*Journal on Interactive Systems, 2024, 15:1,*doi: 10.5753/jis.2024.3944

Handling of Personal Data by Smart Home Equipment: an Exploratory Analysis in the Context of LGPD

Thiago Adriano Coleti [**Universidade Estadual do Norte do Paraná|** ***thiago.coleti@uenp.edu.br***]
Renato Balancieri [**Universidade Estadual do Paraná|** ***renato.balancieri@unespar.edu.br***]
André Menolli [**Universidade Estadual do Norte do Paraná|** ***menolli@uenp.edu.br***]
Omar Ali Mahmoud [**Universidade Estadual do Paraná|** ***omarmahmoud3611@gmail.com***]
Victor Hugo Sotti [**Universidade Estadual do Paraná|** ***vhsotti@gmail.com***]
Michel Yvano [**Instituto Federal do Pará|** ***myvano@usp.br***]
Marcelo Morandini [**Universidade de São Paulo|** ***m.morandini@usp.br***]

Centro de Ciências Tecnológicas, Universidade Estadual do Norte do Paraná, Rod. BR 369, KM 54, Vila Maria, Bandeirantes, PR, 86360-000, Brasil.

Received: 14 December 2023 • **Accepted:** 12 March 2024 • **Published:** 30 March

Abstract: This paper presents an exploratory research that analyzed the Privacy and Security Policies and the Instruction Manuals of 59 home automation equipment for Smart Home in order to verify which personal data was handled and how these documents were providing information about processes performed in personal data. The analysis was conducted with a quantitative approach followed by a qualitative analysis, using content analysis. The surveys identified the following types of personal data: Identification, Financial, Devices and Location. The results presented greater interest in identification data, although financial and location are also used in specific cases. We also concluded that the Privacy and Security Policies present several information that meets the LGPD's guidelines, especially regarding the purpose of using the data and which personal data is used. However, there is a visible lack of information about the benefits provided to data subjects and about sharing data with third parties, such as recipient data and the legal basis for sharing data.

Keywords: Smart Home, Personal Data, General Data Protection Regulation, Privacy and Security

1 Introduction

The ability to automate the operation of domestic/residential equipment, including electronic gates, smart watches, cleaning tools, air conditioning, and energy-use management devices, utilizing computational resources is known as a "smart home". In this context, it also entails customizing how tasks are carried out and how the user interacts with resources so that they meet their needs and preferences (Wanzeler *et al.*, 2016).

These functionalities may require the handling (collection, processing, sharing, etc.) of personal data, which are records that allow the identification of a software user, their preferences and behaviors (Mortier *et al.*, 2016). In the context of Smart Home, personal data can be understood as registration data, voice commands, photos and videos, browsing history, temperature and humidity data, among others. This data can be obtained from control applications, cameras, microphones, sensors and/or actuators (Vavilov *et al.*, 2014). In a way, it can be said that as Smart Home environments become increasingly ubiquitous, the possibilities for collecting and manipulating personal data become broader and more diverse (Chang and Nam, 2021).

However, the handling of this data raises concerns related to the privacy, security and freedom of data holders, i.e., the person to whom these personal data pertain. This is mainly due to the fact that Smart Home devices are connected to the Internet and to the possibility of data exchange between heterogeneous equipment and/or with different computing environments (Guhr *et al.*, 2020). The General Data Protection Law of Brazil (LGPD)[1] and the General Data Protection Regulation (GDPR)[2], of the European Union both reflect this concern with acts involving personal data. These regulations include lengthy documents that outline the proper conduct and rules that controllers[3] and operators[4] should follow while managing personal data.

Transparency in the use of personal data is among the regulatory requirements. According to Filgueiras *et al.* [2019] and Coleti *et al.* [2020], transparency refers to how well an application communicates information about the people and processes involved in the management of personal data in a comprehensible, visible, and accessible manner. In Article 6º, Section VI of the LGPD, transparency is defined as *a guarantee of clear, accurate, and easily accessible information to data subjects regarding the processing and respective processing agents, while maintaining commercial and industrial secrets.*

Mortier *et al.* [2016] argue that users' knowledge about processes or events performed in their personal data serves as a protective mechanism since they can only take action to protect their privacy, security, and freedom if they are aware of and understand the information regarding the handling of their data Typically, this information is provided in

[1] https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm

[2] https://gdpr-info.eu/

[3] Companies or people that define purposes for using personal data.

[4] Companies or people who carry out computational processes on personal data.

Privacy and Security Policies (PSP), which are difficult to understand or completely inaccessible because they are lengthy and filled with technical and legal jargon. The peculiarities of Privacy and Security Policies (PSP) in the context of Smart Homes combined with the users interest in gadgets and their functionalities may discourage them from thoroughly investigating the PSP. The security and privacy of users in their homes or other residential settings may thus be compromised as a result.

There is a growing concern related to the availability and quality of information regarding the handling of data on PSPs, including that of Smart Home devices. Thus, this paper describes a research in which a set of Smart Home equipment PSPs was analyzed. Our aim was to answer two main questions: *Q1 - What personal data are of most interest to Smart Home equipment?* and *Q2 - Are Smart Home devices complying with LGPD requirements regarding the presentation of information on data handling?* To answer these questions, we selected 59 Smart Home equipment sold on the Internet, which had their PSPs analyzed in a quantitative and qualitative way.

The results are presented throughout this paper and point that, about Q1, there was a greater interest in basic data that allows user identification such as name, email and telephone number, in addition to data as connection and device data. Regarding to Q2, there is still no effort to make information available to users, although some aspects of data handling are presented by the PSP, which allow users to understand some actions with their data. All data are available as complementary material to this paper.

This paper is a follow-up to the one titled *Smart Home Technology: What Do They Want to Know About Us?* (Coleti *et al.*, 2023), published in Workshop on the Implications of Computing in Society (WICS), during the 43th Congress of the Brazilian Computing Society (2023).

The organization of this paper is as follows: The theoretical foundation for this research is described in Section 2, the methodology is presented in Section 3, the discussions are shown in Section 4, the limitations and threats to validity are presented in Section 5 and the final considerations are presented in Section 6.

2 Theoretical Background

The theoretical underpinnings of this study are presented in this section, which covers subjects such as Smart Home, Privacy and the Handling of Personal Data and Related Work.

2.1 Smart Home

The idea of a ”smart home” refers to the capacity to manage, customize, and automate home appliances utilizing computing resources. These controls use electronics, algorithms and data communication mechanisms to provide not only automation, but also the ability to provide intelligence and the ability to recognize and adapt to the context (Wanzeler *et al.*, 2016).

The existence of affordable hardware such as Arduino and Raspberry microcontrollers has accelerated the access and development of technologies to build personalized and intelligent Smart Home environments (Huq *et al.*, 2017; Jang *et al.*, 2019). In this context, the use of Artificial Intelligence (AI) techniques to support the implementation of these resources has become common (Freitas *et al.*, 2010; Luor *et al.*, 2015).

Automation in Smart Home environments can identify contexts of usage and make adjustments in order to understand, foresee, and carry out human preferences and behaviors with the goal of delivering a customized user experience (Freitas *et al.*, 2010; Jang *et al.*, 2019). The extensive use of personal data, which is handled by algorithms, Machine Learning methods, Data Mining, Internet, Bluetooth, among other computing and electronic resources — leads to personalized and user-centered experiences (Basarudin *et al.*, 2017). Examples of services provided by Smart Home environments are:

- Video cameras with facial recognition and identification of people and objects for security purposes (Ben Thabet and Ben Amor, 2015);
- Sensors that identify and learn user behaviors such as schedules and preferences, and adjust equipment operation and energy management (Assaf *et al.*, 2012);
- Multimedia resources that identify users' musicals preferences, films and series. These devices can also connect with others in the home and be controlled by cell phone applications (Lamjane and Rojatkar, 2018).

