Department: Head
Editor: Name, xxxx@email

Towards Privacy-Friendly Smart Products

Kimberly García
University of St. Gallen (HSG)

Zaira Zihlmann
University of Lucerene

Simon Mayer
University of St. Gallen (HSG)

Aurelia Tamo-Larrieux
University of St. Gallen (HSG)

***Abstract*—Smart products, such as toy robots, must comply with multiple legal requirements of the country they are sold and used. Currently, compliance with the legal environment requires manually customizing products for different markets. In this paper, we explore a design approach for smart products that enforces compliance with aspects of the European Union's data protection principles within a product's firmware through a case study on a toy robot. This endeavour has taken us through an exchange between computer scientists and legal scholars to determine the relevant data flows, their processing needs, and the implementation decisions that would allow a device to operate while complying with the EU data protection law. By designing a data-minimizing toy robot, we show how the variety, amount, and quality of data that is exposed, processed, and stored outside of a user's premises can be considerably reduced while preserving the device's functionality. In comparison with a robot designed using a traditional approach, where 90% of the collected types of information are stored with the data controller or a remote service, our proposed design leads to the mandatory exposure of only seven out of 15 collected types of information, all of which are legally required by the data controller to demonstrate consent.**

INTRODUCTION. In a globalized world, products move across geographic regions, changing jurisdictions that make them subject to heterogeneous or even incompatible legal environments, e.g., product safety regulations, privacy and data protection regulations. This traditionally leads to the creation of multiple product variants that need to be customized and managed by the manufacturer to serve those different markets [1]. As more and more 'smart' products appear, this

variants management does not only affect the products' hardware, but also the software that these products run, often in the form of pre-installed firmware.

Given that the European Union (EU) is an important marketplace for technology (including social robots) and that EU manufacturers are required to reflect EU standards in their variant management [2], we have taken a toy robot as a use case to investigate how the legal environment of such a smart product is reflected in its firmware implementation.

Although this work focuses on a device's firmware, the creation of software that complies with legal regulations is a current and pressing issue beyond smart products. A recent event demonstrate this dramatically: The social media app TikTok[1] agreed, after the tragic death of a minor involved in a dangerous challenge, to implement measures to re-verify the age of all its users in Italy, in order to block users younger than 13 years old. This case demonstrates possible consequences of omitted age verification and consent measures.

Specific to our use case, we consider the data protection principles of the General Data Protection Regulation (GDPR) of the EU to analyze the legally relevant aspects of a home toy robot's firmware. This analysis is difficult since, in contrast to other jurisdictions [3], in the EU a specific legal ground must be fulfilled to process personal data in accordance with the fundamental principles of data protection law. Thus, there must be an explicit and valid reason for each piece of data to be processed. Furthermore, the GDPR contains a norm on privacy by design and default, which mandates that data controller (DCs) ensure, via technical and organizational measures, that they comply with the fundamental principles of data protection law (Art. 25 GDPR). How to implement privacy-by-design requires a case-by-case analysis and the tailoring of measures to a concrete use case [4], such as data collection by a educational toy robot.

However, the fundamental principles of data protection law are challenged by the data processing capabilities of toy robots. Such robots are equipped with various privacy-sensitive sensors, such as microphones and cameras, that are continuously processing personal data. The robots can easily be moved from one jurisdiction to another and can affect individuals, including vulnerable user groups such as children, in their private homes. Thus, it is not surprising that the privacy implications of such social robots have been explored in various disciplines [5] [6] [7].

In this research, we explore how to design firmware for smart products capable of automatically enforcing the compliance of data protection principles by adapting, at run time, its data processing activities. While various data protection principles apply in our scenario, we focus on the implementation of a toy robot capable of complying with the consent requirement of data protection law. What are the implementation decisions that need to be made to be able to encode these requirements into a toy robot?

Case Study: An Educational Toy Robot

Our toy robot was designed as a (mock) educational tool for young children. This robot roams private family premises, takes pictures of its environment every few seconds, and analyzes them to identify any known individuals (typically children) in its view. If an individual has been identified, the robot stops to perform an educationally valuable action, such as playing a song that motivates the identified individual to sing along to pre-selected personalized content. This content would be relevant for the children's age and current interests. Thus, in order to operate, the toy robot needs to record, process, and store several pieces of data, including personal data.

We have analyzed a traditional software engineering approach in order to identify the data flows that a toy robot needs to process in order to operate. Then, a brief legal analysis is presented to adapt and augment the data flows of the toy robot to fit the consent requirements of Article 4(11) and Articles 7 to 9 of the GDPR, whilst keeping in mind the other data protection principles set out in Article 5, particularly the data minimization principle. The legal requirements are interpreted in a practice-oriented way in order to establish concrete engineering and design recommendations for the implementation of the toy robot. Our main contribution is an exploration of

[1] https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9533424#english_version

a methodology that enables the creation of legally compliant cyber-physical systems.

What Might Be: Traditional Design of a Toy Robot

The following is a "what-might-be" analysis of a toy robot that has been designed following a conventional software engineering approach.

