Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Responses of the European IoT Ecosystem to the European General Data Protection Regulation

Frank Ebbers
Fraunhofer Institute for Systems and Innovation Research ISI
Michael Friedewald (michael.friedewald@isi.fraunhofer.de)
Fraunhofer Institute for Systems and Innovation Research ISI

Research Article

Keywords: GDPR, Privacy Regulations, Internet of Things, IoT, Firmware, Up-To-Dateness, IoT fingerprinting

Posted Date: December 12th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3706521/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

With the proliferation of IoT devices in homes and industry, concerns have arisen about the security of the devices and the privacy of users. The General Data Protection Regulation (GDPR), which went into effect in 2018, aims to protect the rights of IoT (and other) users. One way for device manufacturers to address these rights is through firmware updates. However, the extent of the GDPR's impact on EU member states remains uncertain. This paper presents a comprehensive analysis that examines changes in firmware age, as well as the overall age of devices, following the implementation of GDPR in the EU. The analysis uses approximately 400 terabytes of real-world IoT data from Censys.io, spanning from 2015 to the end of 2021. Using grouped average age values, we perform and compare difference-in-differences analyses for devices deployed in all EU member states and regions. We show that devices deployed in Northern Europe tend to be the most up-to-date, while Eastern Europe lags behind. This work provides initial insights into the impact of the GDPR in the IoT ecosystem, and highlights the need for further research to gain a deeper understanding of the effects.

1 Introduction

By 2023, billions of Internet of Things (IoT) devices found their way into almost every area of life, industry and critical infrastructures [1]. While offering many advantages for people's lives and increased production efficiency, these devices collect and process a significant amount of personal data. This data encompasses audio recordings from smart speakers, video snapshots from IP cameras, health and biometric information from smart wearables, user preferences regarding heating temperatures, and bank account details for purchases [1–4]. This situation has attracted the interest of European policymakers [5]. The General Data Protection Regulation (GDPR), effective since May 2018 across all European Union (EU) member states, seeks to harmonize data protection rules [6].

Nevertheless, it remains uncertain whether the GDPR has indeed yielded such outcomes. Quite recently, in July 2023, the European Data Protection Board (EDPB) called for "[h]armonising procedural rules in cross-border cases" [7]. The GDPR's predecessor, the Data Protection Directive 95/46 from 1995, similarly sought to harmonize data protection laws across the EU [8]. While some countries transposed the text into national law verbatim, "others have deeply modified the structure, added new definitions or […] adopted sectoral or specific legislation" [8].

While the GDPR is not exclusively targeted at IoT devices, it does carry significant implications for "IoT devices and their networks, and IoT device makers should take them into account" [9]. A simple, but necessary, first step is to revise privacy policies [10]. Additionally, manufacturers can ensure compliance with Art. 32 GDPR by implementing new privacy features through firmware (FW) updates and patches (U&P) [11]. However, manufacturers often fail to provide U&P regularly and users often ignore to install them [12].

Evaluating the effectiveness of regulations can be a complex endeavor [13, 14]. We propose one potential approach: conducting a case study that examines the interaction between the GDPR and the IoT ecosystem. This study employs the age of devices (as a measure of up-to-dateness) to identify changes before and after the GDPR was adopted (April 2016) and became effective (May 2018). We utilize a dataset of nearly 400 TB, comprising weekly data on the installed base of IoT devices from Censys.io, spanning from 2015 to the end of 2021. This dataset includes information about the device manufacturer, model, and type, and extracts date information from the devices' web user interface (UI) source code. By calculating the mean age of the devices and categorizing them by country and week, we conduct a difference-in-differences (DiD) analysis to pinpoint significant changes.[1]

The remainder of this paper is structured as follows: Section 2 presents the related work. Section 3 outlines the methodology using the SEMMA process. It is followed by the presentation of the results in Section 4, which are discussed in Section 5. We conclude this paper in Section 6.

2 Related Work

The IoT ecosystem has “fundamentally changed the way information technology and communication environments work” [16], thereby transforming the way information is captured and processed. The omnipresence of IoT devices in both private and industry settings [1, 15] has resulted in the processing of various types of data [1–4]. However, users are often unaware that their IoT devices can intrude upon their privacy [17]. Hence, the GDPR is applicable in the IoT domain as well [9]. Nevertheless, there is a noticeable gap in research concerning the impact of regulations, particularly the GDPR, within the IoT domain.

2.1 Privacy Threats of IoT Devices

IoT devices capture great amounts of (potentially personal) data [1–4], often without the users' knowledge or awareness [17]. Simultaneously, users willingly incorporate devices into their surroundings that “often passively collect data relating to daily activities” [18]. Consequently, the “IoT will magnify privacy threats based on the utilization of multiple devices sharing data across networks” [18]. The lack of transparency concerning data processing compounds these risks, resulting in a scenario where data collection occurs unbeknownst to users [17, 19]. Ultimately, users do not feel in control of their data, as privacy notices or controls are often missing [20].

Furthermore, users show a “setup-and-forget” mentality towards their IoT devices [19], meaning users do not configure their devices anymore once they are running. As a result, software or FW U&P, which are already infrequently provided by many manufacturers, are rarely installed by users [12, 21]. Instead, companies prioritize fast time-to-market over data protection and security [22].