The architecture of Smart Home environments is based on a model proposed by Islam *et al.* [2022], which comprises: (1) a user access interface, which is usually a cell phone app; (2) one or several microcontrollers such as Arduino and Raspberry, or embedded devices with their own microcontroller boards. These controllers are responsible for receiving commands from the user interface and executing them on the equipment; and (3) Bluetooth or Wi-fi for indoor communication, or even the Internet for communicating and manipulating the equipment remotely.

We can conclude that the idea of smart home automation (or ”smart home”) is expanding rapidly, and future prospects are positive. The viewpoints on home automation to support issues including health, safety, and well-being are highlighted by Singh *et al.* [2018]. The use of smart environment is expected to increase, particularly among the elderly for healthcare and aid with potentially difficult chores. Although only 5% of individuals expressed strong reservations about the manipulation and sharing of their private data, the majority believe the advantages outweigh the risks. Privacy concerns and the handling of personal data are covered in the next subsection.

2.2 Human-Data Interaction and Transparency

The development of technological resources, especially the Internet of Things (IoT), has enabled people to become more immersed in technological resources. This has allowed them to adopt these tools for everyday tasks like studying and working as well as automating the execution of various manual actions, which has allowed an evolution in users' experiences with products and services (Jang *et al*., 2019). It has become standard procedure for users to submit personal data that depicts their traits, actions, habits, and preferences when using apps (Toledo, 2020).

According to Bataineh *et al*. [2016], this data provides companies insights into how to offer products and services among other options in order to achieve a competitive edge, financial advantage and decision-making, and bargaining power. It is nearly impossible to connect with an app without having one's personal data captured (Maus, 2015). The handling of personal data occurs through data gathering via sensors, social networks, websites, and mobile applications. The data subject's privacy, security, and freedom, however, can all be substantially compromised by improper or wrong use due to unauthorized processes (Schneier, 2015).

Regulations for the use of personal data firmly mandate the disclosure of information regarding the use of user data since it is a right of persons who have grown more and more concerned with protecting their privacy (Audich *et al*., 2021; Toledo, 2020). Efroni *et al*. [2019] highlights PSPs are well-known and frequently used to inform users about their rights and obligations as well as on actions taken with regard to their personal data.

However, the volume of texts and the intricacy of the content prevent many people from accessing and using them (Zeng *et al*., 2019). As a result, developers of automation tools and applications need to pay close attention to users' privacy needs. In particular, they should be able to provide users with the knowledge and autonomy they need to assess and evaluate the handling of their data (Zeng *et al*., 2019). Some elements of the Brazilian General Data Protection Law (LGPD) are described in the following subsection.

2.3 General Data Protection Law

The processing of personal data in Brazil is governed by the General Data Protection Law (LGPD)[5], Law No. 13.709, which is enforced in the Federal Government, States, Federal District, and Municipalities. The LGPD inherits characteristics from the General Data Protection Regulation (GDPR) of the European Union, and outlines the rights and obligations of controllers, operators, and data subjects in its articles, sections, and paragraphs.

The LGPD presents several definitions, rules, guidelines, rights and duties of data subjects and controllers/operators of personal data in software applications (Toledo, 2020). The LGPD directly impact the way software applications are developed and the final product delivered to the user, which has led companies to rethink or restructure their development strategies (Camêlo and Alves, 2023; Ribeiro and Garcés, 2023). In this sense, initiatives to improve and ensure user privacy have gained attention, in order to promote privacy as an inherent and inseparable element of technology projects such as Privacy by Design (PbD), which describes seven principles to guide regarding privacy as an inherent part of a project (Cavoukian, 2010; Chalhoub *et al*., 2020; Fischer-Hübner *et al*., 2014).

Among the LGPD guidelines and challenges for software application projects is Personal Data Transparency, shown in Article 6: Personal data processing activities must observe good faith and the following principles: and in Section VI, which reads: transparency: guaranteeing data subjects clear, precise and easily accessible information about the processing and the respective processing agents, respecting commercial and industrial secrets. The LGPD does not present a clear definition of what should be considered transparency, but it is possible to find works in the bibliography that define this concept, such as the work of Filgueiras *et al*. [2019], which defines transparency as the degree to which an application provides perceptible, objective and understandable information to data subjects about agents and events involved in the manipulation of personal data.

The GDPR, on its official website[6], highlights in Articles 13 and 14 a list of information that must be presented to users. It is assumed that, by providing this information, software applications are being transparent.Transparency also encompasses challenges related to information presentation strategies, since providing information about events that occurred when handling data involves explaining algorithms and computational techniques in a way that is understandable to data subjects, who may not be experts in the field of technology. It also considers that users are more concerned about their privacy as they realize that their lives are being more controlled and/or impacted by the manipulation of their data (Audich *et al*., 2021; Coleti *et al*., 2020). Among the possible strategies to support design for Transparency, the following can be mentioned:

- **Security and Privacy Policies**: most used approach, as it technically and legally describes all the rights and duties of those involved in using the application. Because it contains long and complex texts, it is not well regarded by users, who are unlikely to read it completely. There are cases in which design improvements seek to improve the readability of content, but maintain the large and complex textual volume;
- **Privacy Icons**: the use of icons with designs created specifically for privacy and Transparency (Holtz *et al*., 2011), but their use and efficiency are controversial and poorly validated (Efroni *et al*., 2019);
- **Dashboards**: panels with visual resources such as flowcharts, tables and lists to allow the handling of information (Bier *et al*., 2016);
- **Traceview, Organization Chart and Timelines**: panels with visual resources such as flowcharts, tables and lists to allow the handling of information (Murmann and Fischer-Hübner, 2017);
- **Tutorials and examples**: content organized in a didactic way to demonstrate to the individual the events involved in data manipulation (Patrick and Kenny, 2003);
- **Models and Guidelines**: strategies to guide developers in building software applications focusing on data subjects (Coleti *et al*., 2020).

However, even with new strategies being studied and validated, it can be said that the transparency of personal data

[5] https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

[6] https://gdpr-info.eu/

is still a major challenge for developers and researchers, as it involves several social, technical and commercial factors, the latter, with the objective of ensuring that trade secrets are still preserved (Christl, 2017).

The next subsection discusses related works we used as the basis for our paper.

2.4 Related Works

For researchers, worry over the handling of personal data has become a constant. Since the dawn of the twenty-first century, when the Internet experienced significant growth and people's interest in software applications peaked, research in this area has been known to exist. Earp *et al.* [2005] discuss the impact of privacy policy texts in informing the user of actions regarding their data and how such information can increase the user's trustworthiness in the system. The authors sought to verify whether privacy policies provided information that users wanted to know. To do so, they analyzed 24 websites and the results revealed that many more needed to be done to give users privacy policies that are more acceptable to the Fair Information Practice (FIP), particularly on North American websites.

Subahi and Theodorakopoulos [2018] used two different methods to analyze privacy policies in their paper. For the research, they put forward eight criteria that PSP makers of Internet of Things (IoT) equipment should follow. They manually analyzed the PSPs of 11 devices to see how well they adhered to the standards. The results of this initial research indicated that the suggested criteria were not sufficiently adequate. They also developed a software app that kept track of data packets traveling between IoT devices and the cloud, enabling them to draw the conclusion that more than 63% of the equipment met the requirements under consideration.