We begin with the setup process of the toy robot and show the flow of data in this process in Figure 1[2]: After unpacking the toy robot, a user is prompted to set up an online account with the robot's supplier i.e., the data controller (DC) according to GDPR, which requires the user to enter an email address and password as well as to complete a 5-min demographic survey (e.g., household members' names, genders, ages, and interests). The user then receives a confirmation email from the website. Subsequently, the user is prompted to enter their robot's serial number and assign a name to it. Next, the user is asked to connect the robot to their home WiFi. To accomplish this, the user switches the robot to a special setup mode that allows the user to connect to the robot's WiFi. Once connected, the user is directed to the robot's locally hosted home page to enter the credentials for their home WiFi. Subsequently, the robot connects to the home WiFi. The user then finalizes the setup through the online portal by entering (for each individual that shall be recognized) a pseudo-identifier along with a set of pictures of the identified individual (i.e., for training a facial recognition algorithm), and configures the robot to provide personalized responses for each recognized individual. This information is then linked to the household members and the training pictures and pseudo-identifiers are sent to a third-party facial recognition service for training (in our implementation, the Google Vision API is used).

Once the setup process has concluded, the robot starts roaming around the user's home. Every five seconds, the robot takes a new picture and uploads it to the DC's website via WiFi. The DC then forwards the picture to a facial recognition service, obtains the identity of the identified individual, and sends this information to the robot. Since the DC aims to become independent of the third-party facial recognition service, it retains all uploaded pictures to bootstrap and improve its own facial recognition algorithm. Finally, the robot performs an action based on the personalized configuration. Moreover, the robot uploads (to the DC) a daily usage report that includes, among others, data on the duration of usage, distance travelled, GPS location, number of classified individuals with classification reliability estimates, and system errors. This data is analyzed by the DC to identify future improvements to the device and the service, and to address further monetary exploitation opportunities.

Table 1 provides an overview of the pieces of data that are processed during the robot setup process and in its subsequent operations. For instance, the table shows that the home network credentials (#7 in Table 1) are the only piece of application-relevant data that does not leave the robot. Moreover, the business strategy of the DC is strongly reflected in the system design, since it shares as little data as possible with the Remote Service (RS) attempting to become independent of it. To achieve this, the DC collects as much information as possible from the user and stores it, ideally indefinitely, to keep improving its product and services, and to enable it to further monetize the acquired information.

Legal Considerations for Data Collection and Processing

In order to design a toy robot that is compliant with the GDPR, we have identified several legal considerations that should be addressed in its design and implementation. Following, we focus on the characteristics of consent and discuss how to remain compliant within our toy robot use case.

Consent should be Timely

Processing of personal data that is based on consent must meet specific requirements laid out by the GDPR. While the specific requirements mostly require a case-by-case analysis, the timing requirement triggers a straightforward engineering implication, since consent must be obtained prior to the data processing activity [8][9]. In the case of smart products, it is thus key that 'set-up-notices' are provided to the user upon initial use of the device [10]. While the focus

[2]This process is aligned with the setup of popular connected toys such as the Tonies ecosystem, see https://tonies.de/