2.2 The GDPR and its Effects in the IoT Ecosystem

Although the GDPR is not exclusively directed at IoT devices, it carries significant implications for “IoT devices and their networks and IoT device makers should take them into account” [9]. For instance, a study by [23] identifies challenges such as consent, transparency, data minimization, and privacy by design as the most significant ones in the IoT domain. The GDPR serves as a valuable research subject for several reasons. Firstly, it is directly applicable in all EU member states [6]. Secondly, researchers have foreseen its global impact [24]. Thirdly, it applies to all kind of devices and services that process personal data [5, 25].

The GDPR protects several rights of individuals, such as the right to be informed or the right to data portability [6]. Like any other web service or site, the IoT ecosystem needs to inform users about its data processing practices and provide options for user control, e.g. by manufactures providing FW U&P for new privacy features [11]. According to a 2016 study (i.e. before the GDPR was adopted), “the majority of connected devices fail to adequately explain to customers how their personal data is processed” [25]. Additionally Art. 32 GDPR calls for “appropriate technical measures” to secure data, including U&P as “data protection authorities clearly indicate that regular software updates and using the latest version are part of the required measures” [26].

A persistent challenge in European data protection legislation since the 1980s has been the harmonization of data protection regulations. The Data Protection Directive from 1995 [27] already had this goal. However, a review in 2005 found that some countries translated the text into national law literally, “whereas others have deeply modified the structure, added new definitions or […] adopted sectoral or specific legislation” [8]. This has led to “many diverging interpretations as regards the fundamental concepts” [8]. Similar challenges emerged with the GDPR, primarily due to its numerous opening clauses, such as those pertaining to the healthcare sector. Countries like Germany, Greece, Sweden and Latvia “define divergent rules as to whether healthcare also comprises administration or quality purposes or only immediate patient care” and “leads to divergent justifications for the processing of the same data categories in the member states” [28]. Ultimately, the practical implementation of the GDPR varies significantly among EU member states, including variations in the severity of fines imposed for similar violations [29].

Moreover, the GDPR has had an impact on existing IoT standards and certifications. For instance, ETSI EN 103 645 introduced the user's right to erasure in its version 2.1.2 released in 2020 [30]. The ioXt Alliance will take the GDPR into account in 2023 [31] and additional security standards that might have been influenced by the GDPR are presented in [32].

There are also specific effects of the GDPR in the IoT domain. In 2016, 25 data protection regulators around the world investigated how manufacturers of “smart electricity meters, internet-connected thermostats and watches that monitor health” inform their customers [33]. They found “59 per cent of devices failed to adequately explain to customers how their personal information was collected, used and disclosed”, “68 per cent failed to properly explain how information was stored” and “72 per cent failed to explain how customers could delete their information” [33]. Ultimately, Trend Micro observes that IoT

regulations have the effect of “clarifying accountability and strengthening information sharing among all entities” [34].

2.3 Measuring the Effectiveness of the GDPR

Indeed, there is a notable research gap when it comes to measuring the effectiveness of regulations, including the GDPR [13, 35, 36]. To evaluate their effectiveness, it is crucial to identify measurable critical success factors (CSFs) [14]. In the realm of IoT, assessing the impact of regulations is particularly challenging due to the “shared responsibility“ of regulators, manufacturers, and users [34], the absence of uniform standards, and the multitude of different devices, usage contexts, and data involved [21]. To the best of our knowledge, CSFs for the GDPR in the IoT domain do not exist yet. This leads us to our research question (RQ): *What impact* d*oes the GDPR have on updating IoT devices in member states of the European Union and are those impacts equal for all member states?*

To address this question, we have chosen to consider the general age of IoT devices and the age of their firmware as measurable CSFs, hereinafter referred to as "up-to-dateness." This is a relevant and comparable indicator since older devices often not only fail to comply with the GDPR (specifically Article 32), but they may also harbor numerous unpatched vulnerabilities that pose a threat to the entire Internet [20]. Consequently, the analysis of both the general age of devices and the age of their firmware allows us to assess the impact of the GDPR effectively.

3 Methodology

3.1 General

Our methodology, an extension to the approach presented in [15], is shown in Fig. 1. It is based on the Sampling, Exploring, Modifying, Modeling, and Assessing (SEMMA) process [37], chosen for its iterative nature and its applicability across various industries, offering solutions for a wide range of business problems. We gathered data from the IoT search engine *Censys.io* [38] and used Google BigQuery for the storage and analysis, applying SQL-like queries. Unlike prior work [e.g., 39], we use the source code of the web user interface (UI) for details about the installed FW date and version. Three elements are part of our analysis:

1. Approximately 400 terabytes of *real-world data* from Censys.io, comprising 290 weekly datasets spanning from October 12, 2015, to November 22, 2021. Out of a total of 1,977 available fields, we extracted data from 25 fields that appeared suitable for providing the necessary information. This included examination of the devices' UI source code to extract FW information.
2. A list of 725 *regular expressions* (regexp) to filter the original data sets, enabling us to identify manufacturers and device models, e.g. “ui-label fn-padl20”, which uniquely identifies the device as one of the manufacturer Dahua. Other regexp were used to delete extraneous information, such as standard HTML tags.

We manually created a *mapping table* containing 37,634 device models from 384 distinct manufacturers, categorized into 17 different device types. This table includes their corresponding FW versions and dates obtained from the manufacturers' websites. This mapping allowed us to associate dates with the installed FW on the respective IoT devices and determine whether the current version is up-to-date or outdated.

3.2 Applying the SEMMA process

Sample Generation and Selection of Search Terms.