Still in the context of analysis of PSPs, in the work of Reis *et al.* [2023], a set of 82 PSP, written in Brazilian Portuguese, from applications available at Google Play Store was analyzed. As results, the authors identified documents that were excessively boring, long and poor in readability. Furthermore, the documents contained several flaws in relation to the LGPD, especially with issues such as objectivity, relevance and perception of information.In addition to work analyzing privacy policies, it is also worth highlighting work aimed at creating automatic PSP analysis approaches, such as Kuznetsov *et al.* [2022], who developed a corpus to be used in analysis tools based on natural language processing.

The aforementioned publications examined privacy laws while also taking into account IoT, which is the category that the Smart Home belongs to. However, no research that focused on the Smart Home and examined a sizable amount of technology with high commercial appeal could be found. The tools and procedures used in this investigation are described in the next section.

3 Methodology

In this section we will delve into the comprehensive approach adopted for conducting the research presented in this study. This includes a detailed exploration of the selection of Smart Home appliances, analysis of the most interesting personal data, and the data analysis method used for handling personal data.

3.1 Selection of Smart Home appliances

The first step involved selecting Smart Home appliances. The selection was random and took place in 2021, with searches via Google for terms such as ”Home automation equipment”. The criteria for inclusion were: (1) it should be commercially available; and (2) it should allow interaction with the user directly, either through apps or by collecting data via sensors. Fifty-nine pieces of equipment were selected, which do not cover all models available on the market, but served as a sample for this research. The list of equipment, classified by type and manufacturer, is presented in Table 1.

Table 1. Types of equipment and its manufacturers

Type	Quantity	Manufactures
Virtual Assistants	8	Amazon (2), Facebook, Google (2), Intelbras, Positivo, and Samsung
Video Cameras	3	Ekaza, Logitech, and Positivo
Pet Equipment	2	PetKit
Locks and Doorbells	5	Elsys, Intelbras (2), Netatmo, and Smarteck
Audio and Video Players	4	Apple, LG (2), and Sonos
Sensors	8	Ecobee (2), Houseeasy, Sensative, Simplehuman, Sonoff (2), and Tuya
Outlets, Lamps, and Connectors	11	Novadigital, Philips, Positivo (2), Ring, Smarteck (2), Sonoff (3), and Tuya
Household Utensils	18	CHEF, Eufy, IRobot, Kohler, LG (4), Philips, Positivo, Rachio, Samsung (2), Sensative, Simplehuman, SmartMi, and Tuya (2)

The selected equipment belongs to twenty-eight different manufacturers, fourteen of which had more than one piece device selected, as shown in Table 2. It is noteworthy that there was no preference for any specific manufacturer, since the selection process was carried out solely and exclusively according to the criteria already mentioned.

Table 2. Number of equipment by manufacturer

Qt.	Manufacturers
01	Apple, CHEF, Ekaza, Elsys Facebook, Houseeasy, Irobot, KHOLER, Logitech, Netatmo, Novadigital, Rachio, Ring, Sonos.
02	Amazon, Ecobee, Google, PetKit, Philipps, Sensative, Simplehuman.
03	Intelbras, Samsung, Smarteck.
04	Tuya.
05	Positivo, Sonoff.
06	LG.

The analysis of the data and the results that will be presented later are limited to the list of equipment mentioned and may change as it change.

Next, we present the analysis of the equipment PSP with the aim of identifying the most interest personal data and whether there was information about the actions carried out with personal data, in accordance with the guidelines of Article 6, section VI of the LGPD.

3.2 Analysis of the most interest personal data

In this stage, the PSPs and User Guides for the chosen equipment were read. We chose to read both of them, as we assumed that information on the handling of personal data would be included in them. These materials basically consist of texts, which led to the identification of objective and clear content; however, we also noticed subjective content. This left doubts regarding to which personal data is handled, such as: "*We may also collect data on your computer*"; or "*Some data is collected to improve the user experience*".

Thus, it was necessary for researchers to interpret and infer the content to decide on its handling context. We also noticed the existence of data with semantic similarity that were classified[7] as a result of the supplied final data. One example: the personal data Address was classified as identification data, as it allows the user to be identified in a software app. The personal data Location was classified as Location Data, as it refers to the user's geographic coordinates in a given context of use and is commonly used to assist them in specific actions, in addition to having a broad range of variability.

We found a total of 30 personal data identified as being subject to handling by Smart Home equipment, which were classified into four groups, as shown in Table 3.

Table 3. Classification of personal data mentioned by Smart Home equipment's PSP

Group	Qty.	Description	Handled Personal Data
Identification	19	Data that allows identifying the user, their actions, habits, and preferences.	Name, address, email, country, nickname, phone number, data about the company they work or study at, position, behavioral data, document numbers, date of birth, gender, phone data, biometric data, images, photo, audio, language, marital status, and specific identifiers (such as Apple ID).
Financial	03	Data that allows identifying financial behaviors of an individual.	Credit card data, purchase details (product, service, amount, installment), and billing information.
Device	05	Data that identifies the device and its usage characteristics.	Device data, connection data, usage data, performance data, and browser history.
Location	03	Data that allows obtaining the user's location when using equipment.	Geographical position (latitude and longitude), location information, and system operational data.

To answer the main question **Q1 - What personal data are of most interest to Smart Home equipment?**, we assumed as a highly interest personal data the one indicated by at least 42 pieces of equipment analyzed (70%). This percentage was chosen by the researchers at random considering that a given piece of data would be handled by the majority of analyzed equipment. The personal data (*and it group*) that met this criterion are shown in Figure 1.

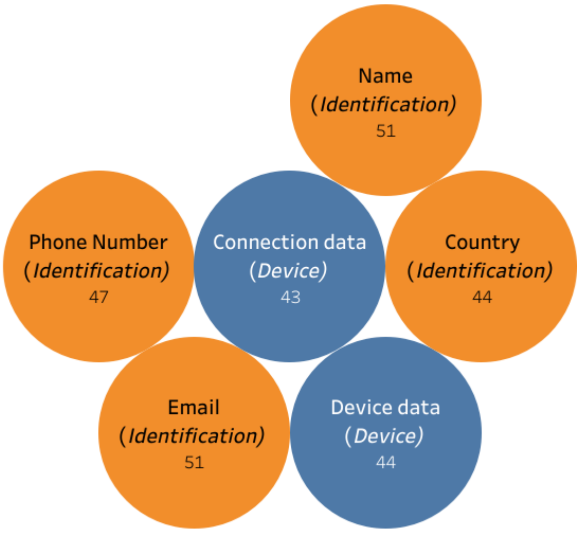

Figure 1. Personal data with the most indications of use

Regarding to Q1 and data presented in Figure 1, we were not surprised to find Name and Email among the most interest piece of data, as it is the minimum necessary for anyone to register for a digital product or service. Of the eight pieces of equipment that did not mentioned the use of the aforementioned fields, three of them did not present any indications regarding the handling of any personal data[8]. For the others, there is an indication of the interest of other data that could eventually replace the data mentioned. For example, instead of Name, they would request Nickname.

As for the personal data E-mail, equipment that did not indicate the handling of this data appears not to request it, as is the case with LED lamps and infrared controls. We also assumed that the equipment is connected to apps that already collect the email. This was also considered for the Country and Telephone Number fields.

Regarding Device Data, we assumed that its handling is necessary to connect equipment with control apps or other equipment. The Location and Financial groups did not present data indicating their use by more than 43 pieces of equipment.

