1007/9783–030-90235–3.
Zhang, W., Liu, L., 2022. Unearthing consumers' intention to adopt ecofriendly smart home services: an extended version of the theory of planned behavior model. J. Environ. Plan. Manag. 65, 216–239. https://doi.org/10.1007/9783–030-90235–3. ⟨http://dx.doi.org/10.1007/978-3-030–90235-3–59⟩.