Sampling aims at consolidating multiple data sources, checking the quality and cleansing [37]. For our work, it was the most important step, as the raw data contained hundreds of terabytes of information. As processing the full count was neither possible, nor advantageous, we decreased the data load by finding only relevant devices, using the following criteria:

- Select only fields that contain information about one of the following: manufacturer, device type, device model, firmware version or date. These were for example "location.country_code", "metadata.manufacturer" or "autonomous_system.name".
- At least one of the fields must not be null.
- We selected only EU member states.
- The source code of the device's web UI must not be null, as it can carry valuable information that are not provided in the fields mentioned above.
- Data from ports indicating not being IoT devices, such as 993 (for IMAP) or 995 (for POP3) were ignored.
- Select 5% of the full data.

In January 2023, we conducted database searches using Google BigQuery. By applying the six criteria outlined above, we were able to create our sample dataset, which consisted of a total of 48,942,402 devices. This dataset served as the foundation for all subsequent processing steps described in the following sections.

Explore - Initial Characterization.

The exploring step aims to find anomalies and data gaps. We found 25 weeks missing. *Censys.io* engineers confirmed that these gaps were irrecoverable. Fortunately, these anomalies did not significantly impede our analysis. Additionally, we encountered certain fields that contained superfluous information. For example, "services.software.update" contained mostly "p1", date fields contained future dates (e.g., 2030) or errors (e.g., "Tue, 16 Feb 02016") we had to control for. Lastly, we noticed that devices from Greece were absent from the dataset. This was attributed to the search for the country code "GR" instead of "EL."

Modify - Adjustments to the Data Set

We made several modifications to the data fields. First, we created a unique identifier (ID) for each entry by calculating a cumulative hash value across all fields to delete duplicates. Next, we focused on missing field values and tried to extract information from the source code using regexp (e.g., "(?:firmware ver\.*\s*\d+(\.*\d*[a-z]*)+)"). This approach was successfully applied by [39] and [15]. Additionally, we extracted general date information (e.g., copyright) and timestamp values from the data. Further modifications involved field merging, for example creating a new manufacturer field composed of entries derived from TLS certificates or hostnames.

Model - Mapping Devices with FW Dates and Analyzing Changes in FW Age over Time.

The final artifact of our SEMMA process is a table containing the following information:

- crawling date of Censys.io data,
- country in which the device deployed in,
- device type,
- device model,
- manufacturer,
- installed FW version as a number (iFV) and
- installed FW date (iFD) as date, derived from the raw data or extracted from the source code,
- age of the installed FW (AiF) as of data set date,
- general age found on the devices' interface.

As not every device provided all information, we applied the rules presented in Table 1. We ignored dates earlier than 1 January 2000, as it is unlikely that it represents a valid date. As a last step, we conducted a DiD regression analysis using STATA to identify the effect of age changes over time.

Assess - Reliability

Usually, a model's reliability is evaluated by applying it to a different set of data. This is impossible in our case, as deployment of the devices is highly volatile due to dynamic IP allocation. However, we conducted a manual review of a sample from the final dataset to ensure that device characteristics were correctly assigned. For the DiD, assessing reliability hinges on the "parallel trend assumption", which we tested using the "graphical examination […] in the pre-treatment period" [40].

Table 1
Rules for modelling the age

Modeling rule				Result in days
iFV	iFD	HTML last-modified tag	General date (e.g. Copyright year)	
	x			Curr_date - iFD = AiF
X			x	Curr_date –copyright year = Age
		x		Curr_date – HTML date tag = Age
			x	Curr_date –copyright year = Age 182 days were added to each year to set the middle of the year for further calculation and avoid bias.
X				No calculation possible

4 Results

4.1 Descriptive Statistics of the Censys.io Raw Data

Our final sample dataset comprised approximately 49 million devices, with an average of 168,766 devices per week (sd = 20,563). Table 2 presents the descriptive statistics. The fluctuations over time are most likely because *Censys.io* randomly selects ports and IP addresses [41]. We do not have an explanation for the low number of devices on 12 Oct. 2015 (~ 74k).

Table 2
Descriptive statistics of the weekly data set (D = date, C = country)

Avg (D)	Min (D)	Max (D)	Avg (C)	Min (C)	Max (C)
168,807	73,631 (12.10.2015)	294,675 (07.02.2017)	6251	143 (Malta)	41,933 (Germany)

4.2 Descriptive Statistics of the Sample Data Set

Frequency Statistics

In total, our approach identified several characteristics of the IoT devices derived from the raw data (Table 3). The disparity between GD and AG is a result of our exclusion of dates earlier than January 1, 2000, as we assume these dates are likely incorrect. The majority of devices are deployed in Germany (24.85%) and France (13.26%), while Malta accounts for 0.08% of the devices.

Table 3
Percentage of identified characteristics

Characteristic	percentage
Precise FW date (FWd)	0.30%
General date (GD)	42.82%
Age of installed FW (AiF)	0.30%
Age based on general date found (AG)	39.50%

When examining the age of the devices, we find that the FW running on them is on average 756 days (d) or approx. 2.07 years (y) old. The average general age of the devices is 751 d (2.06 y). The average general age after the GDPR went into effect went down to 682 d (1.87 y). Devices in Slovakia (1261 d) are the least-up-to-date with FW, while devices in Cyprus are most up-to-date (430 d). The general age of the devices is highest in Luxembourg (1160 d) and lowest in Ireland (462 d).