Among the data that did not have a minimum indication of 70 in the Identification group, the personal data Address appeared in 35 devices and the other data had an indication equal to or less than 16 pieces of equipment. The personal data described above were represented by 24, 11, and 16 pieces of equipment for the Financial Data group. Among Device Data, browser history was reported by 32 devices, while performance and usage data by 13 and 12 devices, respectively. Location Data was also not reported by more than 70 of devices; location data and location information were reported by 26 and 28 devices, respectively, while operational data, by 03 pieces of equipment. We thus infer that the information is valuable to Smart Home devices but is processed to satisfy certain needs, which explains why so few devices reported it.

To better comprehend the data of interest from the equipment, we divided Q1 into three additional sub-questions in order to better determine which personal data is most interest to Smart Home technology, namely:

- *Q1.1 - What personal data is of greatest interest to the brands (manufacturers) of the equipment analyzed?*

[7] Classification carried out based on the researchers' empirical knowledge regarding the final objective of the manipulation.

[8] The discussion about equipment that does not indicate personal data will be made later.

- *Q1.2 - Which three manufacturers handle the largest amount of personal data?*
- *Q1.3 - What personal data is less used, but may have future potential?*

To address **Q1.1 - What personal data is of greatest interest to the brands (manufacturers) of the equipment analyzed?**, we investigated indications for handling personal data categorized by type of equipment (Table 1), allowing us to the conclusion that:

- Regarding the Identification Data group, all types of equipment showed data handling indications similar to the analysis carried out to answer Q1; the exception of was 01 device from the Locks and Doorbells group, which reported the handling of unusual data such as photo, image, and audio;
- The Financial Data group has data handled by groups that have direct interaction with the user, such as Virtual Assistants, Audio and Video Players and Household Appliances. We assumed that this occurs because these devices mediate the acquisition of paid products and services for users. However, devices such as sensors, connectors and sockets also reported the use of financial data, which raises questions about the interest in this data, given the characteristics of these devices;
- 89% of the equipment in the Device Data group indicated the handling of at least one piece of data. Six pieces of equipment showed handling of all the device data, belonging to two categories: (1) Sockets, lamps, connectors; and (2) Sensors;
- Finally, there is predominance of the handling of location data and location information in the Location Data group. The types of devices vary greatly, but they are usually related to appliances with movement features such as cameras, Virtual Assistants, and household appliances. However, group equipment such as sensors, sockets, lamps and connectors also reported the handling of location data. It is noteworthy that only 3 pieces of equipment handle operational data.

Thus, **Q1.1** can be answered considering that, for the personal data indicated in Figure 1, the handling profile is maintained when analyzed by type of equipment. The other data have a more significant variation, since there is no consistent pattern of data indicated, even for similar types of equipment. This might be due to the functionalities of the equipment, and also to the fact that there are appliances that complement others (pure commercial aspect). The fact that more than 70% of devices provided minimal data may be directly connected to the lack of a standard and to a device's special interest in personal data. It can also be considered that the desired information may be available in documents or resources not analyzed by researchers.

We calculated the arithmetic mean of the the count of mentions disclosure in the documents by each manufacturer's equipment to answer **Q1.2 - Which three manufacturers handle the largest amount of personal data?**, the result is shown in Table 4, which the manufacturers Intelbras, Samsung and Logitech were those that reported the largest number of interest personal data. All three manufacturers mentioned data in all groups studied in this paper, with subtle differences in the mention of specific data in each group. Their products are in the Locks and bells, Audio and video players and Household Appliances groups.

Table 4. Average counts of mentions of personal data by manufacturer

Manuf./Avg.	Manuf./Avg.	Manuf./Avg.
Amazon: 12	Apple: 13	CHEF: 6
Ecobee: 6	Ekaza: 13	Elsys: 11
Eufy: 9	Facebook: 10	Google: 11
Houseeasy: 11	Intelbras: 15	Irobot: 9
KOHLER: 10	LG: 12	Logitech: 14
Netatmo: 5	Novadigital: 12	PETKIT: 10
Philips: 5	Positivo: 13	Rachio: 13
Ring: 8	Samsumg: 14	Sensative: 9
Simplehuman: 8	Smarteck: 0	SmartMi: 9
Sonoff: 11	Sonos: 7	Tuya: 10

The manufacturers that presented the lowest number of data handled were Netatmo with 02 pieces of Identification data, 02 of Device data and 01 Location data; and Phillips, with the handling of 05 pieces of data. Netatmo Philips handled 04 pieces of information in the Identification and 01 of Financial data.

To answer **Q1.3 - What personal data is less used, but may have future potential?**, the researchers assessed data with less indication of use (less than 20% of the devices) in order to propose scenarios on how they could be used to benefit the controlling company and/or the user, as well as scenarios where they could be harmful to users if handled incorrectly, presenting great potential for handling and obtaining information. Three pieces of personal data were selected by the researchers based on their prior knowledge in research in Data Science and Human-Data Interaction. The selected data were:

- Photo: Its use would be beneficial considering events such as: secure access control; identification of missing people; identification of facial reactions for product evaluation or health examinations. However, the handling of this data could be extremely invasive, as it would allow locating and/or identifying people in various places or situations in their private life, without their knowledge and consent;
- Purchasing data: this data does not pose a situation of great risk for users, but rather potential for inconvenience, since companies can use purchase information to offer other products, identify customer preferences, trace and predict purchasing intentions, among other commercial aspects. Although the aforementioned actions are common on e-commerce websites, in the context of the equipment analyzed in this research, this data was rarely mentioned;
- Usage data: this data allows identifying user behavior and thus personalizing the user experience. However, inappropriate sharing can compromise the user's privacy, as other equipment, suppliers and services would have information about the person's life inside their homes.

The next subsection presents the method of analysis of information regarding the handling of personal data.

3.3 Data Analysis method regarding the handling of personal data

The data analysis was conducted qualitatively, through content analysis, which is a research method used to systematically analyze the content of various forms of communication, such as text, audio, images or video, and involves the systematic examination of the content and structure of a given communication, searching for patterns, themes and relationships within the data (Bardin, 2011; Krippendorff, 2018). Content Analysis seeks to make inferences, which means perceiving interpretative attitudes based on the evidence and indicators raised, supported by a technical validation structure (Bardin, 2011).

Deductive analysis was used, in which a pre-defined set of categories is created and the data collected is coded according to these categories (Krippendorff, 2018). Table 5 provides a comprehensive overview of the analysis structure employed in this study, outlining the distinct Analysis Categories and their corresponding Recording Units (Codes).

Table 5. Units and codes for deductive qualitative analysis

Categories	Codes	Description
1. Data Flow	1.1 - Processing Flow	Description of actions performed with personal data.
2. Scope and Nature	2.1 - Geographical Area	Geographic area coverage of the processing.
	2.2 - Data Source	Data source used to obtain personal data.
3. Purpose of Processing	3.1 - Purpose	Purpose/objective of the processing.
	3.2 - Legal Basis	Law/regulation ensuring the legality of data processing.
	3.3 - Intended Results	Intended outcomes for the data subject.
	3.4 - Expected Benefits	Expected benefits for the organization, entity, or society as a whole.
	3.5 - Information on Disposal	Information on data disposal/anonymization after the end of the intended use.
	3.6 - Period or interval of personal data processing	Temporal information about the processing of personal data.
4. Sharing	4.1 - Reason for Sharing	Reason for sharing the data.
	4.2 - Recipient Contact Information	Contact information of the data recipient.
	4.3 - List of Shared Data	Information on which data is shared.
	4.4 - Legal Basis for Sharing	Law/regulation ensuring the legality of data sharing.