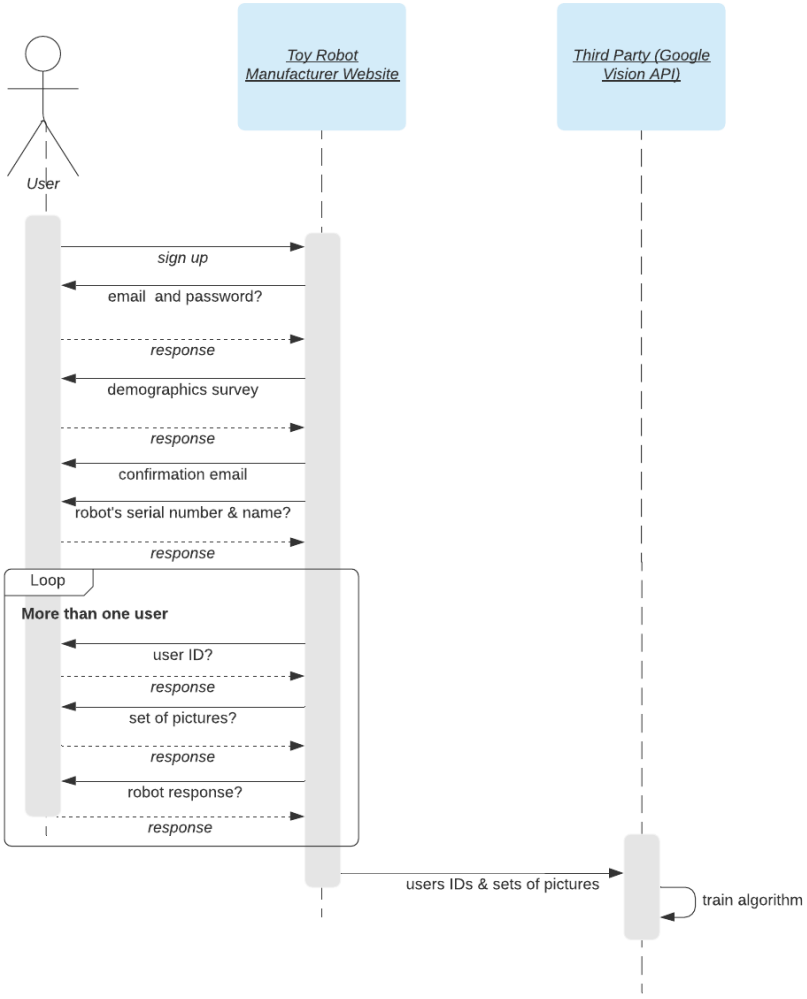

Figure 1. Sequence diagram of traditional toy robot setup

hereinafter is on the process of obtaining consent, it is important to bear in mind that the data subject has the right to withdraw consent. The data subject must be informed thereof prior to giving consent and withdrawing consent should be as simple as giving it [8]. Therefore, a form for withdrawing consent should be available for the user and must be at least as intuitive and simple as the method that is used for providing consent.

Consent should be Freely Given

The other formal requirements of consent are less clear-cut than the timing of consent notices. Following Recital 42 of the GDPR, consent will not be regarded as freely given when the data subject is unable to refuse or withdraw consent without detriment. In order to demonstrate that consent is freely given, the controller may, for example, show that the withdrawal of consent does not lead to a downgrading of the performance of the service to the user's disadvantage. To what extent a downgrading of the performance must be taken into account if the processing occurs in a more privacy-friendly manner, can however only be determined on a case-by-case basis [11]. Moreover, consent is presumed not to be freely given if it is bundled with the acceptance of terms and conditions, or if the performance of a contract is tied to a request for consent to the processing of personal data that is not necessary for the performance of that contract [12][13]. The consent request should therefore be obtained separately from the general terms and conditions [14][15].

#	Data item	Type	Location	Comments
1	User Email Address	String (UTF-8)	DC	The email address is recorded by the DC to enable the user to log in to its portal.
2	Robot Password	String (UTF-8)	DC	The password is recorded by the DC to enable the user to log in to its portal.
3	Demographic Survey	String (UTF-8)	DC	The demographic data is recorded by the DC.
4	Confirmation	Boolean	DC	The confirmation is recorded by the DC to complete the sign-up process (safeguarding from spamming).
5	Serial Number	String (UTF-8)	Local + DC	The serial number is recorded by the DC to link a specific robot to a specific user profile in the portal.
6	Robot Name	String (UTF-8)	DC	The robot name is recorded by the DC for user experience purposes.
7	Home Network Credentials	2 Strings (UTF-8)	Local	The home network credentials are stored locally on the robot.
8	Pseudo-Identifier (per Data Subject)	String (UTF-8)	Local + DC + RS	The DS Identifier is used across all system components.
9	Training Images (per DS)	jpg Files	DC + RS	The training images are used to train the third-party facial recognition service and are stored by the DC to establish its in-house facial recognition service.
10	Robot Response (per DS)	Any (e.g., mp3)	Local + DC	The desired response of the robot to each recognized individual is stored locally and by the DC.
11	Recorded Picture (every five seconds)	jpg File	Local + DC + RS	The recorded images are sent to the third-party facial recognition service and are stored by the DC to establish and improve its in-house facial recognition service.
12	Usage Report (every day)	Any (e.g., Strings)	Local + DC	The daily usage report is used for improvements to the device and service and for further monetary exploitation.

Table 1. Overview of data processed by a toy robot designed with a traditional software engineering approach where data location is *Local* (white), *Data Controller, DC* (red), or *Remote Service, RS* (blue).

Consent should be Specific and Unambiguous

When a service involves different processing operations for multiple purposes, the data subject should have the possibility to express its consent separately for each data processing operation and should not be forced to agree to consent in an all-or-nothing decision [14]. Separating purposes and obtaining consent for each purpose is often referred to as granularity and it is closely aligned with the requirement to obtain specific consent [8], since the specificity criterion prohibits blanket consent and demands the granting of consent linked to a particular processing purpose [16][17]. To provide granularity in consent requests, the options for consent should be distinctive and thereby allow separate consent for different purposes and types of processing [18][8]. This can be accomplished through separate forms or by ticking individual check-boxes [14]. These check-boxes must not be pre-ticked, as consent needs to be unambiguous, i.e. given through an active motion or declaration by the user [11].

Consent should be Informed

In order to ensure that laypersons can understand what they are consenting to, the information provided within a consent form should be presented in an intelligible and easily accessible way [8][10] (e.g., no misleading headings may be used), the language used should be clear and plain [8] (i.e., the message should be understandable to an average user), and the language used should be the local language [19] (e.g., by self-localizing via GPS and choosing the local language of the region). To meet these requirements, an overall balance has to be achieved between the need for clear and plain language, conciseness and the degree of specificity of the information. The objective is to lower the risk that consent forms might be overloaded with information in order to meet the transparency requirements of the GDPR [14].