Statistics of the Age over Time

As the FW age turned out to be very volatile (sd = 428 d), we used grouped mean values of the devices' ages (Fig. 2). Initially we found a significant age drop on 23 October 2018. The mean ages dropped from around 900 d to around 200 d. This is true for every examined country. We examined the reason and found that the *Censys.io* field "updated_at" caused this. From 23 October 2018 onwards, there are about four times more entries showing the exact table date (meaning the date when it was crawled). We do not have a precise explanation for this effect. However, as mentioned in the methodology section, there could be changes in the way *Censys.io* acquired data. We controlled for this by updating all devices' ages to null, where the "updated_at" field was similar to the crawling date. We still see a drop in the data after smoothing the effect, however that drop is not that large anymore (ca. 300 d).

Results of the Difference-in-Differences Analysis

Since the FW age was highly volatile, we used the average general age per week as data for the DiD analysis conducted across all EU member states for two events: adoption of the GDPR (2016) and GDPR in effect (2018). The results do not present a clear-cut picture: while the age of devices increased in some countries (indicating no updates), it decreased in others. This phenomenon holds true for both events, encompassing the same countries. This means that, overall, no significant immediate effect can be identified when the countries are considered separately. The detailed tables with numerical results are available in an online appendix. Thus, we decided to conduct a DiD analysis for four regions of the EU (North, South, East, West) as defined by the UN Statistical Department [42].

Table 4 shows the results after the GDPR went into effect for the respective regions. The findings indicate that devices in Northern Europe are significantly more up-to-date compared to all other regions. For example, when compared to Eastern Europe, the devices age is 203 d lower (Coef. line). In comparison to

Southern Europe, devices are still more up-to-date in Northern Europe, albeit only by 81 d. Devices deployed in Southern Europe are also more up-to-date than compared to Eastern (by 122 d) and Western (by 38 d) Europe. Devices in Eastern Europe are the least updated. All these results are highly significant at $p < 0.001$, except for the comparison between Southern vs. Western Europe with $p = 0.002$.

Table 4
Results of the DiD for the entry into force of the GDPR between EU regions (significance at $p < 0.001$ highlighted)

	Northern vs. Eastern	Northern vs. Western	Northern vs. Southern	Southern vs. Eastern	Southern vs. Western	Eastern vs. Western
F	F(3, 574) = 196.27	F(3, 574) = 304.03	F(3, 574) = 388.10	F(3, 574) = 186.25	F(3, 574) = 462.78	F(3, 574) = 115.02
R^2	0.3917	0.549	0.5816	0.2497	0.532	0.2679
Coef	-202.5921	-118.8052	-80.74207	-121.8501	-38.06316	83.7869
std. Err.	18.14207	14.06948	14.29201	16.93746	12.47771	16.75012
t	-11	-8.4	-5.7	-7.2	-3.1	5
P>|t|	0.00	0.00	0.00	0.00	0.002	0.00
95% conf Int	-238.2251 -166.9592	-146.4392 -91.17129	-108.8131 -52.67106	-155.117 -88.58309	-62.5707 -13.55562	50.8879 116.6859

5 Discussion

According to our analysis, there is a detectable impact of the GDPR on the IoT ecosystem. However, it remains difficult to quantify the impact of the GDPR more precisely and to substantiate it convincingly [13, 35, 36]. Our approach has used the FW age and general device age to investigate changes in EU member states after the GDPR became effective. Age proved to be a useful indicator because it can be measured and compared over time. However, our approach does not account for devices that have not been updated but are newly added to the ecosystem. In theory, this could be addressed by assigning a unique ID to each device and tracking it over subsequent weeks. However, this would only be feasible if the entire dataset were analyzed, rather than just the five percent subset that we worked with.

Our data set is real-world, chronological and large in number. Such data are important to understand impacts and make informed decisions, as “[r]ecorded network traffic alone is typically insufficient for research” [1]. Indeed, data about the actual user would be very interesting to better understand their motivation to (not) install patches. However it is impossible to capture.

The descriptive results demonstrate similar distributions to those identified in [15]. A considerable number of devices are very old and prove the assumption that many devices have been set up and forgotten [19].

Our analysis revealed that, on average, devices are approximately 751 d (2.06 y) old, considering the entire time period. This is a high value and slightly higher than the results from a global study conducted for a data set of April 2020 [15]. Even after the GDPR was enforced, the average general age remained notably high at 1.87 y. One plausible explanation could be that initially, fulfilling the requirements of Article 32 GDPR was not a high priority for manufacturers. Further, according to a 2018 Deloitte study, “44% of respondents [= consumers] believe that organizations care more about their customers’ privacy now that GDPR is in force” [43]. Consequently, consumers may not prioritize their devices' security and privacy as much. Another explanation could be that manufacturers might be too busy to implement the GDPR and U&P are not of a high priority at the moment [22]. Our results suggest that a regulation like the GDPR, even when enforced, may not be sufficient on its own to significantly improve the situation. Consequently, technical means are needed to automate or disburden U&P [15].

Our DiD analysis shows that the entry into force of the GDPR led to a significant age decrease in Northern Europe (Finland, Sweden, Denmark, Estonia, Ireland, Latvia, Lithuania), compared to the other EU region, i.e., more U&P have been carried out there, or more new devices have been installed. One explanation could be the general high digitalization in Northern Europe [44] compared to other EU regions. In these countries digital skills are highly developed [45]. From a security perspective, the Global Cybersecurity Index shows high scores for Northern EU countries (e.g., The Netherlands 97.05/100), compared to Eastern ones (Bulgaria 67.38/100) [46]. However, “the GCI rate does not accurately represent the consumer IoT market” [39].