Each analysis category represents a conceptual dimension of the LGPD created based on the LGPD personal data inventory spreadsheet, which is accessible on the website[9], and the TR-Model model of Coleti *et al.* [2020] both contain information categories that were used to code the contents of the PSP.

Using a magnitude scale to represent the degree of information handling personal data, we categorized the coded contents. The adopted magnitude scale was:

- **Fully address (FA)**: PSP provides information that assists individuals in analyzing and making decisions about their data use, ensuring clarity and objectivity without necessitating the consultation of external sources;
- **Partially addresses (PA)**: PSP provides some information, but it is, incomplete, subjective or necessitates further consultation or research with additional data sources;
- **Does not address (DNA)**: information about the code was not identified;
- **Not Applicable (NA)**: the manufacturer/equipment does not handle personal data for the evaluated context.

Table 6 presents the detailed coding results, indicating how each code was classified for each manufacturer. This table allows a general and individualized view of each coding and the magnitude attributed to it.

The discussion on the identified information is presented in Section 3.4.

3.4 Analysis of information regarding the handling of personal data

This subsection presents the analysys conducted in order to answer the **Q2 - Are Smart Home devices complying with LGPD requirements regarding the presentation of information on data handling?**. The findings and debates from the PSP analysis concerning the level of information display on personal data handling activities are included in this section.

Figure 2 presents the percentages of each magnitude for each manufacturer. The percentage was calculated considering the occurrences of magnitudes (as it is a relatively small number, the percentages were rounded, ignoring decimal places).

Considering the data presented in the Figure 2, the first discussion refers to the fact that a considerable difference was identified in relation to the availability of information about processes with personal data. The pieces of equipment analyzed belonged to companies that trade in Brazil and, consequently, collect data in that territory, which fits them into Article 3 of the LGPD, which highlights: *This Law applies to any processing operation carried out by a natural person or by a legal entity under public or private law, regardless of the medium, the country of its headquarters or the country where the data is located.*

The variation in the availability of information can be exemplified by the Fully Address magnitude, in which three companies had more than 50% of their information classified in it, and two others, very well known, did not reach even 10% of the codes in this magnitude. As for the Partially Address magnitude, the variation in values is smaller, but with the exception of one company in which this magnitude was not indicated, all others had their information classified, which indicates some degree of transparency and/or some effort to make the most informative PSP. As for the Does not Address magnitude, the results highlight the need for improvements on the part of companies, since the values indicated in this magnitude are relatively high, given that all companies had at least 20% of the codes indicated in this class.

The dispersion of values for each magnitude can be seen in the graph in Figure 3.

[9] https://www.gov.br/governodigital/pt-br/seguranca-e-protecao-de-dados/templates-e-ferramentas/template_inventario_dados_pessoais.xlsx

Table 6. Results of magnitude (**FA** - Fully Address, **PA** - Partially Address, **DNA** - Does Not Address, **NA** - Not Applicable) indications for encodings in the PSP of Smart Home Equipment Manufacturers

Manufactures / Codes	**1.1**	**2.1**	**2.2**	**3.1**	**3.2**	**3.3**	**3.4**	**3.5**	**3.6**	**4.1**	**4.2**	**4.3**	**4.4**
Amazon	PA	FA	FA	FA	DNA	FA	DNA	FA	PA	FA	DNA	DNA	DNA
Apple	PA	FA	FA	FA	DNA	FA	DNA	FA	PA	FA	PA	DNA	DNA
CHEF IQ	PA	FA	FA	FA	FA	FA	DNA	FA	PA	DNA	DNA	DNA	FA
Ecobee	FA	DNA	FA	DNA	FA	DNA	DNA	FA	FA	NA	NA	NA	NA
Ekaza	PA	DNA	FA	FA	DNA	FA	DNA	DNA	DNA	FA	DNA	DNA	DNA
Elsys	FA	FA	FA	FA	FA	FA	DNA	FA	FA	FA	DNA	DNA	FA
Eufly	PA	PA	FA	FA	DNA	FA	DNA	PA	FA	FA	DNA	DNA	DNA
Facebook	FA	PA	FA	DNA	PA	FA	DNA	FA	DNA	FA	DNA	FA	DNA
Google	PA	PA	FA	FA	DNA	FA	DNA	PA	DNA	FA	DNA	DNA	DNA
Houseeasy	PA	DNA	FA	FA	DNA	DNA	DNA	PA	DNA	DNA	DNA	FA	DNA
Intelbras	PA	FA	FA	FA	FA	FA	DNA	PA	DNA	FA	DNA	DNA	DNA
Irobot	PA	FA	FA	FA	FA	DNA	DNA	FA	DNA	FA	DNA	DNA	DNA
KOHLER	PA	DNA	FA	FA	DNA	FA	DNA	FA	PA	FA	DNA	DNA	DNA
LG	PA	DNA	FA	FA	DNA	FA	DNA	FA	PA	FA	DNA	DNA	DNA
Logitech	PA	DNA	FA	FA	DNA	FA	DNA	FA	DNA	FA	DNA	DNA	DNA
Netatmo	PA	DNA	FA	FA	DNA	FA	DNA	PA	PA	FA	DNA	DNA	DNA
Novadigital	PA	DNA	FA	FA	DNA	FA	DNA	FA	DNA	FA	DNA	DNA	DNA
PETKIT	PA	DNA	FA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA
Philips	PA	DNA	FA	FA	DNA	FA	DNA	FA	DNA	FA	DNA	DNA	DNA
Positivo	FA	FA	FA	FA	DNA	FA	DNA	FA	PA	FA	DNA	DNA	DNA
Rachio	PA	PA	FA	FA	DNA	PA	DNA	PA	DNA	FA	DNA	PA	DNA
Ring	PA	PA	FA	FA	DNA	DNA	DNA	PA	DNA	PA	DNA	DNA	DNA
Samsung	PA	DNA	FA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA
Sensative	PA	PA	FA	FA	DNA	DNA	DNA	PA	DNA	DNA	DNA	DNA	DNA
Simplehuman	PA	PA	FA	FA	PA	FA	DNA	FA	DNA	PA	DNA	DNA	DNA
Smarteck	PA	DNA	PA	FA	PA	DNA	DNA	DNA	PA	PA	DNA	DNA	DNA
SmartMi	PA	PA	FA	FA	FA	DNA	DNA	FA	PA	DNA	DNA	DNA	FA
Sonoff	PA	DNA	PA	FA	DNA	FA	DNA	PA	DNA	FA	DNA	DNA	DNA
Sonos	PA	PA	PA	FA	FA	DNA	DNA	PA	PA	FA	DNA	FA	DNA
Tuya	PA	PA	FA	FA	PA	FA	DNA	PA	DNA	PA	DNA	DNA	DNA

Next, the data is analyzed considering the percentages indicated for each code. Table 7 presents the percentages of magnitudes by code. By observing this table, it is possible to identify which codes received magnitude indications and consequently, which information is made available more frequently or is missing.