Consent should be Explicit

The processing of special categories of data, such as biometric data, requires not only unambiguous but explicit consent as a lawfulness ground. While Recital 52 of the GDPR states

that the 'processing of photographs should not systematically be considered to be processing of special categories of personal data,' as soon as the photographs are 'processed through a specific technical means allowing the unique identification or authentication of a natural person' they become sensitive data. In our use case, facial recognition technology processes data in order to uniquely identify the respective child. Hence biometric data is being processed [20], requiring explicit consent [21]. As the term indicates, explicit consent means expressed consent (i.e., consent cannot be implied). The data subject must be requested to agree to a particular use of his or her data and the data subject must actively reply to the question in the affirmative [17][21]. The clearest way to ascertain that consent is explicit would be to obtain express confirmation of consent in a written and signed statement [11]. However, since the GDPR does not necessarily require such a statement, it is also possible to gain explicit consent by other means, for instance by having the data subject sign his or her name beneath the statement or tick a box next to it [14].

Further Requirements on Consent

Because the DC bears the burden of proof that valid consent was obtained, the consent process needs to be documented by the DC [8][16] and motivates storing information about the given consent on the premises of the DC [17]. In light of this responsibility, it is often recommended that the DC uses a double opt-in procedure when obtaining consent. Double opt-in requires that users first declare consent, e.g., by ticking a checkbox or entering their email address. The data subject then receives a verification hyperlink via email or another electronic messaging service. By following the verification link, the data subject reconfirms consent [16][22]. Scholars have advised to store the declaration of consent together with the name of the data subject or another reliable identifier (email address, etc.) and the time of the consent ('timestamp') [16]. The obligation to provide proof of consent exists for as long as the data processing activity in question persists. Once the processing activity has ended, the proof of consent should not be kept any longer than strictly necessary [8].

With Article 8 of the GDPR requiring specific protection for children with respect to their personal data, parental consent must be obtained if connected devices, such as toys that target children, process children's personal data [22][23]. Where the toy is offered directly to a child, Article 8(1) states that the child's consent is only valid if the child is at least 16 years old. With respect to this age limit, it is important to note that Article 8(1) of the GDPR allows member states to 'provide by law for a lower age for those purposes provided that such lower age is not below 13 years.' Therefore, the age threshold is fragmented throughout the EU and service providers must comply with the different age thresholds of the member states [24]. Even if Article 8 of the GDPR does not demand the controller to verify the age of the child, it might be inevitable in practice, given that mechanisms for age confirmation can easily be circumvented by indicating a false age [16]. If a child gives consent even though he or she is below the (national) age limit to provide valid consent, the respective data processing is unlawful. Consequently, age verification by the controller is implicitly required [11][25]. If the user indicates that he or she is above the respective age threshold, the controller should carry out an appropriate verification process to check if the indicated information is true [11]. Since different age verification mechanisms exist, when choosing a mechanism, the DC should keep in mind that the privacy of children and other users can be put at risk by requiring the collection of additional personal data [26]. Thus, depending on the assessment of the risk of the processing, in some low-risk situations it may be sufficient to require a user to indicate his year of birth or to declare that she is above a certain age [11][27].

Where the child is below the age threshold, Article 8(2) of the GDPR requires that 'the controller shall make reasonable efforts to verify that consent is given or authorised by the holder of parental responsibility over the child, taking into consideration available technology.' What is considered a 'reasonable effort' depends on the case, taking into account the risk of the processing and the available technology [25]. Verification of parental responsibility via email might be sufficient in low-risk cases, whereas in high-risk cases the DC should require more evidence

in order to verify and retain the information in accordance with Article 7(1) of the GDPR [11]. In this respect various measures can be taken [28], e.g., transmission of a document signed by the holder of parental responsibility or provision of a copy of a parent's government-issued identity card [29]. Although the DC decides what measures may be appropriate in a particular case, the DC must avoid that the chosen verification technology leads to an excessive collection of personal data [11].

Data-Privacy-Compliance-First Design of a Toy Robot

To be able to design and implement a toy robot that considers data protection law, the identified data flows from the previous analysis are augmented with a legal layer that guides the way the data is processed to comply with GDPR. Figure 2 shows an overview of the toy robot setup that is designed to comply with the data protection requirements.

The user first unpacks the robot, starts its controller, and connects it to his or her computer. At start-up, the robot activates its on-board GPS module and acquires its physical location, i.e., a GPS coordinate. From this location, the robot computes its current logical location (i.e., a country code). The determined country code is stored on the device. The robot furthermore launches a local Web server. At this point, the robot does not have its own Internet connection, thus the user is prompted for the credentials of the user's home wireless network (i.e., SSID and password) so the robot can communicate with remote services. Next, the user directs a Web browser to the (fixed) URL that is connected to the local Web server to access the robot's welcome and configuration dashboard.