Furthermore, our results revealed that devices in Eastern Europe are the least updated. Again, this can be explained with the relatively low digitalization in those regions [44]. Looking at the digital skills of Eastern countries, [44] indicates that of the 16–74 aged inhabitants only 28% in Romania and 31% in Bulgaria possess at least basic digital skills. This might directly correlate with the outdated devices, as users must “take an active lead toward protecting their devices” [39] and it “is common for device owners to be unaware of proper safety techniques” [39].

In summary, the results show that the high level of digitalization in Northern Europe does correspond with up-to-date IoT devices. Since we have only observed correlations in our study, our hypotheses for the explanation need to be examined more closely in further research. Our work contributes to the quantitative assessment of regulatory impact, utilizing the GDPR and the IoT domain as a case study. It adds to IoT security research, which “is still at its infancy” [11]. Not to forget that every identifiable IoT devices is considered as a target of an attack vector.

Limitations of this work include the selection bias due to the selected regexp. However, we find this is the best available data source. Not using regexp would have led to a high number of devices that are unnecessary for our analysis. Another limitation is that we cannot control for other (demographic or psychographic) variables that might influence the device age and are not related to the GDPR. Since this work is a pioneering effort for EU member states, it is challenging to provide definitive conclusions about

the results and potential side effects. Instead, it should be considered a preliminary step in addressing the issue. Consequently, we suggest further research in that direction.

Future work could involve the identification of newly added devices and the analysis of FW release notes to investigate whether changes can be attributed to GDPR compliance. Further it would be interesting to allocate the manufactures' legal origins to each device and analyze if there are remarkable changes. Further work could also analyze changes based on different device types.

6 Conclusion

The increasing usage of IoT devices poses privacy risks to users. As the GDPR is considered a comprehensive regulation to ensure fair data processing, evaluating the effect in the IoT domain seems important. To address this, we identified the age of the installed FW and general age of IoT devices, using a real-world data set spanning from 2015 to 2021. Using this data, we conducted a difference-in-differences analysis for EU member states to investigate changes in the age of devices after the GDPR was adopted, and after it went into force. The results indicate that devices deployed in Northern Europe are the most up-to-date, while Eastern Europe lags behind. However, despite the explanation related to the high level of digitization in Northern Europe, we recognize the need for further research in this area. Our analysis represents an initial step and only scratches the surface of this complex issue.

Declarations

Compliance with Ethical Standards

Not applicable. No human participants were involved.

Funding

This work is partially funded by the German Ministry of Education and Research within the project 'Forum Privacy and Self-determined Life in the Digital World' (https://www.forum-privatheit.de) and by the Google Education Credits Program. Google had no influence on the result. The scholarship was only limited to the use of the Google infrastructure (Big Query).

Research Data Policy and Data Availability Statement

The data is provided by Censys.io and hosted on the Google Cloud. Data can be accessed using a paid account from Censys.io. For research they offer a free (but limited) access to their repository on the Google Cloud. Crawling the data must be done using the Google Cloud facilities, i.e. BigQuery. Costs for querying varies based on the amount of data. The authors do not have downloaded the data to own hard drives due to lack of storage and missing facilities to do so usable. Further, the authors do not have the ownership of the original raw data. The authors will provide the analyzed data, e.g., list of devices per country, mean age of the devices over time.

For initial searches (without downloading bulk data), there is an online search engine provided at: https://search.censys.io/.

The results of the DID are provide in the appendix file.

Online appendix available at: http://dx.doi.org/10.24406/fordatis/277

References

1. Danny Y. Huang, Noah Apthorpe, Frank Li, Gunes Acar, and Nick Feamster. 2020. IoT Inspector: Crowdsourcing Labeled Network Traffic from Smart Home Devices at Scale. *Proc. ACM IMWUT* 4, 2, 1–21. DOI: https://doi.org/10.1145/3397333.
2. Frank Ebbers and Murat Karaboga. 2023. Influencing Factors for Users' Privacy and Security Protection Behavior in Smart Speakers: Insights from a Swiss User Study. In *ESORICS 2022 Internat. Workshops*, Sokratis Katsikas, Frédéric Cuppens, Christos Kalloniatis, John Mylopoulos, Frank Pallas, Jörg Pohle, M. A. Sasse, Habtamu Abie, Silvio Ranise, Luca Verderame, Enrico Cambiaso, Jorge Maestre Vidal, Marco A. Sotelo Monge, Massimiliano Albanese, Basel Katt, Sandeep Pirbhulal and Ankur Shukla, Eds. Springer, 195–211. DOI: https://doi.org/10.1007/978-3-031-25460-4_11.
3. Christoph Lutz and Gemma Newlands. 2021. Privacy and smart speakers: A multi-dimensional approach. *The Information Society* 37, 3, 147–162. DOI: https://doi.org/10.1080/01972243.2021.1897914.
4. Lydia Manikonda, Aditya Deotale, and Subbarao Kambhampati. 2018. What's up with Privacy? In *Proc. AAAI*. ACM, 229–235. DOI: https://doi.org/10.1145/3278721.3278773.
5. Sandra Wachter. 2018. Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer Law & Security Review* 34, 3, 436–449. DOI: https://doi.org/10.1016/j.clsr.2018.02.002.
6. European Parliament. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council*: *GDPR*.
7. European Commission. 2023. *New rules to ensure stronger enforcement of the GDPR* (2023). Retrieved July 13, 2023 from https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3609.
8. Yves Poullet. 2006. EU data protection policy. The Directive 95/46/EC: Ten years after. *Computer Law & Security Review* 22, 3, 206–217. DOI: https://doi.org/10.1016/j.clsr.2006.03.004.
9. Sharon Shea and Wigmore. Ivy. 2022. *What is IoT Security?* (2022). Retrieved December 9, 2022 from https://www.techtarget.com/iotagenda/definition/IoT-security-Internet-of-Things-security.
10. Razieh N. Zaeem and K. S. Barber. 2021. The Effect of the GDPR on Privacy Policies. *ACM Trans. Manage. Inf. Syst.* 12, 1, 1–20. DOI: https://doi.org/10.1145/3389685.
11. Nataliia Neshenko, Elias Bou-Harb, Jorge Crichigno, Georges Kaddoum, and Nasir Ghani. 2019. Demystifying IoT Security: An Exhaustive Survey on IoT Vulnerabilities and a First Empirical Look on