Table 7. Percentage of magnitudes (**FA** - Fully Address, **PA** - Partially Address, **DNA** - Does Not Address, **NA** - Not Applicable) by deductive qualitative analysis codes

Codes	**FA**	**PA**	**DNA**	**NA**
1.1 - Treatment flow	13%	87%	0%	0%
2.1 - Geographical area	20%	40%	40%	0%
2.2 - Source	90%	10%	0%	0%
3.1 - Purpose	87%	0%	13%	0%
3.2 - Legal basis	27%	13%	60%	0%
3.3 - Expected Results	63%	3%	34%	0%
3.4 - Benefits	0%	0%	100%	0%
3.5 - Disposal information	50%	37%	13%	0%
3.6 - Period	10%	33%	57%	0%
4.1 - Reason for sharing	60%	17%	20%	3%
4.2 - Contact information	0%	3%	94%	3%
4.3 - Shared Data List	7%	3%	87%	3%
4.4 - Shared legal basis	10%	3%	84%	3%

Regarding the **Fully Address (FA)** magnitude, this indicates that PSP provides information that assists individuals in analyzing and making decisions about their data use, ensuring clarity and objectivity without necessitating the consultation of external sources, in codes *2.2-Source*, *3.1 - Purpose*, *3.3 - Expected Results*, *3.5 - Disposal Information* e *4.1 - Reason for Sharing*. This information is made available by the PSP, becoming an indication of the analysis and decision-making capacity of data subjects. In a way, this information is considered basic on the handling of personal data and the lack of it can cause difficulties and distrust among users in relation to the equipment. The codes mentioned, when added to the *Partially Address* percentages, reach considerable figures, which indicates that this information is made available, even partially to users.

For the **Partially Address (PA)** magnitude, this means PSP provides some information, but it is, incomplete, subjective or necessitates further consultation or research with additional data sources, in the code *1.1 - Treatment flow*. This information can be considered as partially because it involves explaining computational processes within the PSP text, which is considered a challenge given the characteristics of both information. This code refers to very important information for data subjects to understand what will happen to their data, but it lacks methods and techniques to display information with technical/computational aspects in a simple and accessible language.

Regarding the **Does Not Address (DNA)** magnitude, this means information about the code was not identified, the con-

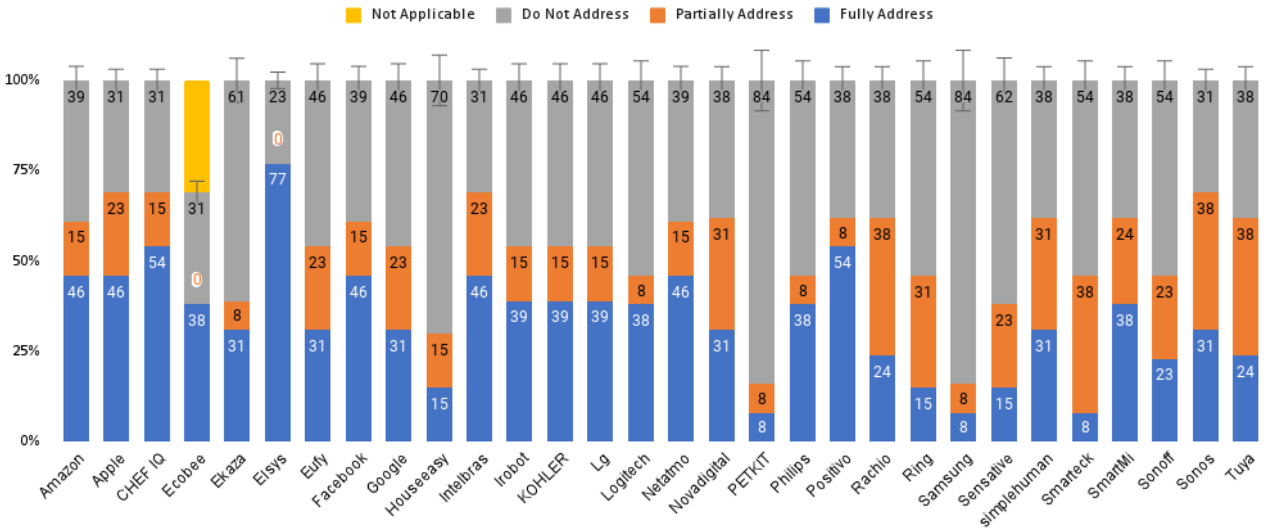

Figure 2. Percentages of magnitudes for each manufacturer.

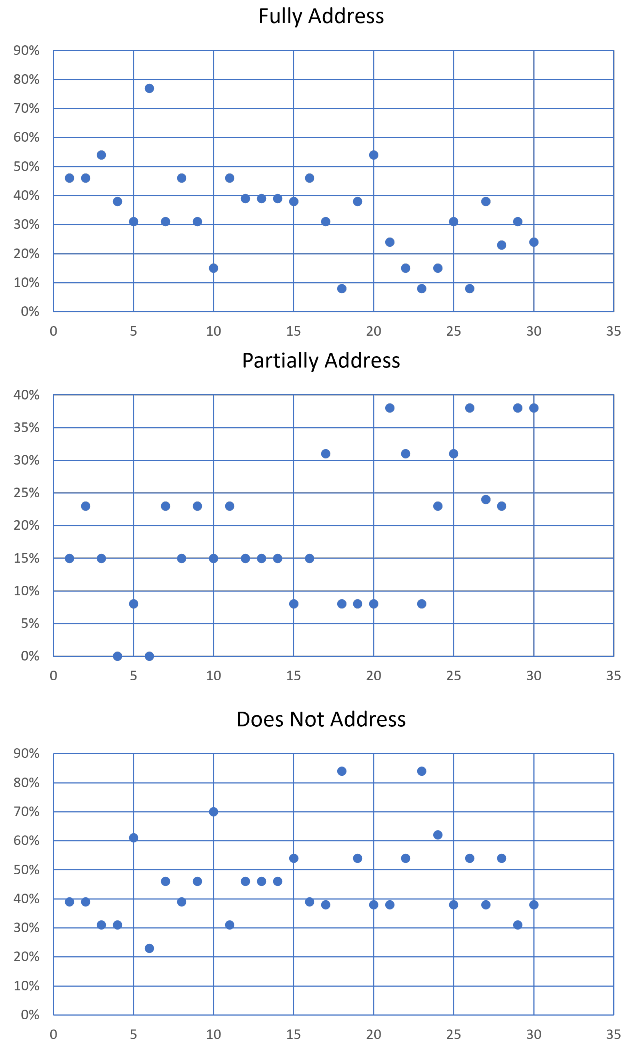

Figure 3. Dispersion of magnitude values

cern was related to some codes that represent information with a certain sensitivity to the user's privacy. Code *3.2 - Legal basis* was not addressed in 60% of the PSP, which indicates a failure to inform data subjects about the legality of data processing, which raises concerns about the need to comply with Art. 6, Section of the LGPD. Furthermore, by informing the law that guarantees the handling of data, the controlling company conveys greater reliability to its users. Still at this magnitude, there was a lack of information detailing the sharing and transfer of personal data between controllers. Codes *4.2 - Contact Information*, *4.3 - Shared Data List* e *4.4 - Shared legal base* were not identified in most PSP. This concern stands out, since data sharing is one of the main fears of data holders, as discussed in Coleti *et al.* [2020] and Filgueiras *et al.* [2019].

However, the biggest concern/surprise for the *Does Not Address* magnitude was for code *3.4 - Benefits*, since not one of the PSP analyzed indicated the benefits of handling personal data for the holders. The lack of this information leads to the idea that there may be no benefit for the user, but only for the company that controls the software, which can seriously affect the privacy, security and freedom of users. Indeed, the benefit to the controller is guaranteed in the LGPD when considering Art. 10, but in practice, the user should be reminded, since it provides inputs for the production of information that may interfere with their routine.