Through the configuration dashboard, the user accesses the consent forms—consent is timely —which are made available in the local language by default (based on the determined country code). The system also informs users that they may withdraw consent or modify their consent settings at any time via the robot's configuration page—consent is informed.— When entering this configuration page for the first time, the user is prompted to change the system's password to ensure that the locally stored data is secure. Next, the user is prompted to enter pseudo-identifiers of the data subjects (i.e., of the individuals that shall be recognized by the system). The user is encouraged to enter pseudonyms or nicknames instead of the real names of the data subjects. Then, for each named data subject the following **consent process** is executed:

- The subject is prompted to enter its date of birth; the user is instructed that the real date of birth is required. The parental consent requirement is then based on the entered date of birth and the local jurisdiction regarding the legal protection age of the data subject
- An additional form is displayed later in the process if parental consent needs to be obtained.
- The user provides facial recognition consent, stating that he or she is informed about the purpose of data processing—consent is specific and unambiguous.— The toy robot data processing is restricted to facial recognition capabilities, and other processing capabilities (e.g., audio for voice recognition) would need to be consented to by the data subject through a consent process that is similar to, but separate from, the consent process regarding facial recognition. To ensure freely given consent, the user is furthermore given the choice between three options of how pictures taken by the robot are processed: No processing, offline processing, and online processing. Short statements inform the user about the consequences of selecting either of these options, including that the accuracy of offline processing (i.e., processing on the robot itself) is generally inferior to online facial recognition, and that no processing of data will provide generic educational content that is not tailored to the configured subjects.
- After selecting one of the processing options, the user is prompted again to affirm the online or offline processing of images taken by the robot.
- The robot then prompts the user for his or her email address and sends an email to that address that contains a hyperlink to verify the consent settings. This hyperlink contains the serial number of the robot together with a hash of the data subject's consent choices. The

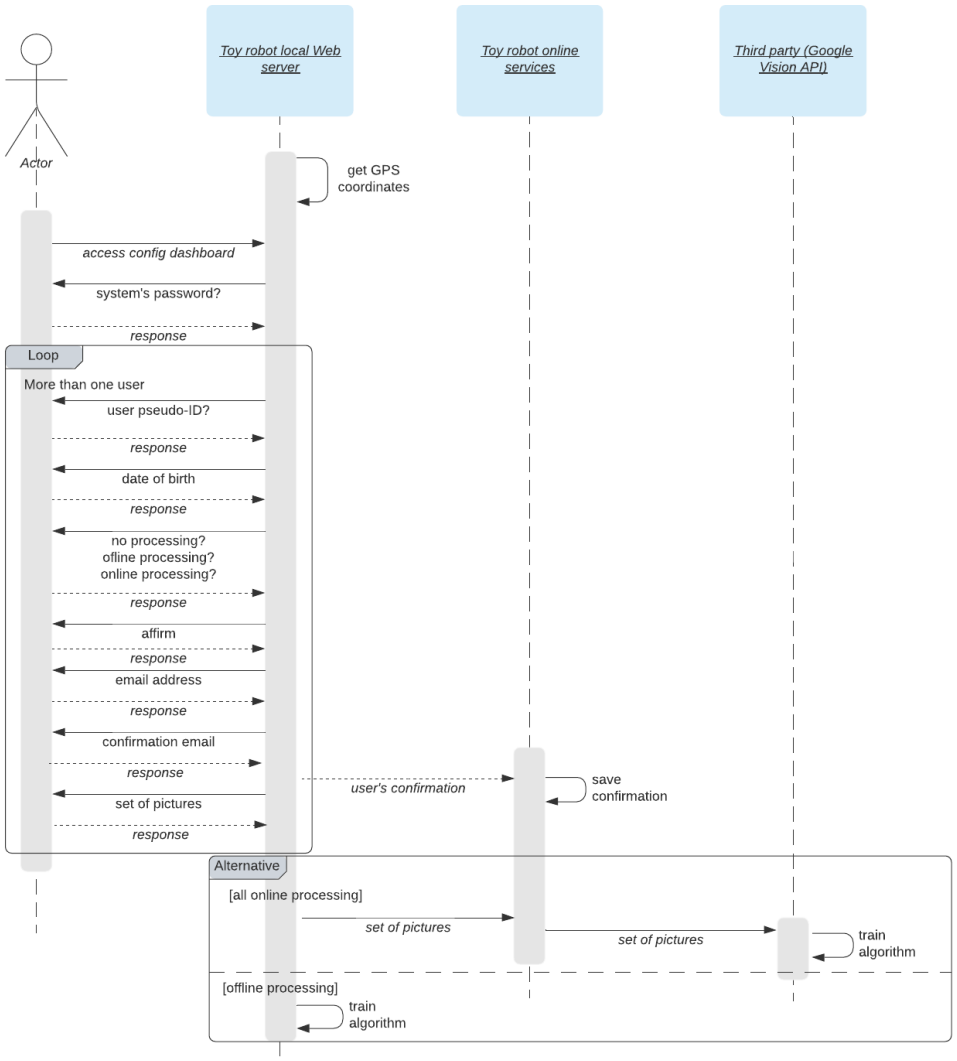

Figure 2. Sequence diagram of a toy robot setup designed under data privacy law

user clicks the received hyperlink to confirm his/her consent settings—consent is explicit. Finally, in this way, user consent settings are stored locally on the robot as well as on the DC's back-end—and the DC can demonstrate consent of the data subject. This concludes the consent process for each data subject

In case a user wants to update their consent settings later on, the previous process needs to be repeated. If the data subject opts for online or offline processing of images, the system finally prompts the user to upload training images to the robot, for each to-be-recognized data subject. Depending on the data subjects' (collective) choices, the robot then either uses these images to train a local classification model or uploads them to an online image classification service. When all data subjects agree to using an online classification service, the user is redirected to the robot's welcome page which displays a 'Start' button.