Internet-Scale IoT Exploitations. *IEEE Commun. Surv. Tutorials* 21, 3, 2702–2733. DOI: https://doi.org/10.1109/COMST.2019.2910750.

12. Michael Roytman and Jay Jacobs. 2019. The complexity of prioritising patching. *Network Security,* 7, 6–9. DOI: https://doi.org/10.1016/s1353-4858(19)30082-0.
13. Cary Coglianese. 2012. *Measuring Regulatory Performance* (2012). Retrieved February 2, 2023 from https://www.oecd.org/gov/regulatory-policy/1_coglianese%20web.pdf.
14. Jose M. Torres, Jose M. Sarriegi, Javier Santos, and Nicolás Serrano. 2006. Managing Information Systems Security: Critical Success Factors and Indicators to Measure Effectiveness. In *Information Security*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Sokratis K. Katsikas, Javier López, Michael Backes, Stefanos Gritzalis and Bart Preneel, Eds., 530–545. DOI: https://doi.org/10.1007/11836810_38.
15. Frank Ebbers. 2022. A Large-Scale Analysis of IoT Firmware Version Distribution in the Wild. *IIEEE Trans. Software Eng. DOI: https://doi.org/*10.1109/TSE.2022.3163969.
16. Mark M. Ogonji, George Okeyo, and Joseph M. Wafula. 2020. A survey on privacy and security of Internet of Things. *Computer Science Review* 38. DOI: https://doi.org/10.1016/j.cosrev.2020.100312.
17. Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, 1–20. DOI: https://doi.org/10.1145/3274469.
18. C. B. Foltz and Laura Foltz. 2020. Mobile users' information privacy concerns instrument and IoT. *ICS* 28, 3, 359–371. DOI: https://doi.org/10.1108/ICS-07-2019-0090.
19. Constantinos Kolias, Georgios Kambourakis, Angelos Stavrou, and Jeffrey Voas. 2017. DDoS in the IoT: Mirai and Other Botnets. *Computer* 50, 7, 80–84. DOI: https://doi.org/10.1109/MC.2017.201.
20. Imran Makhdoom, Mehran Abolhasan, Justin Lipman, Ren P. Liu, and Wei Ni. 2019. Anatomy of Threats to the Internet of Things. *IEEE Commun. Surv. Tutorials* 21, 2, 1636–1675. DOI: https://doi.org/10.1109/COMST.2018.2874978.
21. Devender Bhushan and Rashmi Agrawal. 2021. Security and Privacy in IOT. In *Proc. RICE*, Raghvendra Kumar, Nguyen H. Quang, Vijender Kumar Solanki, Manuel Cardona and Prasant K. Pattnaik, Eds. Springer, 673–679. DOI: https://doi.org/10.1007/978-981-15-7527-3_63.
22. Ari E. Waldman. 2018. Designing Without Privacy. *Houston Law Review* 55, 3.
23. Daniel Bastos, Fabio Giubilo, Mark Shackleton, and Fadi El-Mousa. 2018. GDPR Privacy Implications for the Internet of Things. *4th Annual IoT Security Foundation Conference*.
24. Nicholas Martin and Frank Ebbers. 2022. When Regulatory Power and Industrial Ambitions Collide: The "Brussels Effect," Lead Markets, and the GDPR. In *Privacy Symposium 2022*, Stefan Schiffner, Sebastien Ziegler and Adrian Quesada Rodriguez, Eds. Springer, 129–151. DOI: https://doi.org/10.1007/978-3-031-09901-4_8.