The next section presents the discussion of the results.

4 Discussion

Regarding the most interest personal data for Smart Home, our analysis concluded that despite the Smart Home equipment studied has great capacity and possibilities for data handling, they collected a relatively small amount of personal data. Identification data, such as name, email, country and telephone number were the most observed, followed by device data and connection data. This data is very important, as it is the minimum necessary for a user to register, configure and use the device.

In this sense, although Smart Home devices have a wide capacity to offer customized experiences to their users, we assumed that the data reported is simple and would not always allow learning, prediction and execution of customized experiences. We expected, for example, the handling of certain types of data considered common elements such as audios, photos, document numbers and credit/debit card data; however, these were rarely mentioned. For these data, the following inferences were made by the researchers to justify the low number of notes:

- The handling of other data is done in a very specific way in each equipment; although several personal data have been reported, few are actually handled by the equipment;
- Information about which data is handled was not identified and/or it may be in other documents; also the fact that manufacturers are not fully adjusted to transparency policies in the use of personal data of regulations such as LGPD and GDPR may create difficulties for a user to identify and analyze how their data is handled;
- Data handling may be outsourced, using cell phone apps, tablets and websites. In these cases, these devices work as actuators to turn electronic components on and off. This scenario is widely considered by researchers, since apps present on users' cell phones have a greater capacity for collecting and handling data, not to mention a possible combination with cloud services;
- The fourth possibility would be the fact that, even with minimal personal data provided, the combination of two or more pieces of data, in specific situations, could also produce relevant information about individuals. For this case, we assumed that algorithms could perform combinations and associations between data in order to produce the desired information with the minimum amount of data collected.

Regarding the analysis of information on the handling of personal data, the results presented that there is still no effort to make information available to users, although some aspects of data handling are highlighted by the PSP, which allow users to understand some actions with their data. However, what was noteworthy in this stage was: i) the lack of information about the benefits of processing personal data for its users, information required in the LGPD data inventory; and ii) lack of information about sharing recipients of personal data, an aspect of great concern to individuals.

Also, it can be said that the PSP of Smart Home equipment present certain information about the handling of personal data, which already allows the data subject some insight/knowledge about what is happening. However, when considering the requirement for personal data arising from the personal data inventory model and the information needs requested in the TR-Model, it can be said that there is still a long way to go for appropriate transparency for the data subjects, so that they can interact with simple, objective and relevant information, which supports decision-making.

Furthermore, the lack of information about certain codes may be related to the text of Art. 6, Section VI of the LGPD, as this text highlights the need for Transparency, respecting commercial secrets, but does not specify the minimum amount of information necessary nor how to present it, which allows us to say that there is a great deal of subjectivity in what should be made available as transparency for users.

The next section presents the limitations and difficulties of this research.

5 Limitations and threats to validity

This work was carried out through the analysis (reading) of the PSP of Smart Home equipment. Therefore, this work is limited by the fact that the information extracted was influenced by social, physical and cognitive aspects of the readers. Accordingly, this analysis is subjective, depending exclusively on the reader's interpretation.

This work was carried out through the analysis (reading) of the PSP of Smart Home equipment. Therefore, this work is limited by the fact that the information extracted was influenced by social, physical and cognitive aspects of the readers. Accordingly, this analysis is subjective, depending exclusively on the reader's interpretation.

Finally, the analysis of privacy policies to extract information about the data of interest was carried out in PSP available in 2021. For the analysis of information transparency, the versions available in September and October of the year 2023 were used. Considering that the PSP may undergo updates due to political, corporate, social factors, among others, it is possible that newer versions may be made available, which leads to the existence of a set of information different from that presented in this work.

6 Final Considerations

This paper presented a research in which a set of PSP of Smart Home equipment was analyzed, with the aim of understanding which personal data were of greatest interest from these equipment and whether they were transparent regarding the processes of handling personal data as established in the LGPD. To this end, exploratory analyzes were carried out on equipment documents in which we sought to identify data of interest for processing, as well as information on handling procedures. Quantitative analysis were adopted with subsequent discussion of the results for the data of interest. As for the processes carried out with personal data, qualitative analysis was conducted with content analysis, in which codes from deductive qualitative analysis were used.

In the first step of the research, in which the data of interest from Smart Home equipment were analyzed, there was a greater interest in basic data that allows user identification such as name, email and telephone number, in addition to connection and device data. We presumed that interest in this data because of the need for a minimum identification of each user so that they may enjoy the functionalities and, since there is the possibility of using emails registered on Google, Microsoft platforms, among others, the identification ends up being a simplified process. Financial, device and location data, although with several indications, were of less interest and their use was related to supporting specific equipment functionalities. Thus, the equipment´s data of interest may allow the identification of various characteristics of individuals. However, improper use of data may become invasive and impact the privacy, security and freedom of data subjects, but proper use can significantly improve users' experience with the equipment and its services.

Considering that this work analyzed PSP of 59 Smart Home equipment, a considerable number based on the analyzes available in the literature, it can be concluded that, despite the LGPD being in force and requiring controlling companies to provide transparency with personal data, there is still a lack of improvements regarding the existence of certain information and its quality. This lack of transparency is harmful to users, as they can use products and services without knowing the impact on their personal lives.

It is also worth noting that the trend of data handling by Smart Home equipment is growing, which means that in the future, there will be a need for methods and techniques to guarantee user privacy and transparency in the manipulation of personal data. This is because manipulating personal data involves strong relationships between people, environments, and businesses.

The spreadsheets containing the data utilized in this research are available on Zenodo[10] for verification by developers and users.

7 Acknowledgment

We would like to thank the Araucária Foundation of the State of Paraná and the State University of Paraná (UNESPAR) for the financial support for this research.

References

Assaf, M. H., Mootoo, R., Das, S. R., Petriu, E. M., Groza, V., and Biswas, S. (2012). Sensor based home automation and security system. In *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, pages 722–727. DOI: https://doi.org/10.1109/I2MTC.2012.6229153.

Audich, D. A., Dara, R., and Nonnecke, B. (2021). Improving readability of online privacy policies through doop: A domain ontology for online privacy. *Digit.*, 1:198–215.

Bardin, L. (2011). Content analysis. *São Paulo: Edições*, 70(279):978–8562938047.

Basarudin, N. A., Yeon, A. L., Yusoff, Z. M., Dahlan, N. H. M., and Author, N. M. (2017). Smart home users' information in cloud system: A comparison between Malaysian personal data protection act 2010 and EU general data protection regulation. *Malaysian Construction Research Journal*, 2(2):209–222.

Bataineh, A. S., Mizouni, R., Barachi, M. E., and Bentahar, J. (2016). Monetizing personal data: A two-sided market approach. *Procedia Computer Science*, 83:472–479. DOI: https://doi.org/10.1016/j.procs.2016.04.211.

Ben Thabet, A. and Ben Amor, N. (2015). Enhanced smart doorbell system based on face recognition. In *2015 16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pages 373–377. DOI: https://doi.org/10.1109/STA.2015.7505106.

Bier, C., Kühne, K., and Beyerer, J. (2016). Privacyinsight: The next generation privacy dashboard. In *Privacy Technologies and Policy*, pages 135–152, Cham. Springer International Publishing.

Camêlo, M. N. and Alves, C. (2023). G-priv: A guide to support lgpd compliant specification of privacy requirements. *iSys - Brazilian Journal of Information Systems*, 16(1):2:1–2. DOI: 10.5753/isys.2023.2743.