Once configured, the robot roams the user's premises. In case the user selected offline processing of images, the robot takes a new picture every five seconds and classifies it using the local classification model which returns the classified category (i.e., the name of the data subject that was entered before). If the user selected online processing of images, new pictures are instead submitted to an online classifier which returns the category. In both cases, the robot's response is to perform a personalized action for the identified

individual (e.g., playing a music file). In case the user gave no permission to process recorded images, the robot does not take pictures and does not activate the camera-relevant features of its software and hardware.

Data Privacy Vocabulary

The W3C Data Privacy Vocabulary (DPV)[3] was taken as a foundation to design and implement our toy robot that is compliant with GDPR. DPV provides a machine-readable representation of terms relevant to personal data handling, in adherence to the EU GDPR. This vocabulary is a specification that, if broadly standardized, it could be a strong basis for software engineers and designers to better understand the GDPR, which will allow them, together with legal experts, making implementation decisions that best suit a device's operations while complying with the law. Moreover, since the machine-readable representation of this vocabulary is an ontology, it could be seamlessly linked to other legal regulations; for example geographic, opening the possibility of having a robot that remains compliant even when moving within jurisdictions: On this basis, the robot could for instance autonomously request local restrictions regarding parental consent from a Web service in order to obtain consent from users in a compliant way.

In the same way as Table 1 documents the data flows for the baseline implementation of our robot use case, Table 2 shows the data processed by a version of the toy robot that we implemented following a data-privacy-first approach. For each data item, its storage location (local or remote) is shown along with the mapping of the item to the DPV vocabulary and a comment. In comparison to the robot designed following a traditional approach, the proposed design presented in this section clearly minimizes the variety, amount, and quality of data that leaves the premises of the user: Instead of storing over 90% of the collected types of information (some of which point towards large amounts of data, e.g. 'usage reports') the privacy-by-design robot uploads under 70% of the collected types of information. Moreover, the user is required to explicitly consent to the upload of three types of information and five of them are only collected optionally. Regarding the information types that are processed locally, three of them (#1-3 in Table 2) are used to facilitate the consent process for users and to better secure the user's local data. These three items are furthermore either processed transiently or stored encrypted; the other two (#12 and #15 in Table 2) are optional in this design, although #15 is required to enable the robot to fulfill its application purpose. All the information types that are uploaded to the DC (#4-10 in Table 2), are required by the DC to demonstrate consent. Regarding the information types shared with the third-party service, the derived category (#14) has been decoupled from the data subjects' pseudo-identifiers and the usage of the service has been made optional, by outfitting the robot with an additional local (i.e., offline) facial recognition system. In case of such offline processing, the robot uses the training images to train its algorithms, but does not store them. If the user selects online processing, these images are uploaded to the DC and the remote classification service, but again neither stored locally nor by the DC. Moreover, the data items used to personalize content for a specific user (#11,#13-#14) are always optional. Since, the user can select to play pre-defined content according to the child's age.

[3]https://dpvcg.github.io/dpv/

Conclusion

In this paper, we present the use case of a toy robot's firmware that follows the GDPR data protection principles. This was achieved by carefully analyzing the different data flows needed for the robot to provide its functionalities; in this case, recognizing children to play educational content for them according to their age and specific preferences. This analysis led us to implement variants of the functionality that require more or less data and different levels of data exposure (i.e., online vs. offline processing). This enables in particular granular and specific consent of users to the handling and processing of their data. The consent forms were designed to be easily understood and informative. Moreover, our implementation uses the DPV vocabulary as a foundation to open up the possibility to integrate other regulations, and to enable the automatic compliance with laws when the robot is moved between jurisdictions.