25. Davide Borelli, Ningxin Xie, and Eing Kai Timothy Neo. n.n. *The Internet of Things: Is it just about GDPR?* (n.n). Retrieved March 7, 2023 from https://www.pwc.co.uk/services/risk/technology-data-analytics/data-protection/insights/the-internet-of-things-is-it-just-about-gdpr.html.
26. Alana Maurushat and Kathy Nguyen. 2022. The legal obligation to provide timely security patching and automatic updates. *Int. Cybersecur. Law Rev.* 3, 2, 437–465. DOI: https://doi.org/10.1365/s43439-022-00059-6.
27. European Parliament. 1995. *Directive 95/46/EC*.
28. Fruzsina Molnár-Gábor, Julian Sellner, Sophia Pagil, Santa Slokenberga, Olga Tzortzatou-Nanopoulou, and Katarina Nyström. 2022. Harmonization after the GDPR? *Seminars in cancer biology* 84, 271–283. DOI: https://doi.org/10.1016/j.semcancer.2021.12.001.
29. Martin Holzhofer. 2023. *Fines for violations of the GDPR* (2023). Retrieved July 17, 2023 from https://www.dsgvo-portal.de/gdpr-fine-database/.
30. ETSI. 2020. TS 103 645 - V2.1.2.
31. Brett Brune. 2023. *Cybersecurity Label for U.S. Coming as Early as April* (2023). Retrieved September 27, 2023 from https://www.eetimes.com/cybersecurity-label-for-u-s-coming-as-early-as-april/.
32. Nickson M. Karie, Nor M. Sahri, Wencheng Yang, Craig Valli, and Victor R. Kebande. 2021. A Review of Security Standards and Frameworks for IoT-Based Smart Environments. *IEEE Access* 9, 121975–121995. DOI: https://doi.org/10.1109/ACCESS.2021.3109886.
33. ICO. 2016. *Privacy regulators study finds Internet of Things shortfalls* (2016). Retrieved March 7, 2023 from https://web.archive.org/web/20170906011408/https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2016/09/privacy-regulators-study-finds-internet-of-things-shortfalls/.
34. Trend Micro. 2018. *Securing the Internet of Things Through Effective Regulation* (2018). Retrieved April 16, 2023 from https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/securing-the-internet-of-things-through-effective-regulation.
35. Knut Blind, Crspin Niebel, and Christian Rammer. 2022. The Impact of the EU General Data Protection Regulation on Innovation in Firms. *SSRN Journal. DOI: https://doi.org/*10.2139/ssrn.4257740.
36. Crispin Niebel. 2021. The impact of the general data protection regulation on innovation and the global political economy. *Computer Law & Security Review* 40, 105523. DOI: https://doi.org/10.1016/j.clsr.2020.105523.
37. SAS Institute. 2020. *Introduction to SEMMA* (2020). Retrieved August 28, 2022 from https://documentation.sas.com/.
38. Censys. 2022. *Search* (2022). Retrieved August 8, 2022 from https://search.censys.io/.
39. Antonio Mangino, Morteza S. Pour, and Elias Bou-Harb. 2020. Internet-scale Insecurity of Consumer Internet of Things. *ACM Trans. Manage. Inf. Syst.* 11, 4, 1–24. DOI: https://doi.org/10.1145/3394504.
40. Landon Gibson and Frederick Zimmerman. 2021. Measuring the sensitivity of difference-in-difference estimates to the parallel trends assumption. *RMM* 2, 4, 148–156. DOI:

https://doi.org/10.1177/26320843211061306.

41. Tom Simonite. 2015. *Search Engine Censys Knows the Internet's Dirty Little Security Secrets* (2015). Retrieved April 7, 2020 from https://www.technologyreview.com/s/544191/a-search-engine-for-the-internets-dirty-secrets/.
42. UNSD. 2023. *Standard country or area codes for statistical use (M49) - Geographic Regions* (2023). Retrieved July 12, 2023 from https://unstats.un.org/unsd/methodology/m49/.
43. Deloitte. 2018. *A new era for privacy - General Data Protection Regulation ("GDPR") six months on* (2018). Retrieved May 26, 2023 from https://www2.deloitte.com/uk/en/pages/risk/articles/gdpr-six-months-on.html.
44. European Commission. 2023. *The Digital Economy and Society Index (DESI)* (2023). Retrieved July 17, 2023 from https://digital-strategy.ec.europa.eu/en/policies/desi.
45. eurostat. 2022. *How many citizens had basic digital skills in 2021?* (2022). Retrieved July 17, 2023 from https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220330-1.
46. *Global Cybersecurity Index 2020 - D-STR-GCI.01-2021-PDF-E.pdf*. Retrieved July 18, 2023 from https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-GCI.01-2021-PDF-E.pdf.

Footnotes

1. Parts of this text (sect. 1, 3) have already been published in an article with a focus on a one-time snapshot, analyzing influencing factors on the IoT up-to-dateness [15].