Cavoukian, A. (2010). Privacy by Design. *Identity in the Information Society*, 3(2):1–12.

Chalhoub, G., Flechais, I., Nthala, N., Abu-Salma, R., and Tom, E. (2020). Factoring user experience into the security and privacy design of smart home devices: A case study. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–9, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3334480.3382850.

Chang, S. and Nam, K. (2021). Smart home adoption: The impact of user characteristics and differences in perception of benefits. *Buildings*, 11(9). DOI: https://doi.org/10.3390/buildings11090393.

Christl, W. (2017). How companies use personal data against people. automated disadvantage, personalized persuasion, and the societal ramifications of the commercial use of personal information. *CrackedLabs*.

Coleti, T., Mahmoud, O., Sotti, V., Menolli, A., Morandini, M., and Balancieri, R. (2023). Equipamentos para smart home: O que eles querem saber sobre nós? In *Anais do IV Workshop sobre as Implicações da Computação na Sociedade*, pages 26–37, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/wics.2023.230083.

Coleti, T. A., Corrêa, P. L. P., Filgueiras, L. V. L., and Morandini, M. (2020). TR-Model. A Metadata Profile Application for Personal Data Transparency. *IEEE Access*, 8(1):75184–75209. DOI: https://doi.org/10.1109/ACCESS.2020.2988566.

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

Efroni, Z., Metzger, J., Mischau, L., and Schirmbeck, M. (2019). Privacy icons: A risk-based approach to visualisation of data processing. *European Data Protection Law Review*, 5(3):352–366. DOI: https://doi.org/10.21552/edpl/2019/3/9.

Filgueiras, L. V. L., Leal, A. S. F., Coleti, T. A., Morandini, M., Correa, P. L., and Alves-Souza, S. N. (2019). Keep System Status Visible: Impact of Notifications on the Perception of Personal Data Transparency. *Human-Computer Interaction. Perspectives on Design*, 1:513–530.

Fischer-Hübner, S., Angulo, J., and Pulls, T. (2014). How can cloud users be supported in deciding on, tracking and

[10] https://zenodo.org/doi/10.5281/zenodo.10372205

controlling how their data are used? In *Privacy and Identity Management for Emerging Services and Technologies*, pages 77–92, Berlin, Heidelberg. Springer Berlin Heidelberg.

Freitas, C. C. S., Mesquita, B. D. R., Pereira, C. E., and Farias, V. J. C. (2010). Automação residencial - uma abordagem em relação as atuais tecnologias e perspectivas para o futuro. In *V Congresso Norte-Nordeste de Pesquisa e Inovação (CONNEPI)*, pages 1–8. DOI: 10.13140/2.1.3590.3683.

Guhr, N., Werth, O., Blacha, P. P. H., and Breitner, M. H. (2020). Privacy concerns in the smart home context. *SN Applied Sciences*, 2(247):1–12. DOI: https://doi.org/10.1007/s42452-020-2025-8.

Holtz, L. E., Nocun, K., and Hansen, M. (2011). Towards displaying privacy information with icons. *IFIP Advances in Information and Communication Technology*, 352 AICT:338–348. DOI: https://doi.org/10.1007/978-3-642-20769-3_27.

Huq, S. M., Rahman, M. A., and Saleh, S. M. (2017). Application for integrating microcontrollers to internet of things. In *2017 20th International Conference of Computer and Information Technology (ICCIT)*, pages 1–4. DOI: https://doi.org/10.1109/ICCITECHN.2017.8281837.

Islam, R., Rahman, W., Rubaiat, R., Hasan, M., Reza, M., and Rahman, M. M. (2022). Lora and server-based home automation using the internet of things (iot). *Journal of King Saud University - Computer and Information Sciences*, 34(6, Part B):3703–3712. DOI: https://doi.org/10.1016/j.jksuci.2020.12.020.

Jang, I., Lee, D., Choi, J., and Son, Y. (2019). An approach to share self-taught knowledge between home iot devices at the edge. *Sensors*, 19:833. DOI: https://doi.org/10.3390/s19040833.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Kuznetsov, M., Novikova, E., Kotenko, I., and Doynikova, E. (2022). Privacy Policies of IoT Devices : Collection and Analysis. *Sensors*, 22(5):1–23.

Lamjane, K. S. and Rojatkar, D. V. (2018). Amazon Alexa Based Home Automation Using Particle Photon. *International Journal of Scientific Research in Science, Engineering and Technology IJSRSET*, 4(May):80–84.

Luor, T., Lu, H.-P., Yu, H., and Lu, Y. (2015). Exploring the critical quality attributes and models of smart homes. *Maturitas*, 82(4):377–386. DOI: https://doi.org/10.1016/j.maturitas.2015.07.025.

Maus, G. (2015). Decoding, hacking, and optimizing societies: Exploring potential applications of human data analytics in sociological engineering, both internally and as offensive weapons. In *Proceedings of the 2015 Science and Information Conference, SAI 2015*, pages 538–547. DOI: https://doi.org/10.1109/SAI.2015.7237195.

Mortier, R., Haddadi, H., Henderson, T., Mcauley, D., Crowcroft, J., and Crabtree, A. (2016). Human-Data Interaction. *Interaction Design Foundation - IxDF*, pages 1–48.

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

Patrick, A. S. and Kenny, S. (2003). From privacy legislation to interface design: Implementing information privacy in human-computer interactions. In *International Symposium on Privacy Enhancing Technologies*.

Reis, V. Q. d., Rabello, M. E. R., Lima, A. C., Jardim, G. P. S., Fernandes, E. R., and Brefeld, U. (2023). Data practices in apps from brazil: What do privacy policies inform us about? *Journal on Interactive Systems*, 14(1):1–8. DOI: 10.5753/jis.2023.2954.

Ribeiro, J. and Garcés, L. (2023). Especificação de requisitos de design de software para sistemas de iot conforme a lgpd: Resultados de aplicação em um sistema de assistência para pacientes com diabetes mellitus. In *Anais Estendidos do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 37–42, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbcas_estendido.2023.229693.

Schneier, B. (2015). *Data and Goliath. The hidden battles to collect your data and control your world*. Norton, New York.

Singh, D., Psychoula, I., Kropf, J., Hanke, S., and Holzinger, A. (2018). Users' perceptions and attitudes towards smart home technologies. In *Smart Homes and Health Telematics, Designing a Better Future: Urban Assisted Living*, pages 203–214, Cham. Springer International Publishing.

Subahi, A. and Theodorakopoulos, G. (2018). Ensuring compliance of iot devices with their privacy policy agreement. In *2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud)*, pages 100–107. DOI: 10.1109/FiCloud.2018.00022.

Toledo, M. D. E. (2020). *Lei Geral de Proteção de Dados. um guia completo*. Emprendedorismo Legal.

Vavilov, D., Melezhik, A., and Platonov, I. (2014). Reference model for smart home user behavior analysis software module. In *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, pages 3–6. DOI: https://doi.org/10.1109/ICCE-Berlin.2014.7034262.

Wanzeler, T., Fülber, H., and Merlin, B. (2016). Desenvolvimento de um sistema de automação residencial de baixo custo aliado ao conceito de Internet das Coisas (IoT). In *Anais do XXXIV Simpósio Brasileiro de Telecomunicações*, pages 40–44. DOI: https://doi.org/10.14209/sbrt.2016.176.

Zeng, E., Mare, S., and Roesner, F. (2019). End user security & privacy concerns with smart homes. In *Proceedings of the 13th Symposium on Usable Privacy and Security, SOUPS 2017*, pages 65–80.