#	Data Item	Type	Location	DPV Relevant Concept	Comments
1	GPS Coordinate	3 Floating-point numbers	Local	dpv:GPSCoordinate	Since GPS is a self-localization system, the robot's location is not learned by any external entity. This data is stored transiently only.
2	Country Code	2 Characters (UTF-8)	Local	dvp:Country	The robot uses a reverse geo-coding algorithm that is implemented locally and therefore does not require any information to leave the device. This data is stored transiently only.
3	Robot Password	String (UTF-8)	Local	No human related→ no GDPR relevance	The robot password is stored locally, encrypted.
4	Pseudo-Identifier (per Data Subject)	String (UTF-8)	Local + DC	dpv:UserName	The DS Identifier is recorded by the DC as part of the consent process.
5	Date of Birth (per DS)	3 Integer numbers	Local + DC	dpv:age	The date of birth of each DS is required to enable the system to comply with its data protection obligations regarding parental consent.
6	Selected Option (per DS)	Short Integer number	Local + DC	dpv:PrivacyPreference	The selected processing option for images is recorded by the DC to enable it to demonstrate consent.
7	Affirmation (per DS)	Boolean	Local + DC	dpv:PrivacyPreference	The affirmation is recorded by the DC to enable it to demonstrate consent.
8	Email Address (per DS)	String (UTF-8)	DC	dvp:EmailAddress	The email address is recorded by the DC to enable it to demonstrate consent and to enable unambiguous consent.
9	Serial Number	String (UTF-8)	Local + DC	dpv:MACAddress	The serial number is recorded by the DC (together with the consent settings) to enable it to demonstrate consent.
10	Confirmation (per DS)	Boolean	DC	dpv:Consent	The confirmation is recorded by the DC to enable it to demonstrate consent.
11	Training Images (per DS)	jpg Files	RS	dpv:Picture	In case of local classification, the training images are used to train the local classifier and are subsequently deleted. In case of the usage of an online service for classification, the images are transmitted to that service.
12	Home Network Credentials	2 Strings (UTF-8)	Local	dpv:Password→ However, it will not leave the robot	The home network credentials are stored locally on the robot and deleted whenever the user consent settings are modified and this information is not anymore required (i.e., when the user chooses a processing option different from online processing).
13	Recorded Picture (every five seconds)	jpg File	RS	dvp:Picture	In case of local processing, recorded pictures are stored locally until they are successfully classified. Then they are deleted. In case of online processing, recorded pictures are transmitted to the remote service.
14	Derived Category (every five seconds)	Integer number	RS	dpv:Identifying	In case the image classification process takes place using a local and locally trained model (our implementation is based on an Inception-v3 Convolutional Neural Network), the classified category is stored locally to trigger the robot response to that individual. In case the image classification takes place using an online service (in our implementation, the Google Vision API is used), the classified category is derived and transmitted by that service. The service thus has access to this information. The category is decoupled from the pseudo-identifier to ensure anonymity.
15	Robot Response (per DS)	Any (e.g., mp3)	Local	dpv:FavoriteMusic	The desired response of the robot to each recognized individual is stored locally.

Table 2. Overview of data processed by a toy robot designed to consider data privacy law where data location is *Local* (white), *Data Controller, DC* (red), or *Remote Service, RS* (blue).

As the toy robot shows, even when minimal data is required to provide a personalized service, such as playing educational content for a kid, it is undoubtedly necessary that designers and software engineers establish an interdisciplinary dialogue with legal professionals to ensure that functionality requirements are appropriately balanced with data privacy requirements. Although this incurs additional effort, our work shows that it can lead to a considerable reduction in the amount of personal data that is shared with the DC and other third parties. Moreover, smart products designed in this manner could flexibly adapt to a user's given consent, which could further reduce the amount of data shared with the DC while preserving essential user experience aspects, albeit at lower quality.

Moreover, the proposed design provides an opportunity to promote transparency. By adopting privacy-by-design and utilizing machine-readable law vocabularies, standardized and understandable documentation could be made available to end users to learn about the way their data is handled.