Figures

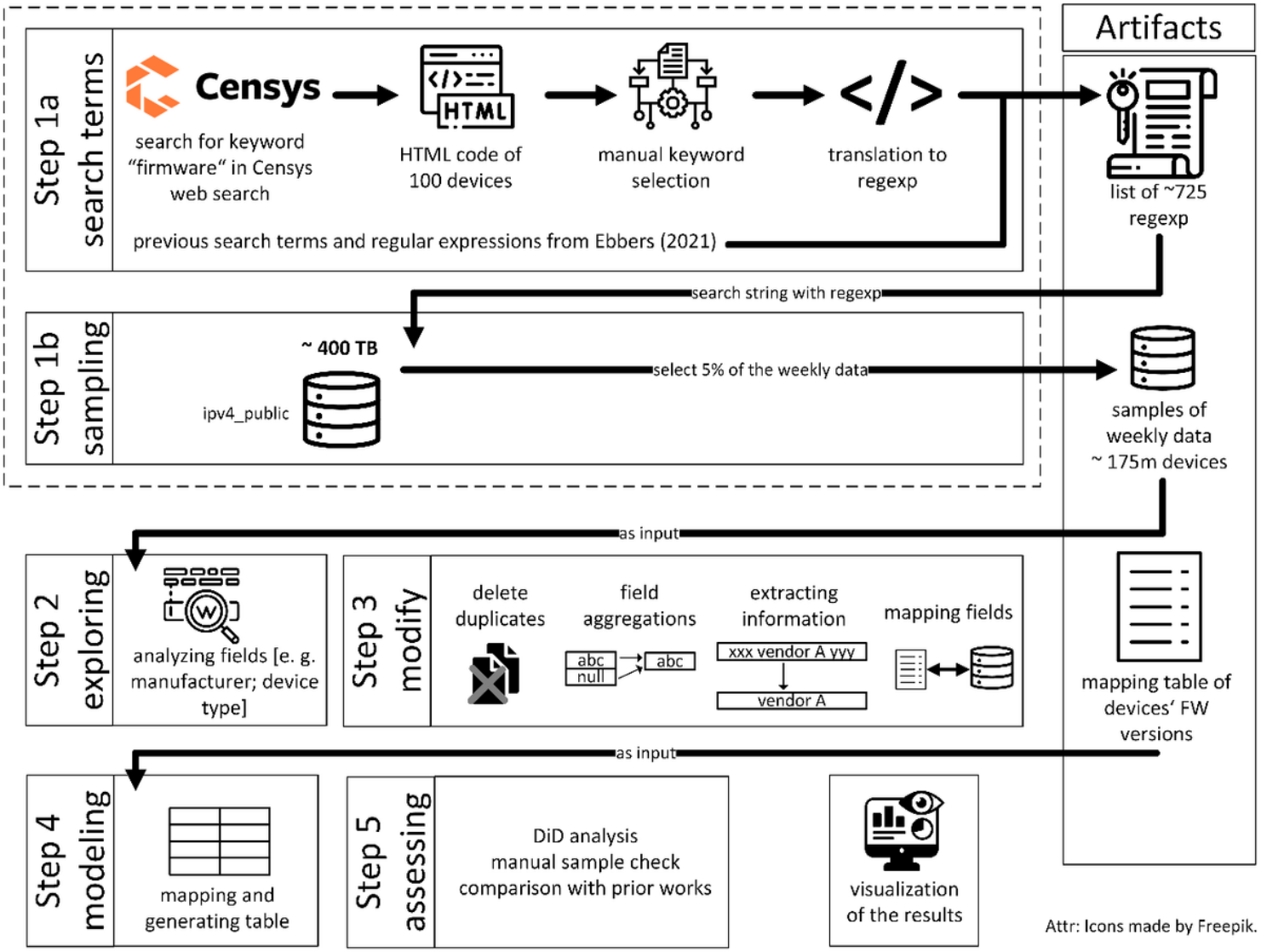


Figure 1

Methodology

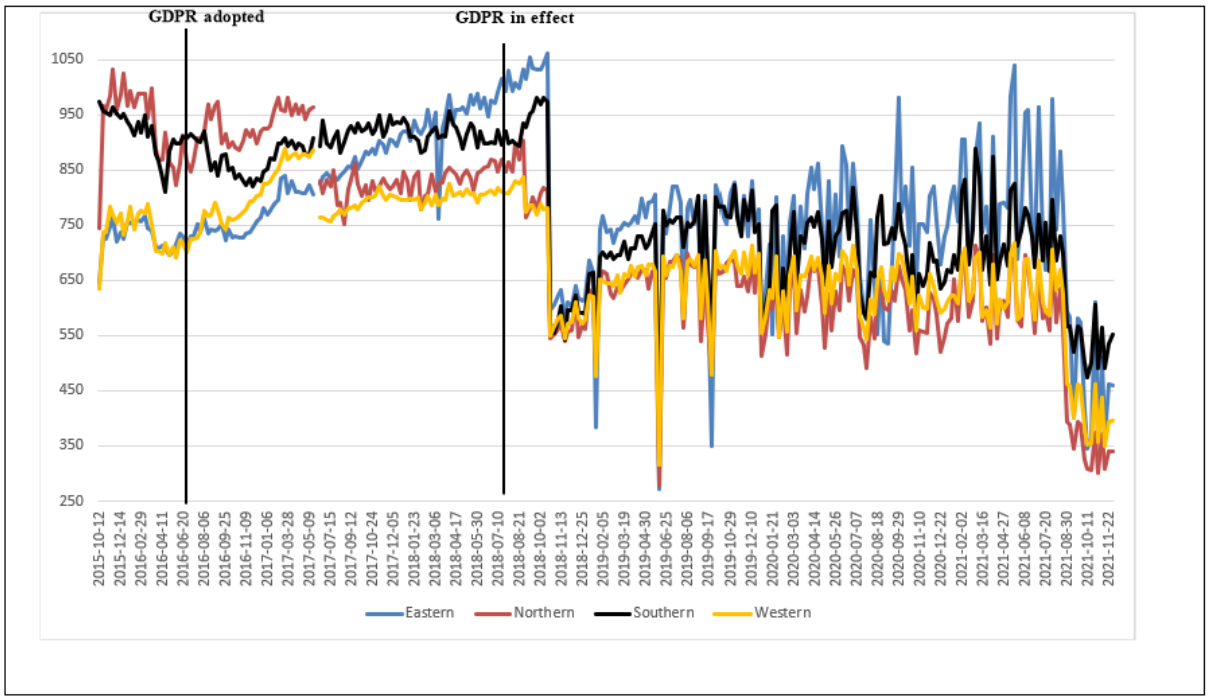


Figure 2

General age over time by region (white space on 23rd May 2017 due to missing input data)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- OnlineAppendixDiDResults.pdf