REFERENCES

1. B. Avak, "Variant management of modular product families in the market phase," Ph.D. dissertation, ETH Zurich, 2006.
2. European Commission, "Statement by executive vice-president margrethe vestager on the launch of a sector inquiry on the consumer internet of things," 16 July 2020. [Online]. Available: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/speech_20_1367/SPEECH_20_1367_EN.pdf
3. S. Gray, P. Sanderson, and K. Ringrose, "Comparison of the proposed 2020 washington privacy act (ssb-6281) to: Gdpr, ccpa, california ballot initiative, and the 2019 wa proposal," 2020. [Online]. Available: https://fpf.org/wp-content/uploads/2020/02/fpf_comparison_of_wa_ssb-6281_to_gdpr__ccpa__cpra__and_2019_version_-_v1.0_feb_12_2020-1.pdf
4. A. Tamò-Larrieux, *Designing for privacy and its legal framework*. Springer, 2018.
5. M. Rueben, C. M. Grimm, F. J. Bernieri, and W. D. Smart, "A taxonomy of privacy constructs for privacy-sensitive robotics." [Online]. Available: https://arxiv.org/pdf/1701.00841
6. E. Fosch-Villaronga, C. Lutz, and A. Tamò-Larrieux, "Gathering expert opinions for social robots' ethical, legal, and societal concerns: Findings from four international workshops," *International Journal of Social Robotics*, vol. 12, no. 2, pp. 441–458, 2020.
7. U. Pagallo, "Robots in the cloud with privacy: A new threat to data protection?" *Computer Law & Security Review*, vol. 29, no. 5, pp. 501–508, 2013.
8. European Data Protection Board, "Guidelines 05/2020 on consent under regulation 2016/679: Version 1.1." [Online]. Available: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf
9. M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "Dark patterns after the gdpr: Scraping consent pop-ups and demonstrating their influence." [Online]. Available: https://arxiv.org/pdf/2001.02479
10. F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A design space for effective privacy notices," in *The Cambridge handbook of consumer privacy*, E. Selinger, J. Polonetsky, and O. Tene, Eds. Cambridge: Cambridge University Press, 2018, pp. 365–393.
11. Article 29 Working Party, "Guidelines on consent under regulation 2016/679': Wp 259 rev.01."
12. E. Kosta, "Article 7. conditions for consent," in *The EU general data protection regulation (GDPR)*, C. Kuner, L. A. Bygrave, L. Drechsler, and C. Docksey, Eds. Oxford: Oxford University Press, 2020.
13. E. Gil González and P. de Hert, "Understanding the legal provisions that allow processing and profiling of personal data—an analysis of gdpr provisions and principles," *ERA Forum*, vol. 19, no. 4, pp. 597–621, 2019.
14. Information Commissioner's Office, "Consent." [Online]. Available: https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/consent-1-0.pdf
15. B. Custers, F. Dechesne, W. Pieters, B. Schermer, and S. van der Hof, "Consent and privacy," in *The Routledge handbook of the ethics of consent*, ser. Routledge handbooks in applied ethics, A. Müller and P. Schaber, Eds. Boca Raton, FL: Routledge an imprint of Taylor and Francis, 2018.
16. S. Dienst, "Lawful processing of personal data in companies under the general data protection regulation," in *New European general data protection regulation*, T. Kugler and D. Rücker, Eds. München and Baden-Baden and Oxford and München: C.H. Beck and Nomos and Hart and Verlag C.H. BECK oHG, 2018.
17. L. A. Bygrave and L. Tosoni, "Article 4(11). consent," in *The EU general data protection regulation (GDPR)*, C. Kuner, L. A. Bygrave, L. Drechsler, and C. Docksey, Eds. Oxford: Oxford University Press, 2020.

18. O. Drozd and S. Kirrane, “I agree: Customize your personal data processing with the core user interface,” in *Trust, Privacy and Security in Digital Business*, ser. Security and Cryptology, S. Gritzalis, E. R. Weippl, S. K. Katsikas, G. Anderst-Kotsis, A. M. Tjoa, and I. Khalil, Eds. Cham: Springer International Publishing and Imprint: Springer, 2019, pp. 17–32.
19. J. Taeger, “Art. 7 Bedingung für die Einwilligung,” in *DS-GVO - BDSG*, ser. Kommunikation & Recht, J. Taeger and D. Gabel, Eds. Frankfurt am Main: Fachmedien Recht und Wirtschaft dfv Mediengruppe, 2019.
20. Y. Welinder and A. Palmer, “Face recognition, real-time identification, and beyond,” in *The Cambridge handbook of consumer privacy*, E. Selinger, J. Polonetsky, and O. Tene, Eds. Cambridge: Cambridge University Press, 2018, pp. 102–124.
21. L. Georgieva and C. Kuner, “Article 9. processing of special categories of personal data,” in *The EU general data protection regulation (GDPR)*, C. Kuner, L. A. Bygrave, L. Drechsler, and C. Docksey, Eds. Oxford: Oxford University Press, 2020.
22. P. Voigt and A. von dem Bussche, *The EU General Data Protection Regulation (GDPR): A practical guide*. Cham, Switzerland: Springer, 2017. [Online]. Available: http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=029688307&sequence=000002&line_number=0002&func_code=DB_RECORDS&service_type=MEDIA
23. D. Kelleher and K. Murray, *EU data protection law*. London: Bloomsbury Professional, 2018.
24. Information Commissioner’s Office, “Children and the gdpr.” [Online]. Available: https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/children-and-the-gdpr-1-0.pdf
25. E. Kosta, “Article 8. conditions applicable to child’s consent in relation to information society services,” in *The EU general data protection regulation (GDPR)*, C. Kuner, L. A. Bygrave, L. Drechsler, and C. Docksey, Eds. Oxford: Oxford University Press, 2020.
26. Unicef, “Children’s online privacy and freedom of expression.” [Online]. Available: https://issuu.com/unicefusa/docs/unicef_toolkit_privacy_expression?e=29613278/60947364
27. Information Commissioner’s Office, “Age appropriate design: a code of practice for online services.” [Online]. Available: https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/age-appropriate-design-a-code-of-practice-for-online-services-0-0.pdf
28. Centre for Information Policy Leadership, “Gdpr implementation in respect of children’s data and consent.” [Online]. Available: https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_-_gdpr_implementation_in_respect_of_childrens_data_and_consent.pdf
29. R. Kazemi, *Die EU-Datenschutz-Grundverordnung in der anwaltlichen Beratungspraxis*, ser. AnwaltsPraxis. Bonn: Deutscher Anwaltverlag & Institut der Anwaltschaft GmbH, 2017.